A Prototyping Method to Simulate Wearable Augmented Reality Interaction in a Virtual Environment - A Pilot Study

Günter Alce[1], Klas Hermodsson[1], Mattias Wallergård[2], Lars Thern[2], Tarik Hadzovic[2]
[1]Sony Mobile Communications
Mobile vägen 4, Lund, Sweden 221 88
Gunter.Alce@sonymobile.com; Klas.Hermodsson@sonymobile.com
[2]Lund University, Department of Design Sciences
P.O. Box 118, Lund, Sweden 221 00
Mattias.Wallergard@design.lth.se; Lars.Thern@gmail.com; Hadzovic.Tarik@gmail.com

Abstract- *Recently, we have seen an intensified development of head mounted displays (HMD). Some observers believe that the HMD form factor facilitates Augmented Reality (AR) technology, a technology that mixes virtual content with the users' view of the world around them. One of many interesting use cases that illustrate this is a smart home in which a user can interact with consumer electronic devices through a wearable AR system. Building prototypes of such wearable AR systems can be difficult and costly, since it involves a number of different devices and systems with varying technological readiness level. The ideal prototyping method for this should offer high fidelity at a relatively low cost and the ability to simulate a wide range of wearable AR use cases.*

This paper presents a proposed method, called IVAR (Immersive Virtual AR), for prototyping wearable AR interaction in a virtual environment (VE). IVAR was developed in an iterative design process that resulted in a testable setup in terms of hardware and software. Additionally, a basic pilot experiment was conducted to explore what it means to collect quantitative and qualitative data with the proposed prototyping method.

The main contribution is that IVAR shows potential to become a useful wearable AR prototyping method, but that several challenges remain before meaningful data can be produced in controlled experiments. In particular, tracking technology needs to improve, both with regards to intrusiveness and precision.

Keywords: Augmented Reality, Virtual Reality, Prototyping, User Interaction.

1. Introduction

Recently, we have seen an intensified development of head mounted displays (HMD), i.e. display devices worn on the head or as part of a helmet. Two of the most well-known examples are Google Glass [1] and Oculus Rift [2]. The HMD form factor facilitates Augmented Reality (AR), a technology that mixes virtual content with the users' view of the world around them [3]. Azuma [4] defines AR as having three characteristics: 1) Combines real and virtual, 2) interactive in real time and 3) registered in 3-D. One of many interesting use cases that illustrate this is a smart home in which a user can interact with consumer electronic devices through a wearable AR system. For example, such an AR system could help the user discover devices, explore their capabilities and directly control them.

Building prototypes of such wearable AR systems can be difficult and costly, since it involves a number of different devices and systems with varying technological readiness level. The ideal prototyping method for this should offer high fidelity at a relatively low cost and the ability to simulate a wide range of wearable AR use cases. Creating a prototyping method which fulfils these requirements is problematic due to underdeveloped or partially developed technology components, such as display technology and object

tracking. Also the lack of development tools and methodologies is a hindrance [5]. In particular, it is difficult to achieve prototypes that offer an integrated user experience and show the full potential of interaction concepts.

There are numerous examples of prototyping methodologies and tools used for prototyping AR interaction. Some offer low fidelity at low cost, e.g. low fidelity mock-ups [6] and bodystorming [7], whereas some offer high fidelity at high cost e.g. a "military grade" virtual reality (VR) system [8]. In between these two extremes there is a huge variety of prototyping methods. For example, a prototyping method widely used within human-computer interaction is Wizard of Oz (WOZ), in which a human operator simulates undeveloped components of a system in order to achieve a reasonable level of fidelity.

This paper presents a proposed method for prototyping wearable AR interaction in a virtual environment (VE). From here on we refer to the method as IVAR (Immersive Virtual AR). IVAR was developed in an iterative design process that resulted in an adequate setup in terms of hardware and software. Additionally, a small pilot study was conducted to explore the feasibility of collecting quantitative and qualitative data from the proposed method.

The main contribution of this paper is to present IVAR, a method for exploring the design space of AR interaction using a VE.

In the next section relevant related work is presented. The method is described in the section called the IVAR method, which is followed by pilot experiment, results, discussion and conclusions.

2. Related Work

Prototyping is a crucial activity when developing interactive systems. Examples of methods which can be used for prototyping AR systems include low-fidelity prototyping, bodystorming and WOZ.

Each method has its advantages and disadvantages. For example, low fidelity prototyping such as paper prototyping can be very effective in testing issues of aesthetics and standard graphical UIs. However, to do so when designing for wearable AR interaction, a higher fidelity is likely to be required [9].

As already mentioned, a well-known prototyping method widely used within human-computer interaction is WOZ. For example, the WOZ tool WozARd [10] was developed with wearable AR interaction in

mind. Some advantages of using WozARd are flexibility, mobility and the ability to combine different devices.

Carter et al. [9] states that WOZ prototypes are excellent for early lab studies but they do not scale to longitudinal deployment because of the labor commitment for human-in-the-loop systems.

Furthermore, WOZ relies on a well-trained person, the wizard, controlling the prototyping system during the experiment. The skill of the wizard is often a general problem with the WOZ method since it relies on the wizard not making any mistakes [11].

VE technology has been used as a design tool in many different domains, such as architecture, city planning and industrial design [12]. A general benefit of using a VE to build prototypes of interactive systems is that it allows researchers to test systems or hardware that do not actually exist in a controlled manner. Another advantage, compared to the WOZ method, is that functionality in the prototype can be handled by the VE, instead of relying on an experienced human operator to simulate the technical system's behaviour. One of the main drawbacks when it comes to simulating AR interaction in a VE is related to the fidelity of the real world component in the system, i.e. the simulated real environment and objects upon which augmented information is placed [13]. Examples of issues include the lack of tactile feedback in the VE and physical constraints due to limitations of VE navigation compared to real world movements. However, these issues depend on the goal of the study and are likely to be less of an issue for some use cases.

Ragan et al. [13] used VEs to simulate AR systems for the purposes of experimentation and usability evaluation. Their study focused on how task performance is affected by registration error and not on user interaction. They conducted the study using a four-sided CAVE™ with an Intersense IS-900 tracking system for head and hand tracking and their setup can be considered to be a high cost system.

Baricevic et al. [14] present a user study, evaluating the benefits of geometrically correct user-perspective rendering using an AR magic lens simulated in a VE. Similar to Whack-A-Mole, the participants were asked to repeatedly touch a virtual target using two types of magic lenses, phone-sized and tablet-sized. The focus of the study was on which rendering method the participants preferred and not on the user interaction.

Lee et al. [15] used a high-fidelity VR display system to achieve both controlled and repeatable mixed reality simulations of other displays and environments.

Their study showed that the completion times of the same task in simulated mode and real mode are not significantly different. The tasks consisted of finding virtual objects and reading information.

Despite these research efforts, there are no AR prototyping tools, to the knowledge of the authors of this paper, that focuses on using a VE to prototype interaction concepts for wearable AR.

3. The IVAR Method

We reason that a method for prototyping wearable AR interaction in a VE should have the following characteristics in order to be effective:

- It should offer a degree of immersion, i.e. "*the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding, and vivid illusion of reality*" [16]. Immersion is believed to give rise to presence a construct that can be defined by [17]:
  - The sense of "being there" in the environment depicted by the VE.
  - The extent to which the VE becomes the dominant one, that is, that the participant will tend to respond to events in the VE rather than in the "real world."
  - The extent to which participants, after the VE experience, remember it as having visited a "place" rather than just having seen images generated by a computer.

  It has been suggested that with a higher degree of presence a VR system user is more likely to behave as he/she would have done in the corresponding real environment [18]. This is very important in order to achieve realistic user behavior and accordingly meaningful data with the IVAR method.

- The interaction with the VE (navigation and manipulation of objects) should be intuitive and comfortable. 3D interaction is difficult and users often find it hard to understand 3D environments and to perform actions in them [19]. Part of the problem is that VEs lack much of the information we are used to from the real world. According to Bowman [20], "the physical world contains many more cues for understanding and constraints and affordances for action that cannot currently be represented accurately in a computer simulation". Difficult and awkward VE interaction would bias the results of prototype evaluations, making it very difficult to say if user performance is due to

the VE interaction or due to the prototyped AR interaction.

With the fast development of off-the-shelf technology it has become increasingly easier to build VR systems that live up to the two characteristics described above. For example, low price HMDs for VR are now sold by a number of companies. The development of input devices for VR is even faster and has produced products such as Microsoft Kinect [21], Razer Hydra [22] and Leap Motion [23], which offer different opportunities for tracking user input depending on the underlying technology.

To build a setup for the IVAR method, a number of off-the-shelf input/output devices (available in late 2013 when this study was initiated) were tested in different configurations and their advantages and disadvantages were reflected upon. Through this exploratory development the final IVAR system components were chosen (Figure 1). An important design guideline during the development was to map the physical world to the VE in an attempt to create higher immersion and more intuitive VE interaction (Figure 3).

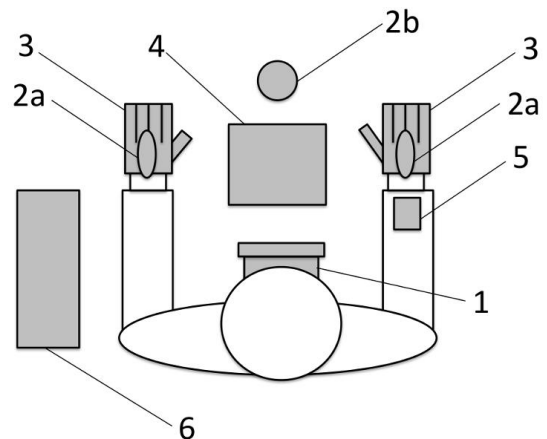The total effort spent in designing, implementing and testing the set-up was 45 person weeks.

Figure 1. IVAR system components, 1) Oculus Rift [2], 2) Razer Hydra [22], 3) 5DT Data Glove 5 Ultra [24], 4) Sony Xperia Tablet Z [25], 5) Smartphone, 6) Desktop computer.

1. Oculus Rift Development Kit [2]. A VR HMD, having head tracking in three rotational degrees of freedom (DOF) and approximately 100^0 field of view.
2. Razer Hydra [22]. A game controller system that tracks the position and orientation of the two wired controllers (2a) in six DOF relative to the

base station (2b) through the use of magnetic motion sensing. The two controllers are attached to the back of the user's hands. This enables the system to track location and orientation of the user's hands, and in extension via inverse kinematics the flexion and rotation of the user's arms.

3. 5DT Data Glove 5 Ultra [24]. A motion capture glove, which tracks finger joint flexion in real time.

4. Sony Xperia Tablet Z [25]. An Android powered 10'' tablet. The tablet is placed on a table in front of the user in alignment with the location of a tablet in the VE (Figure 2). The tablet allows the system to capture and react to touch input from the user. Additionally, it offers tactile feedback, which is likely to result in higher immersion and more intuitive VE interaction.

5. Android powered smartphone. This device is attached to the wrist of the user's dominant arm and is used to simulate a wristband that gives vibration feedback. The location of this feedback device is aligned with the virtual wristband in the VE (Figure 2).

6. Desktop computer with a powerful graphics card. This computer executes and powers the VE through the use of the Unity [26] game engine. The computer is powerful enough to render the VE at a high and steady frame rate.

Figure 2. The VE.

Most of the IVAR system components are wired, making this setup unsuitable for interaction where the user needs to stand up and walk around. However, the setup works for use cases that involve a seated user. For this reason, it was decided to implement a VE based on a smart living room scenario in which a user sitting in a sofa can interact with a set of consumer electronic devices (Figure 2). A tablet device is lying on the table in front of the user. The tablet is one of the input devices that can be used to control the living room. A TV is hanging on the wall in front of the user where media can be played back. Other consumer electronic devices including speakers, audio receiver, game console and printer, are located throughout the living room. The user is presented with overlaid information as spatially placed, visual feedback while interacting with the VE. In a real world scenario, this augmentation could originate from an optical see-through HMD, ceiling mounted projector or similar.

It has been suggested that physical awareness is important for effective interaction [27] in a VE. Therefore, some of the furniture and devices in the VE were mapped with corresponding objects in the real world (Figure 3). For instance, 1) The virtual table was mapped with a physical table, 2) The virtual tablet was mapped with a physical tablet, and 3) the virtual smartband was mapped with a smartphone (Figure 3). The physical table and tablet provided tactile feedback and the smartphone provided haptic feedback.
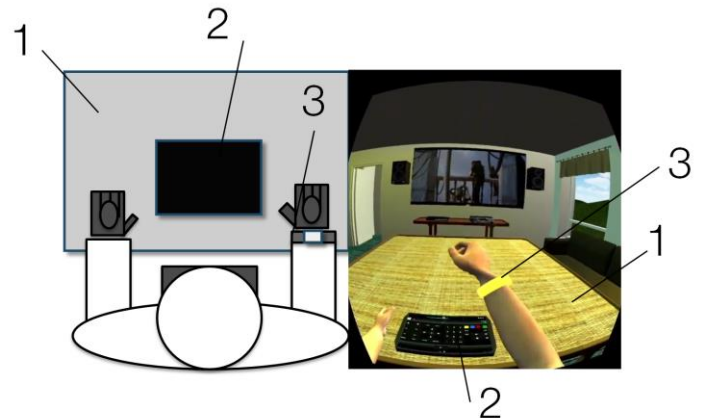

Figure 3. The physical devices mapped with virtual devices, 1) table, 2) tablet, 3) smartband.

In order to facilitate immersion, ease of interaction and physical awareness, the VE was equipped with a virtual representation of the user's own body. The virtual body was based on a 3D model of a young male whose two virtual arms could be moved in six DOF by the user. The ten virtual fingers could be moved with one rotational DOF. This was the most realistic input that could be achieved with the available tracking devices.

Four well-known interaction concepts with relevance for wearable AR were implemented in the VE. The concepts support two tasks that can be considered fundamental for a smart living room scenario: device discovery and device interaction, see video[1]. The purpose of the interaction concepts was to use them for starting exploring what it means to collect quantitative and qualitative data with a prototyping method like IVAR. The four interaction concepts and their usability in relation to wearable AR are not the focus of this paper but are instead described and reflected upon in more detail in [28].

Device discovery - gesture: This concept allows the user to discover the identity and location of consumer electronic devices by moving the dominant hand around, as if "scanning" (Figure 4). When a discoverable device is pointed at by the user's hand for more than one second, the user receives vibration feedback from his/her wristband and a window with information of the device is displayed. When the user is no longer aiming at the device with his hand, the window disappears.
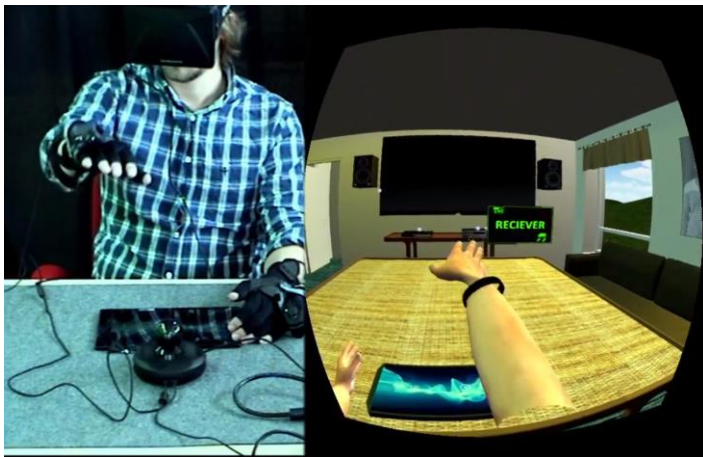

Figure 4. Device discovery – gesture.

Device discovery - gaze: This concept works like gesture except that the user is using his gaze instead of his hand movements to discover devices. Another difference is that the user does not receive vibration feedback when the information window appears. The hardware system does not include a gaze tracking component, instead the centre of the display view is used to approximate the focus of the user's gaze.

Device interaction - grab: This concept allows the user to select a playback device with a grabbing gesture. The user first selects an output device by reaching towards it with an open hand and then clenching the fist. If a device was correctly selected, the virtual wristband gives vibration feedback and also changes colour from black to yellow. The device remains selected as long as the user keeps making a fist. The user can then place the hand above the tablet and unclench the fist, which makes the tablet render a UI for media playback.

Device interaction - push: This concept allows the user to first select media content on the tablet and then select an output device by making a flick gesture towards it (Figure 5). The user then interacts with the tablet UI to control the playback of the content.

Figure 5. Device interaction – push.

4. Pilot Experiment

In an attempt to create an initial understanding of what it means to evaluate wearable AR interaction with the IVAR method, a basic pilot experiment was conducted. The purpose of the pilot experiment was to let a group of participants carry out tasks with the four interaction concepts while collecting qualitative and quantitative data about their performance.

4. 1. Setup

The pilot experiment was conducted in a usability lab with audio and video capturing capabilities. The sessions involved a participant, a test leader and a test assistant (Figure 6).

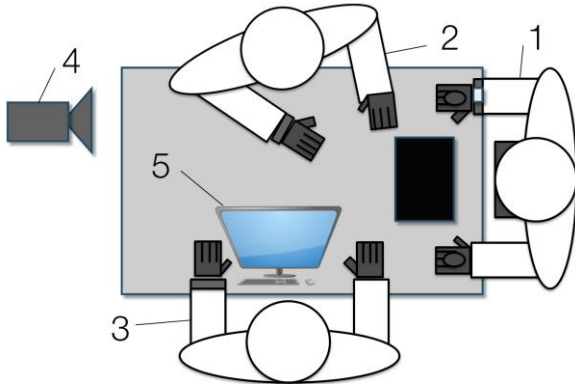[1] Concepts at youtube: http://goo.gl/zewYjm

Figure 6. The experimental setup. 1) Participant, 2) Test leader, 3) Test assistant, 4) Video camera, 5) Desktop computer running the VE.

4. 2. Procedure

First, the participant was informed about the purpose of the experiment and signed an informed consent form. The participant also filled in a background questionnaire, which included age, gender, occupation and 3D gaming experience.

Next, the participant was given a brief introduction to the idea of future types of interaction beyond the desktop paradigm. Thereafter, after the VR equipment had been put on the participant, he/she was given approximately five minutes to get acquainted with the VR system and the living room VE (Figure 2).

The participant was then informed about the task to perform in the living room VE. The task was to 1) find a device for video playback and then 2) start playback of a film on that device. In the first part of the task, the participant used the two device discovery concepts, i.e. gesture and gaze. In the second part, the two device interaction concepts, i.e. grab and push, were used by the participant. In other words, each participant tried all four interaction concepts. The order in which the concepts were tested was counterbalanced in order to address learning effects. During each test, task completion time, performed errors and error recovery time were recorded and logged by the system. Furthermore, multiple audio/video feeds and one video feed from the VE were captured.

After each interaction concept, the participant filled out a NASA-TLX questionnaire [29]. NASA-TLX is commonly used [30] to evaluate perceived workload for a specific task. It uses an ordinal scale on six subscales (Mental Demand, Physical Demand, Temporal Demand, Performance, Effort and Frustration) ranging 1 - 20.

The questionnaire was followed by a short semi structured interview.

When the participant had completed all four interaction concepts, he/she was asked questions about perceived differences between the concepts for device discovery and device interaction respectively.

The participant's comments from the experiment were transcribed and analyzed. Total perceived workload was calculated for each participant based on the NASA-TLX data. A Wilcoxon signed rank test for two paired samples ($p < 0.05$) was used to find eventual differences between interaction concepts with regards to the NASA-TLX data.

4. 3. Participants

Participants were enrolled among university students; the majority of them studied engineering. 24 persons (9 female) with mean age 24.5 (19 - 37 years) participated in the device discovery part. Only 20 of these participants (9 female) performed the device interaction part of the experiment due to technical problems.

5. Results

This section presents quantitative and qualitative data from the pilot experiment. Overall, all participants managed to complete the tasks and the majority of them showed signs of enjoyment. The following data is divided in to three sections, device discovery, device interaction and possible sources of error.

5. 1. Device Discovery

All participants managed to perform the device discovery part of the experiment.

5. 1. 1. Quantitative Data

The NASA-TLX data is presented in Table 1 and Figure 7, respectively. The statistical significant differences are marked with an asterisk.

As for task completion time, no statistically significant difference between gesture and gaze could be found (median gesture = 28 and median gaze = 26.5 with range (8 - 142) and (6 - 93) respectively), ($Z = -0.51, p < 0.05$).

Table 2 presents the total number of errors and the average recovery time for each device discovery concept. An error was defined as a faulty action made by the participant. E.g. if the participant was asked to play a video on the TV but instead chose the PlayStation it was considered a fault. The recovery time was defined as the time from the faulty action until the correct action was initiated. No statistically significant

differences between the two concepts with regards to number of errors and average recovery time could be found (Table 2).

Table 1. Device discovery: Median and Z-value for NASA-TLX values. Perceived workload on top followed by the six sub scales.

Device Discovery	Gesture	Gaze	Z-value
Workload	5.13	3.03	3.40*
Mental Demand	5.00	2.50	1.32
Physical Demand	5.50	1.50	4.07*
Temporal Demand	3.00	3.00	1.16
Performance	5.00	3.00	2.23*
Effort	5.00	2.50	2.81*
Frustration	4.00	2.50	-2.50*

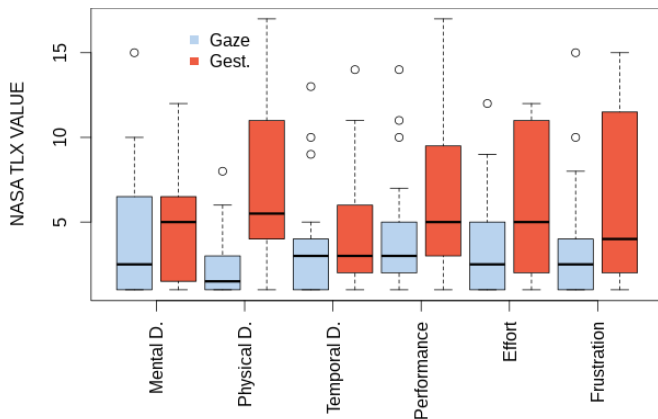*   Significant at the 0.05 probability level

Figure 7. Boxplots for NASA-TLX data, gaze and gesture.

Table 2. Number of errors and average recovery time for the device discovery concepts.

Device discovery	Number of errors	Average recovery time
Gesture	16, $SD = 1.16$	15.06, $SD = 8.21$
Gaze	32, $SD = 1.63$	6.13, $SD = 3.25$

5. 1. 2. Qualitative Data

In general, participants tended to describe gaze as "natural" and "intuitive".

Several participants commented that the augmented information appearing on the discoverable consumer electronic devices might become intrusive. One participant commented that ''When the cards kept appearing I got overwhelmed'' and another related that ''It can be annoying with too much information, popping up all the time''.

Several participants stated that they felt more comfortable and in control with the gesture concept.

Two illustrative comments were: ''It is comparable to point towards objects one is interested in'' and ''I feel more comfortable and secure, I am more in control''. Two negative comments about the gesture concept were: ''Stuff like this... look cool in movies but if it was my living room I would not want to move too much... the hand moving is a lot to ask'' and ''Feels weird, not a fan''.

5. 2. Device Interaction

All participants managed to finish the two device interaction tasks. However, the data of four of the participants could not be used due to technical problems.

5. 2. 1. Quantitative Data

The NASA-TLX data is presented in Figure 8. No significant differences between the two device interaction concepts could be found in the NASA-TLX data. Nevertheless, there was a statistically significant difference between grab and push with regards to task completion time (grab = 28.20, push = 12.88 and median 18 and 10.75 respectively, with a range (8 - 75) and (6 - 28.5) respectively), ($Z = 2.39, p < 0.05$).

No statistically significant results were found for number of errors and average recovery time (Table 3). The definition of the errors and the recovery time has been explained in the device discovery section (Section 5.1.1).

Table 3. Number of errors and recovery time by the participants for the device interaction concepts.

Device Interaction	Number of errors	Average recovery time
Grab	4, $SD = 0.70$	8.88 , $SD = 2.51$
Push	6, SD = 0.66	4.33, $SD = 1.99$

5. 2. 2. Qualitative Data

Overall, the participants seemed to find the push concept easy and convenient to use. Two comments were: ''A very convenient way to select and start'' and ''I think it is comfortable, it is intuitive to send things in the right direction''.

Nevertheless, some of the participants seemed to find the push concept awkward and felt insecure about where the pushed content would end up. One participant reasoned that ''I think it is more logical to start with the (output) device''. Another participant imagined a scenario with a large amount of consumer

electronic devices: ''When having many devices I would not have felt comfortable to send things like this''.
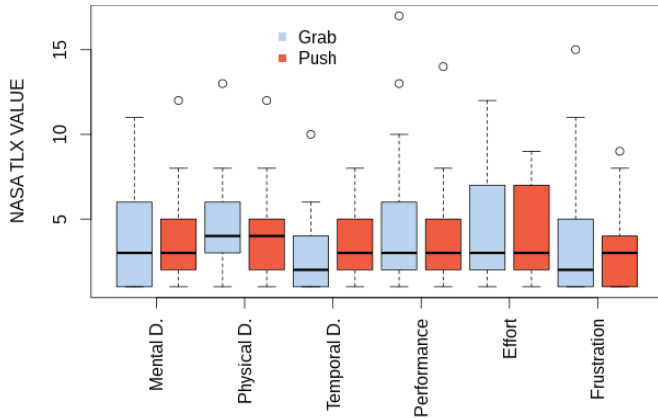

Figure 8. Boxplots for NASA-TLX data, grab and push.

Overall, the participants seemed to find the grab concept intuitive. Two comments that illustrate this were: ''Cool idea and it felt very intuitive'' and ''If you just want to watch TV then just grab''. However, also negative comments appeared: ''It felt somewhat unnatural... you want to feel what you have grabbed in the hand, here you grab in the thin air'' and ''This is not as easy as push''.

5. 3. Possible Sources of Error

In this section, some possible sources of errors that could have influenced the pilot experiment data are presented.

Gloves. The Razer Hydra controllers were attached to the participant's hands by fastening them on thin gloves on top of the 5DT Data Gloves. These gloves were only available in one size, which created problems for participants with smaller hands. The problem was that the Razer Hydra controllers could move a bit in relation to the hand, which had a negative effect on the tracking accuracy for these participants. The problem mainly appeared for Gesture and Grab but also to some extent for Push.

Cables. The VR system setup included a number of cables. The cables of the 5DT Data Gloves and the Razer Hydra controllers were grouped together on the participant's arm whereas the cable of the Oculus Rift HMD went behind the participant's back. It could be observed that the cables to some extent restricted the participant's movements. The problem was particularly apparent for participants with smaller hands since the weight of the cables attached to the participant's arm

induced even larger movements of the Razer Hydra controller in relation to the hand.

6. Discussion

In this paper we have presented IVAR, a method for prototyping wearable AR interaction concepts in a VE. A basic pilot experiment was performed to create an initial understanding of what it means to evaluate wearable AR interaction with this type of method.

Overall, the results from the pilot experiment can be described as mixed. On the one hand, the participants in general managed to solve the tasks in the VE. Furthermore, the qualitative data suggests that the participants had reached a level of understanding for the interaction concepts, even though they were not familiar with either the concepts or the medium, to be able express preference and evaluative statements. On the other hand, it is very difficult to say something about the effectiveness of IVAR based on the quantitative data. Some very rough tendencies could be observed in the NASA-TLX data, especially when comparing gesture and gaze. This could eventually be interpreted as a result of these two interaction concepts being quite different to their nature. However, it was apparent that some participants experienced problems during the experiment, inducing a considerable effect on task performance time, number of errors and recovery time. In some cases this was clearly due to the somewhat error prone tracking. Also, some participants seemed to occasionally loose track of the experimental tasks and appeared more interested in enjoying the VE and the VR experience.

What it boils down to is that the validity of a method based on participants' perceptions and actions inside a VE must be carefully considered. One could argue that the proposed method constitutes a sort of Russian nested doll effect with ''a UI inside a UI''. This raises the question: are observed usability problems caused by the UI or by the VR technology, or by both? Providing a definitive answer to this question is well beyond the scope of this paper, but based on our results the IVAR method appears to have an adequate potential, making continued exploration and development of the method worthwhile for the purpose of prototyping AR interaction. With continued development, methods such as IVAR could have the potential of being usable both for experiments in academic research and prototype development in industrial research and development. However, IVAR would probably be more useful in an industrial

environment since it facilitates relatively fast prototyping of several parallel design concepts, something which is increasingly important in today's agile development.

To accurately evaluate and compare different AR interaction concepts, a VR setup with higher fidelity than the one presented in this paper would be needed. From a validity point of view, it is crucial that participants behave as they would have done in a corresponding real situation and that their intents and actions are translated to the VE in a precise manner. A key element for achieving this is immersion (as described earlier). Examples of factors that contribute to immersion include field of view, resolution, stereoscopy, type of input, latency etc. In the VR system described in this paper a relatively high degree of immersion was achieved with a wide FOV HMD (100 degrees diagonal) and low latency head tracking. Nevertheless, there is still big room for improvement with regards to immersion factors of the VR system. For example, the version of the Oculus Rift used (DK1), lacks the ability to detect translational movement, restricting the user's head movements to three rotational DOF. Head positional tracking can improve depth perception and thus immersion in a VE due to motion parallax [31]. Furthermore, the display resolution of Oculus Rift DK1 is only 640*800 pixels per eye. Low resolution makes especially text hard to read and it is not possible to faithfully reproduce the crispness of a device UI inside the VE. Also, low pixel density and noticeable pixel persistence inhibit visual fidelity and consequently the degree of immersion.

In the case of user input, there is big room for improvements with regards to tracking hands and fingers. The setup used in this paper was relatively cumbersome with several tracking and mobile devices attached to the participant, resulting in a ''tangle'' of cables and straps. This probably had a negative effect both on immersion and the precision by which participants could perform tasks in the VE. This type of setup is therefore likely to introduce variables that significantly interact with what is being studied using the IVAR method, i.e. different wearable AR interaction concepts. It is also important to note that the precision problem most likely affected the data to a lesser extent in the present study since none of the four concepts involved fine motoric movements. Prototyping and comparing different interaction concepts based on e.g. single-finger flicking gestures would require a much more precise tracking system in order to be meaningful.

An alternative setup could consist of Leap Motion's Dragonfly [32] mounted at the front of the Oculus Rift DK2 [33]. Such a VR system would not require the participant to wear any tracking devices on his/her upper limbs, which would lead to much less movement restrictions. It would also facilitate significantly more precise tracking not least by offering several rotational DOFs for each finger. Nevertheless, other problems could appear due to the restricted field of view of the Leap sensor.

The four interaction concepts used in the pilot experiment were for a particular use case that assumes ''interaction from a couch'' with a stationary user. For this use case, the tethered hardware only limits the user's actions to some extent. A use case with more mobility involved would render this VR-based method hard to use. This would either require a portable VR setup or one that allows intuitive VE navigation, e.g. an omnidirectional treadmill. However, this is expensive hardware that can be difficult to use and maintain. Nevertheless, low-cost hardware targeting VR gaming can be expected to appear soon. One example is Virtuix Omni [34] which uses a slippery platform to allow the user to turn, walk or run in place and translate it to similar motion inside a VE.

7. Conclusions

The main conclusion of the research described in this paper is that IVAR shows potential to become a useful prototyping method, but that several challenges remain before meaningful data can be produced in controlled experiments. In particular, tracking technology needs to improve, both with regards to intrusiveness and precision.

8. Future work

To further explore the proposed prototyping method, more controlled experiments should be performed. Most importantly, experiments with a control group are needed in order to learn more about the validity of the method. A possible setup for such an experiment could be to let a control group perform tasks in a real world scenario and compare their performance with that of a group doing the same tasks in a VE.

As has been discussed above, the VR setup described in this paper has big improvement potential in terms of hardware. For example, a new version of the Oculus Rift Development kit, which feature positional tracking, a resolution of 960*1080 per eye and low pixel

persistence was released in the summer of 2014 [33]. An alternative VR HMD is also being developed by Sony under the codename Project Morpheus [35]. An alternative display solution for the proposed prototyping method could be see-through displays, such as Microsoft HoloLens [36]. This opens up the possibility to prototype person-to-person interaction as well as multiple user consumer electronics interaction based on AR.

Furthermore, future VR setups for prototyping wearable AR interaction could exploit the benefits of wireless tracking systems such as the upcoming STEM wireless motion tracking system [37] or Myo [38], a gesture control armband.

Finally, it is important to explore how more quantitative data can be generated with a prototyping method such as IVAR. One way could be to exploit gaze tracking functionality which has started to appear in some HMD models such as SMI eye tracking glasses [39]. The eye movement data could then be used to produce e.g. scan paths and heat maps [40], in order to understand how the participants guide their visual attention. Another idea could be to make use of the movement data generated by the tracking system that controls the virtual representation of the user's own body. The data could be matched with anthropometric databases in order to perform basic ergonomic analyses of e.g. gestures that are part of an AR interaction concept.

9. Acknowledgments

References

[1] (2015, April 10). Google Glass. [Online]. Available: https://developers.google.com/glass/design/ui.

[2] Oculus. (2015, April 29). Oculus Rift - Virtual Reality Headset for Immersive 3D Gaming. [Online]. Available: http://www.oculusvr.com/rift/.

[3] E. Barba, B. MacIntyre and E. D. Mynatt, "Here We Are! Where Are We? Locating Mixed Reality in The Age of the Smartphone," in *Proceedings of the IEEE*, 2012, vol. 100, no. 4, pp. 929–936.

[4] R. Azuma, "A survey of augmented reality," *Presence*, vol. 4, no. 4, pp. 355–385, August 1997.

[5] N. Davies, J. Landay, S. Hudson and A. Schmidt, "Guest Editors' Introduction: Rapid Prototyping for

Ubiquitous Computing," *IEEE Pervasive Computing*, Oct. 2005, vol. 4, no. 4, pp. 15–17.

[6] M. De Sá and E. Churchill, "Mobile augmented reality: exploring design and prototyping techniques," *Proceedings of the 14th International Conference on Human-computer Interaction with Mobile devices and Services.*, 2012, pp. 221–230.

[7] A. Oulasvirta, E. Kurvinen and T. Kankainen, "Understanding contexts by being there: case studies in bodystorming," *Personal Ubiquitous Computing*, vol. 7, no. 2, pp. 125–134, Jul. 2003.

[8] D. Courter and J. Springer, "Beyond desktop point and click: Immersive walkthrough of aerospace structures," *Aerospace Conference 2010 IEEE*, 2010, pp. 1–8.

[9] S. Carter, J. Mankoff, S. Klemmer and T. Matthews, "Exiting the Cleanroom: On Ecological Validity and Ubiquitous Computing," *Human-Computer Interaction*, vol. 23, no. 1, pp. 47–99, 2008.

[10] G. Alce, K. Hermodsson and M. Wallergård, "WozARd: A Wizard of Oz Tool for Mobile AR," in *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services - MobileHCI '13*, 2013, pp. 600–605.

[11] J. Lazar, J. H. Feng and H. Hochheiser, *Research Methods in Human-Computer Interaction*. West Sussex: John Wiley & Sons Ltd, 2010.

[12] R. C. Davies, "Applications of Systems Design Using Virtual Environments," in *Handbook of Virtual Environments - Design, Implementation, and Applications*, K. M. Stanney, Ed. Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey, 2002, pp. 1079–1100.

[13] E. Ragan, C. Wilkes, D. A. Bowman and T. Hollerer, "Simulation of augmented reality systems in purely virtual environments," *Virtual Reality Conference 2009. VR 2009. IEEE*, pp. 287–288, 2009.

[14] D. Baricevic, C. Lee and M. Turk, "A hand-held AR magic lens with user-perspective rendering," *Measurement Science and Technology and Augmented Reality (ISMAR), IEEE International Symposium*, pp. 197–206, 2012.

[15] C. Lee, G. Rincon, G. Meyer, T. Hollerer and D. A. Bowman, "The effects of visual realism on search tasks in mixed reality simulation," *Visualization and Computer Graphics, IEEE Transaction*, vol. 19, no. 4, pp. 547–556, 2013.

[16] M. Slater and S. Wilbur, "A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments,"

Presence Teleoperators Virtual Environments 6.6, pp. 603–616, 1997.

[17] M. Slater, "Measuring Presence: A Response to the Witmer and Singer Questionnaire," *Presence Teleoperators Virtual Environments*, vol. 8, no. 5, pp. 560–566, 1998.

[18] J. Freeman, S. E. Avons, R. Meddis, D. E. Pearson and W. IJsselsteijn, "Using Behavioral Realism to Estimate Presence: A Study of the Utility of Postural Responses to Motion Stimuli," *Presence Teleoperators Virtual Environments*, vol. 9, no. 2, pp. 149–164, Apr. 2000.

[19] K. P. Herndon, A. van Dam and M. Gleicher, "The Challenges of 3D Interaction: A CHI '94 Workshop," *SIGCHI Bull.*, vol. 26, no. 4, pp. 36–43, October 1994.

[20] D. A. Bowman, E. Kruijff, J. J. LaViola Jr. and I. Poupyrev, *3D user interfaces: theory and practice*. Addison-Wesley, 2004.

[21] Microsoft. (2015, April 29). Microsoft Kinect [Online]. Available: https://www.microsoft.com/en-us/kinectforwindows/.

[22] Sixense. (2015, April 29). Razer Hydra | Sixense [Online]. Available: http://sixense.com/hardware/razerhydra.

[23] I. Leap Motion. (2015, April 29). Leap Motion [Online]. Available: https://www.leapmotion.com/.

[24] 5DT. (2015, April 29). 5DT Data Glove 5 Ultra [Online]. Available: http://www.5dt.com/products/pdataglove5u.html.

[25] Sony. (2015, April 29). Sony Xperia Tablet Z [Online]. Available: http://www.sonymobile.com/se/products/tablets/xperia-tablet-z/.

[26] U. Technologies. (2015, April 25). Unity - Game Engine [Online]. Available: http://unity3d.com/.

[27] W. E. Marsh and F. Mérienne, "Nested Immersion: Describing and Classifying Augmented Virtual Reality," in *IEEE Symposium on 3D User Interfaces*, 2015.

[28] G. Alce, L. Thern, K. Hermodsson and M. Wallergård, "Feasibility Study of Ubiquitous Interaction Concepts," in *Procedia Computer Science,* 2014, vol. 39, pp. 35–42.

[29] N. ARC. (2015, April 29). NASA Task Load Index [Online]. Available: http://humansystems.arc.nasa.gov/groups/tlx/.

[30] S. G. Hart, "NASA-task load index (NASA-TLX); 20 years later," *Proceedings of the Human Factors Ergonomics Society Annual Meeting*, 2006, vol. 50, no. 9, pp. 904–908.

[31] K. Bystrom, W. Barfield and C. Hendrix, "A conceptual model of the sense of presence in virtual environments," *Presence Teleoperators Virtual Environments,* vol. 8, no. 2, pp. 241–244, 1999.

[32] SixSense. (2015, April 29). Leap Motion Sets a Course for VR [Online]. Available: http://blog.leapmotion.com/leap-motion-sets-a-course-for-vr/.

[33] Oculus. (2015, April 29). The All New Oculus Rift Development Kit 2 (DK2) Virtual Reality Headset [Online]. Available: http://www.oculusvr.com/dk2/.

[34] Virtuix. (2015, April 29). Omni: Move Naturally in Your Favorite Game [Online]. Available: https://www.kickstarter.com/projects/1944625487/omni-move-naturally-in-your-favorite-game.

[35] Sony. (2015, April 29). Project Morpheus, exclusive PlayStation TV shows and the best in fluffy dog games [Online]. Available: http://blogs.sonymobile.com/2014/03/21/project-morpheus-exclusive-playstation-tv-shows-and-the-best-in-fluffy-dog-games/?rl=cn.

[36] Microsoft. (2015, March 04). Microsoft HoloLens [Online]. Available: http://www.microsoft.com/microsoft-hololens/en-us.

[37] S. S. E. Inc. (2015, April 29). STEM System | Sixense [Online]. Available: http://sixense.com/STEM.

[38] T. Labs. (2015, April 29). Myo - Gesture control armband by Thalmic Labs [Online]. Available: https://www.thalmic.com/en/myo/.

[39] (2015, June 10). SMI Eye tracking glasses [Online]. Available: http://eyetracking-glasses.com/.

[40] J. Wilder, K. G. Hung, M. Tremaine Mantei and M. Kaur, "Eye Tracking in Virtual Environments," in *Handbook of Virtual Environments - Design, Implementation, and Applications*, K. Stanney, Ed. Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey, 2002, pp. 211–222.