Defining Size Parameters for Touch Interaction in Substitutional Reality Environments

Christian Mai(✉), Christian Valenta, and Heinrich Hußmann

LMU Munich, Media Informatics, Munich, Germany
{Christian.Mai,Hussmann}@ifi.lmu.de
http://www.medien.ifi.lmu.de

Abstract. The physical support of touch interaction for a 2D interface when wearing a fully immersive head-mounted display (HMD), e.g., by using the kitchen table in a home environment, improves the user's quality of interaction. To define interface parameters - button size, adaption over time- we conducted a user study. In two experiments with 30 participants in total, we compared the ability of the HMD user's pointing to targets on a 2D surface without visual feedback, with visual feedback of the touched position and a real-world baseline. As a result, we give estimates for button dimensions, interaction design based on the learning curve of the user and present insights on the tested feedback modalities. We show that providing no feedback has limitations, presenting the touched position helps to increase accuracy and a head-mounted finger tracker has advantages but also comes with restrictions.

Keywords: Head-mounted displays · Touch interaction
Pointing task · Haptic feedback · User interface design

1 Introduction

The growing distribution of consumer-grade head-mounted displays, like the Oculus Rift[1], leads to many situations in which HMDs are used in real-world (RW) environments. In contrast to the past decades these environments are not highly equipped laboratories, but the user's offices or homes. These environments offer limited space with a number of physical objects in the movement area and the motion tracking system is mostly limited to controllers and head. Further, the user is surrounded by people acting in the RW which might interfere with the users VR experience.

The limited space for physical moves of the HMD user in these everyday environments is one of the major challenges. Several concepts are addressing this: (A) research on locomotion in virtual reality (VR) [16], (B) redirected walking techniques [24] and (C) augmented virtuality [30] or substitutional [27]

[1] https://www.oculus.com/rift.

L. T. De Paolis and P. Bourdot (Eds.): AVR 2018, LNCS 10851, pp. 683–701, 2018.
https://doi.org/10.1007/978-3-319-95282-6_48

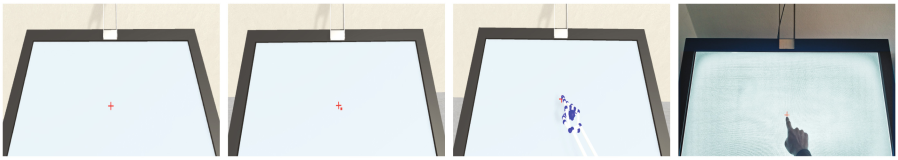

Fig. 1. Conditions during the haptical supported pointing task from left to right: No Feedback (NO_FB), rendering the touched position (TOUCHPOINT_FB), continuous visual feedback of the hand provided by handtracking (HAND_FB) and the real world baseline (RW_BASELINE)

environments. We focus on the concept (C), as (A) often leads to an unnatural abstraction of walking, e.g., swinging the arms to mimic the walking movement [25] and (B) still needs more movement space then a standard living room can offer.

There are several examples using concept (C) like Intel's project Alloy[2], which maps virtual objects onto furniture of a living room. Another example is the integration of a large touch surface to support interaction with the virtual world when controlling a large ship[3]. The authors argue in this paper that the lifespan of a ship is several decades, therefore instead of changing the whole physical bridge, the visualization in an HMD is easy to replace, and the RW needs some new touch surfaces. In a household, the touch surface might be even simpler by using a smartphone or tablet, or it might incorporate a transportable interactive projector like the Sony Xperia touch[4]. So it makes good sense to assume that there are physical touch surfaces available, even with touch sensing technology, but the user is interacting with a visual overlay from a virtual reality simulation.

When using existing physical surfaces to support touch interaction, the user immersed in VR will benefit from (1) a gain in precision - the haptic feedback of the physical table limits one degree of freedom for the touch -, (2) a more consistent and larger space for movement - the integration of physical objects in the VR shows the HMD user where to move without danger - (3) an increased feeling of being present in the VR - more senses are integrated into the VR experience [28] - (4) an environment supporting combined interaction between real and virtual world - the physical touch surfaces can be used as a shared frame of reference for RW bystanders and HMD users by either knowing what kind of interaction is coupled with a physical object or by presenting the same information on the surface for the HMD and the RW user.

Existing research and industrial development show that technically substitutional environments are possible, as the merged reality demo from Intel's project Alloy. Therefore we focus on the unaddressed question of the user interface (UI)

[2] https://www.youtube.com/watch?v=Ku9gjx5ECuY.
[3] http://e2c2.ict.usc.edu/blueshark-environment/.
[4] https://www.sonymobile.com/de/products/smart-products/xperia-touch/.

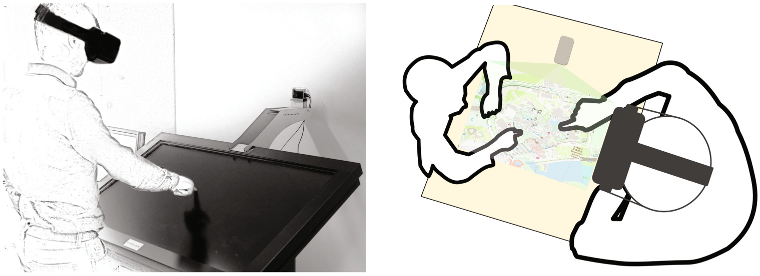

Fig. 2. Left: An HMD user acting in the virtual world with haptic feedback. Right: A short term interaction between a RW user and a HMD user with a shared frame of reference by using a mobile projector with touch recognition.

design for these types of systems in detail. In particular, we are addressing the precise definition of the button size and take into account the learning curve of the user for the UI design. The primary challenge arising is that the HMD blocks the users' view on their real hands and therefore the pointing accuracy is reduced. So we are interested in the requirements on the interface design and the compensation possibilities for the blocked view by different feedback methods, described further down.

We expect that touch interaction is possible in principle even if the HMD user does not see his/her own hands, as the human visuomotor system can compensate for such disturbances. We also expect that humans can deal with perceiving the hand in a position not matching precisely with the RW position. Such differences can arise when using a head-mounted finger tracker, e.g., due calibration reasons or the visual limitations introduced by the properties of the HMD [1,3,12,17].

To define button sizes and consider the adaption phase of the user to the respective systems constraints, we conducted a user study. We measured the pointing accuracy in a carefully calibrated system on a 2D touch surface (Fig. 2).

We decided on three feedback conditions plus a baseline condition (Fig. 1). The three feedback conditions differed in how we presented visual feedback to the user during a pointing movement, as described by MacKenzie [21]:

NO_FB: The user does not get any visual feedback of the hand or the touch point on the 2D surface.
TOUCHPOINT_FB: The user gets feedback on the interaction by a visual presentation of the point on the 2D surface where the touch happened.
HAND_FB: The users' hands are tracked by a head-mounted finger tracker and presented in the VR.
RW_BASELINE: The user conducts the task in the RW environment.

With our study, we were able to derive the following contributions:

- Detailed insights into the ability of the user to adapt the hand-eye coordination to the limitations introduced by the respective condition.

- Guidelines that help researchers and practitioners in designing user interfaces for substitutional virtual reality environments with targets positioned relative to physical surfaces.

2 Related Work

The usage of an HMD introduces two restrictions to the users' hand-eye coordination. One restriction is related to perceptional issues, generating an underestimation of distances when looking at a stereoscopic picture within an HMD [1,23] and the other one is not seeing the own hand when touching an object. Recent studies focusing on interaction with stereoscopic images on 2D touch surfaces in the RW looked into the effects when touching at a 2D surface while the object is rendered with positive or negative parallax, which means it looks like floating behind or in front of the 2D surface of the touchscreen monitor. The focus in this RW interaction is different from the usage of HMD environments, as they focus on occlusion problems and parallax effects and the real hand is visible all the time [7,32].

Studies that use HMD displays to assess interaction with the virtual world mostly focus on different aspects of factors influencing size and distance estimation, as summarized in the literature review of Renner et al. [26]. But it could be shown in different pointing tasks that an adaptation for the systems disturbances by the user is possible [1,4,13,22,29]. These studies do not match our target system of touching at a 2D surface as some of them did not offer a physical target for haptic feedback and/or the participants in the pointing task used a stylus and not their fingers.

2.1 Background: Hand-Eye Coordination in 3D Pointing Tasks

The concept we are aiming for is based on the idea of pointing at objects while not being able to see one's hand directly. In this chapter, we introduce how human pointing movements work to motivate that quasi-blind pointing at objects is possible. Further, the description of pointing movements helps us to understand and design the system, as we derive the feedback systems from it.

A goal-directed ballistic movement of the hand to touch a particular point can be divided into two phases. The first phase does not need visual guidance, the second phase, the actual touch, needs visual guidance for readjustment of the finger to the target [21]. Therefore in an RW scenario with undisturbed visual guidance the highest accuracy is expected. We can expect a high-end absolute tracking system to give a comparable accuracy. However, these systems are costly and difficult to use, which is why we do not consider them to be available in consumer grade hardware.

In a pointing movement without visual feedback of the hand, an open loop task, the proprioceptive system, is used to lead the hand to the target location [21]. In this case the users only have the visual information about the target location to plan the ballistic movement of the hand. Without continuous visual feedback a person does not have the possibility to refine the hand and finger position in the second phase of the pointing movement [21], which creates a deviation in the touch accuracy according to the precision in the first phase.

In our system, we always have the haptic feedback given by the touch surface. Further, when using touch-sensitive surfaces like tabletop displays or tablets, the system receives the touched position on this surface, so it can help the user to adapt to the distortions introduced by the HMD. The possible usage of this principle will be discussed in the following section.

2.2 Influence of Feedback on a Pointing Task

Many studies suggest that feedback helps the user to adapt the distance perception to the distortions introduced by the HMD system in an egocentric pointing task [1,3,12,17]. In contrast to that, the lack of visual feedback lets the users' accuracy drift [6] and therefore will not be described further but also tested in the condition NO_FB as it might get important when either tracking systems do not operate reliable or are not available at all. For us, the relevant feedback modalities are visual feedback, haptic feedback, and combinations of them.

Most of the studies using a pointing task present the visual feedback of the hand during the whole touch gesture [1,4,12] which therefore matches a closed loop pointing task. It was shown that visual feedback helps to improve the adaptation of the visuomotor system to the HMD distortions in a minimal amount of attempts [1,17] to a significantly higher accuracy [1,5,17]. Therefore continuous optical feedback would give the highest possible accuracy after some touch attempts and is tested as the condition HAND_FB.

A purer form of feedback is to present visual feedback of the touch position on the touched surface only (Fig. 2, bottom left). This does not create a closed loop for pointing, as the user has no continuous visual feedback of the fingertips. But we assume that there will be an adaptation of the visuomotor system for the visual disturbances in the first phase of the ballistic movement that does not need visual guidance [21]. Furthermore, users remember the touched position and adapt their motor planning at the following touch to the visually estimated delta [14,20]. This might compensate for the lack of the second phase of the pointing task to a certain amount and is tested in the condition Touchpoint_FB.

If visual and haptic feedback are presented together, it is essential to know how they influence the visuomotor systems adaptation process and how they influence each other. On the one hand, there is some tolerance for the divergence between visual and haptic depth cues [9,31]. This might help to compensate some issues that are introduced by the inaccuracies of the finger tracking system (Fig. 2, bottom right). Further, it is known for a 3D pointing gesture that giving haptic feedback is essential to determine the depth at the end of the correction phase, the second phase of the pointing task [31]. These effects might be helpful for all conditions, but we are not focusing on examining them. However, they are used in different applications to provide haptic feedback for the user [2,9].

3 Experiments

The experimentation on defining the parameters for touch interaction in a substitutional environment was divided into two experiments. This decision was based on a pre-study in which the participants reported to feel the physical effort in their arms after having gone through all conditions after another. In the following, we describe the visual parameters and the apparatus that was the same for both experiments. Both experiments included the Touchpoint_FB condition to indicate the validity of the experiment.

Visual Parameters. The findings on factors influencing the perception of distance and size in a virtual environment are very well described by Renner [26]. In fact, there are several factors known to affect size and distance estimation, but it is difficult to put them on a simple numeric scale. Therefore we only report on the parameters we considered when adjusting the HMD system to each user and the method we used.

We measured the interpupillary distance (IPD) in two ways. The first is by using the integrated IPD measuring method of the Oculus runtime (version 0.8). The second measurement was by using an IPD measuring template from Eye-Net Ltd.[5]. Both were conducted according to the given instructions. The mean of both methods was used. Further, all adjustable parameters of the Oculus were checked to be the same in the runtime environment as set at the hardware to guarantee a matching field of view (FOV) between the virtual cameras and the perceived FOV by the user.

The subjects were tested for their leading eye by using the Dolman, also called hole in the card, test[6]. The subjects had normal or corrected to normal vision during the experiment. None of the subjects reported any known disorders with their visual system. We used the graded circle test of a Rendot® stereotest[7] to measure the stereoacuity of the participants without any salience.

Apparatus. In our experiment we used a 42 multi-touch table with WXGA resolution with 1366 Pixel resolution in width (x-axis) and 768 Pixel in height (y-axis) and an optical tracking system with mm accuracy for touch recognition just above the surface. The size of the display and touch surface was 1.015 m × 0.57m. The table was slightly tilted towards the user (Fig. 3). We used the Oculus Rift Development Kit 2 [11] with the Oculus Head tracking system and the SDK Version 0.8. The camera was attached to the screen as shown in Fig. 3, left. The frame rate of the system was 75 Hz throughout the experiment. For the HAND_FB condition, a Leap Motion finger tracker was used with the official head mount and the Orion Beta SDK 3.1.1. We decided to use a head-mounted finger tracker, as they are not as intrusive and complicated to use as gloves and

[5] https://www.eye-net.com/media/cms/pdf/anl_neu_optiker.pdf.
[6] https://www.usaeyes.org/lasik/library/Dominant-Eye-Test.pdf.
[7] http://www.visionassessment.com/1005.shtml.

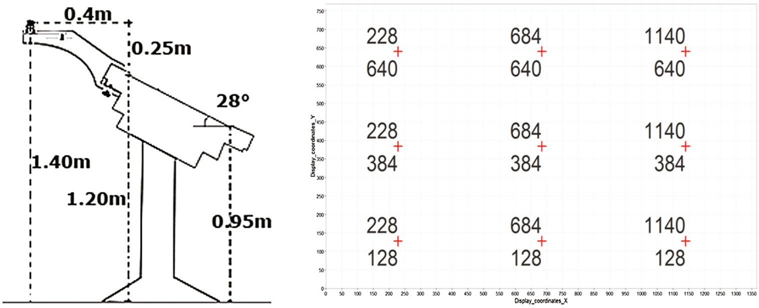

Fig. 3. Drawing of touch table dimensions (left) and the target positions in pixel coordinates on the screen (right). (Color figure online)

they are not bound to a specific area like external tracking systems, e.g., Vicon optical tracking[8], as they move with the user.

The table, wall, floor, and room were 3D-modeled in their dimensions according to their real counterpart. The textures were approximated manually. The visualization for the Oculus Rift and the touch table was rendered with Unity 5.3.1. The visualized background on the screen in the real task and the virtual representation was a 50% grey. The target cross was 30 * 30 pixels in red.

To map the virtual representation of the touch surface to the real surface used in our study, the tracking camera was mounted in a defined position relative to the table. The virtual and the RW were overlayed with their origin of coordinate systems in the center of the camera.

Nine different target positions on the screen are defined by coordinates. The screen coordinates start with (0,0) at the bottom left (Fig. 3, right).

During the experiment the target ID and the according position of a pointing event to that target was recorded in pixel values. The metric distance to the target was calculated by using the scale known from the value pixels per centimeter.

3.1 Experiment on Comparing NO_FB and TOUCHPOINT_FB

Participants. 14 male and 6 female with a mean age of 28 (SD = 9.9) participated in the experiment. The subjects were students and employees from different departments at the LMU Munich. 13 of the participants had minor experiences with HMDs, and the rest had no experience. One subject was left handed.

Between the different conditions, the subjects had time to take a break for a few minutes. The total time for each subject to participate in the experiment was about 40 min including pre-questionnaire, instructions, experiment, and breaks.

8 https://www.vicon.com/.

Method. The experiment was conducted as a 2×9 within-subject design. The independent variables are the two feedback conditions (NO_FB vs. TOUCHPOINT_FB) and the nine target positions (Fig. 3). The dependent variable was the distance from the touch position by the user on the touch screen to the presented target cross center. In between the two feedback conditions the RW condition was conducted as a baseline for the system. In between conditions the participants took off the HMD and played with a tennis ball to avoid carry-over effects and the so called negative aftereffect [15]. All subjects completed all feedback conditions.

The experiment started with the welcoming of the subjects and the explanation of the procedure of the experiment and its goal. They were informed that no personal data is recorded during the study. They were instructed to touch with the fingertip of the index finger of their dominant hand exactly at the center of the cross, lift their hand after touching and go on to the next target. They were asked to concentrate and touch as accurately as possible. The subjects should move quickly, but still feel comfortable. If they got hectic or too slow, they were requested to adapt their speed and keep the focus on precision. They were allowed to rest in between the conditions. After that their dominant eye and stereo acuity was determined. They were told about simulator sickness and possible symptoms and were informed to stop their session at every time if they felt uncomfortable [18].

Then they put on the HMD while standing in front of the touch table. They had to touch at the center of the red target cross. The target disappeared after the touch happened and the next target appeared. Therefore only one target was visible at the same time. All nine targets were touched in a randomized order before one target was touched a second time. This we call a block which will be relevant to identify the adaption phase. The same target was never requested to touch two times in a row. Every subject performed 180 touches per feedback condition.

3.2 Results on Comparing No_FB and TOUCHPOINT_FB

Outliers. The collected data included unintended touches that occurred by accidentally hitting the touch surface in between two touches. We chose a distance of more than 20 cm to be an unintended touch, as this is the shortest distance between two target centers. 42 touches were excluded for the condition NO_FB, 33 for Touchpoint_FB and 38 for the RW_BASELINE.

The data was processed by using a repeated measures ANOVA. As Mauchly's test of sphericity was violated in all cases, the degrees of freedom were corrected by the Greenhouse-Geisser estimates to compensate for this. As posthoc a Tukey multi comparison ($p = .5$) was calculated with Bonferroni correction.

Identifying the Adaption Phase. We started the analysis of the results with an analysis of adaption phases, as we want to exclude the adaption phase from the analysis for the final UI design. As one can see for the condition

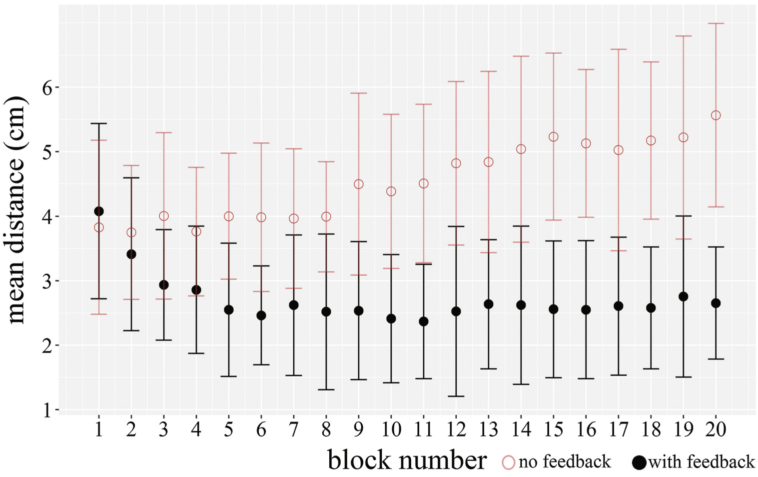

Fig. 4. Touch accuracy over blocks during the NO_FB condition with mean and 95% confident intervals. Each block comprises a touch to each target and therefore 9 touches. (Color figure online)

TOUCHPOINT_FB in Fig. 4 (black, solid circle), there seems to be a adaption phase, as the mean distance value drops from the first to the fourth block. A repeated measures ANOVA comparing the mean touch distance between the trials showed a significant difference in touch accuracy between the blocks ($F(5.915, 112.377) = .505,\ p < .05$; ŋˆ2 =.225).

The post hoc tests showed that block 1 was highly significantly different from all following blocks, except the block number 2 at $p < .05$. Block 2 showed a significant difference to most other blocks but block 1, 3 and 4 at $p < .05$. Block 3 only showed a significant difference to block 1 at $p < .05$. Block 4 showed no difference to the other blocks. The mean touch distances declined from 4.1 cm (SD = 1.4 cm) in block 1, 3.4 cm (SD = 1.2 cm) in block 2 to 2.9 cm (SD = 0.9 cm) in block 3. This is an improvement of 30% from touching each target for the first time (block 1) to the third time (block 3). Therefore we expect 2 touches to a target to be sufficient to learn to compensate for the introduced disturbances.

The first two blocks will be excluded for the further analysis to describe requirements on the UI design. The resulting dataset for the TOUCHPOINT_FB condition includes 3207 touches. This dataset without block 1, block 2 and without outliers will be called the dataset after the training phase.

Comparison of Overall Touch Accuracy of the Conditions. To quantify the differences when implementing on-screen feedback after the users adapted to the system, we compared the touch distributions between the two conditions. A Quantile-Quantile plot showed that the processed data has no normal distribution. A Wilcoxon Signed Rank test indicated that the scores for the NO_FB were

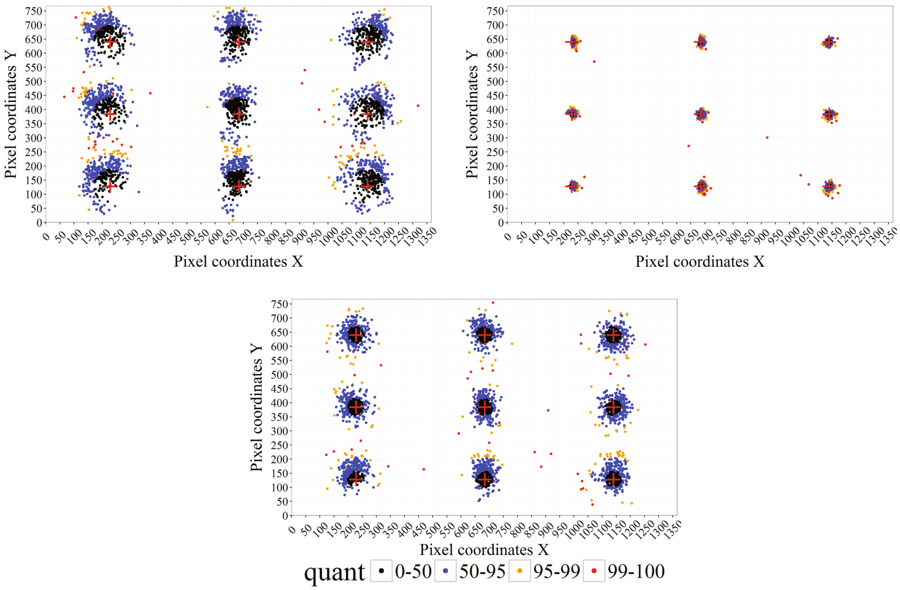

Fig. 5. Distribution of all touches, without outliers and the adaption phase. Top left: NO_FB, Top right: RW_BASELINE, Bottom: TOUCHPOINT_FB. The shown target crosses are oversized for better visibility

significantly higher than the scores for the TOUCHPOINT_FB ($p < .001$). The mean touch accuracy without feedback was 4.5 cm (SD = 2.3 cm), with feedback it was 2.7 cm (SD = 1.9 cm).

As one can see in Fig. 4 (brown circle, no fill), there is no adaptation of the user to the condition NO_FB over time, as there is a drift to higher touch deviation. Further, as no threshold is visible, we assume that users are not aware of the drift and the deviation gets worse over time. Therefore we did not consider the NO_FB condition in the further analysis as it has a high potential for failure. As regards to the completeness, in the RW_BASELINE condition, the users could achieve in average 0.7 cm (SD = 0.4 cm) in accuracy for 99% of the touches (Fig. 5, top right).

Influence of Target Position on Touch Precision. Using an ANOVA, as described at the beginning of this section, we found a significant difference in touch accuracy between the positions ($F(3.668, 69.697) = 3.334, p < .05$; ŋˆ2 = .149).

The Bonferroni corrected post hoc comparison showed that there is a significant difference between the lower left target (M = 3.0 cm, SD = 1.1 cm) and the middle target (M = 2.2 cm, SD = 0.8 cm) for the condition TOUCHPOINT_FB in the dataset after the training phase.

The difference is rather small, therefore we do not expect a relevant effect for the user's interaction precision.

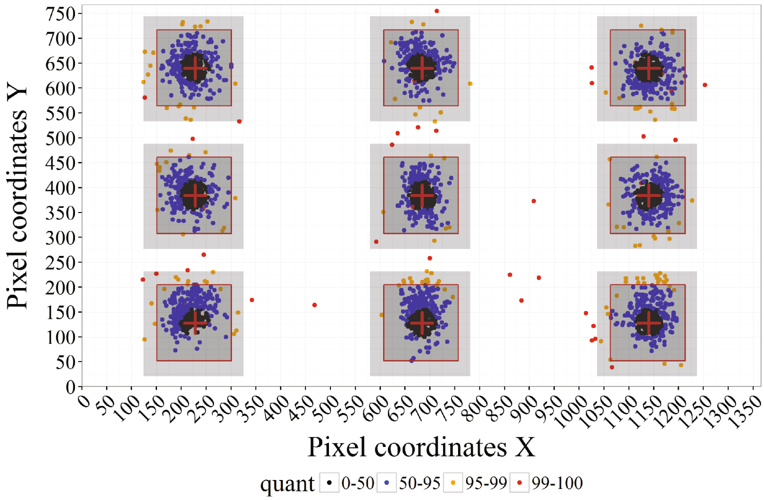

Fig. 6. Distribution of all touches, without outliers and without the adaption phase for the TOUCHPOINT_FB condition. The squares represent the size for buttons depending on the accepted failure rate.

Variability of Touches in Horizontal and Vertical Display Direction and Resulting Target Size. In the previous sections, the touch distribution was calculated as circular shapes around the center of the targets. But as one can see in Fig. 5 TOUCHPOINT_FB (bottom), the centroid of touches to the targets might be shifted from the actual target, which would give the possibility for improvement by individual calibration [8]. Also, the variability of the touch precision could be bigger on the y-axis compared to the x-axis, which could be supported by buttons of other shapes than squares.

The mean deviation of the centroid of touches to each target in the condition with feedback along the x-axis is 0.4 cm (SD = 0.3 cm), and along the y-axis it is 0.1 cm (SD = 0.2 cm). In comparison to the touch area this deviation is very small, therefore we do not expect a benefit from implementing such a system.

The acceptance of missing a button can help to reduce the target size. As one can see in Fig. 6, the acceptance of 5% touches that would miss the button will influence the necessary size for a button visible.

To generate a minimal target size we compared the maximum deviation of the touches in x-direction to the deviation of the touches to the different targets in y-direction for the 99% closest points. A t-test between the two directions showed significant differences ($t(8) = -1.1663$, $p = 0.027$)) with a mean of 9.8 cm (SD = 1.3 cm) deviation in x- and 10.7 cm (SD = 0.6 cm) in y-direction. For 99% hit success the bounding box would need to be x = 14.9 cm and y = 15.6 cm (Fig. 6). For 95% successful touches the button would need to be x= 11.2cm and y= 11.4cm.

3.3 Experiment on Comparing TOUCHPOINT_FB and HAND_FB

Participants. 6 male and 4 female subjects with a mean age of 26 (SD = 3) participated in our second experiment. The subjects were students and employees from different departments at the LMU Munich. Two of the participants had no experiences with HMDs, and the other subjects had minor experience. Between the different conditions, the subjects had time to take a break for a few minutes. The total time for each subject to participate in the experiment was about 40 min including pre-questionnaire, instructions, experiment, and breaks.

Methods. The method in this experiment was the same as in the experiment described above. In the condition HAND_FB the visualization of the tracked fingers and the touch position on the touchscreen was compared by visualizing the touch position on the touch surface and the fingers within the HMD. In case there was a mismatch between the physical position of the fingers on the screen and the virtual representation it was manually calibrated by putting the real finger on the middle of a target and moving the virtual finger to match the touched position visually.

The rest of the experiment is analog to the methods of the experiment on comparing NO_FB and TOUCHPOINT_FB in the section above.

3.4 Results of Experiment on Comparing TOUCHPOINT_FB and HAND_FB

Outliers. The collected data includes unintended touches that occurred by accidentally touching the surface in between two touches. We chose a distance of more than 20 cm to be an unintended touch. 13 touches were excluded for the condition TOUCHPOINT_FB and 19 for the condition HAND_FB.

Insights on Validity of Results Between the Experiments. To get an insight on the validity of the two experiments we compared the TOUCHPOINT_FB condition from both experiments. There was not a significant difference in the scores for TOUCHPOINT_FB from the first (M = 2.6 cm, SD = 3.5 cm) and TOUCHPOINT_FB from the second (M = 2.7 cm, SD = 2.8) experiment ($t(4143) = 1.7895$, $p = 0.07$). Further a t-test only tests for differences and does not proof equality, the difference of 0.1 cm with the given sample size is very small. Therefore we expect a good validity for the experiment. Also, the number of outliers is similar to 1.7 outliers per participant in the experiment presented above and 1.3 in this experiment.

Identifying the Adaption Phase. Similarly to the experiment presented above, a adaption phase can be detected for the TOUCHPOINT_FB condition on the table. A paired t-test showed significant difference between the distance to the target between the first (M = 3.12 cm, SD = 1.9 cm) and second (M = 2.6 cm, SD = 1.5 cm) block of touches ($t(89) = 2.0969$, $p = 0.039$). For the condition with finger representation, no adaption phase could be determined.

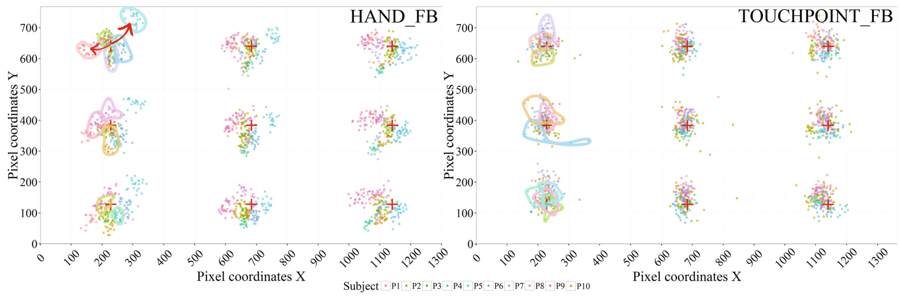

Fig. 7. Touches per subject. Left: HAND_FB. Right: TOUCHPOINT_FB. The contours show exemplarily the distribution of touches for each subject.

Relative Spreading of Touches and Shift of Touch Centroids. The plot of the touch distribution for each subject in Fig. 7 shows that the area needed to cover all touches appears to be much bigger in the HAND_FB condition (Fig. 7, left), then for the TOUCHPOINT_FB condition (Fig. 7, right). A paired t-test shows a highly significant difference for the mean touch distribution between the TOUCHPOINT_FB condition (M = 4.3 cm, SD = 1.9 cm) and HAND_FB condition (M = 5.6 cm, SD= 2.7 cm) ($t(320) = -5.5486$, $p = 0.001$). But the distribution of touches for every single subject is smaller in the condition HAND_FB, see the following paragraph.

It is visible in Fig. 7 that the touches in the TOUCHPOINT_FB condition are spread around the targets equally for all subjects (Fig. 7, right). In the HAND_FB condition, the subjects show clusters of touches that do not overlap as much as in the TOUCHPOINT_FB condition. A drift of the centroid of touches to a specific direction and amount is visible (Fig. 7, left). This drift is individual for each subject.

Variability of Touches in Horizontal and Vertical Display Direction and Resulting Target Size for HAND_FB. In the following we only analyze the condition HAND_FB, as the TOUCHPOINT_FB condition was compared in a previous section. A paired t-test of comparing the distribution of touches in x- (M = 3.8 cm, SD = 1.1 cm) and y- (M = 4.9 cm, SD = 2.3 cm) direction shows a significant difference ($t(89) = -3{,}97$, $p = 0.001$). For the acceptance of 5% missed buttons this leads to a target size of 7.8 cm in x- and 10.1 cm in the y-direction.

3.5 Limitations

Although our user study was designed with great care, there are some limitations to it. The mounting for the leap motion finger tracker on the is costume made based on the instruction provided by Leap Motion. But we still had to calibrate for each subject manually. This calibration was done before the experiment and carefully tested. However, during the experiment the calibration might have been lost when the HMD moved relative to the users head.

Further, the number of participants was rather small. As we did not find a remarkable difference between the first and the second study, we argue that our results still give a very good estimation for the pointing accuracy.

Also system inherent attributes might influence the results. Other HMD models might have a different influence on the size and distance estimation. Also pointing to a surface above the head height and different orientation of the surface to the user might lead to deviating results. However, the latest are extreme cases and might be considered for further studies.

4 Discussion of the Study Results

The results of the experiment give indications for the definition of button sizes as well as the interaction design for a system enabling touch interaction in substitutional virtual reality environments scenarios. In the following we present guidelines based on the results of our experiment that help to design such systems closely coupled to the users' needs.

Keep in Mind the Limitations of the Feedback Modality. In our study, we tested the conditions without restrictions of the HMD (RW_BASELINE) and the three different visualization possibilities no feedback at all (NO_FB), feedback about the touched position on the touch surface (TOUCHPOINT_FB) and continuous visual feedback given by a head-mounted finger tracking system (HAND_FB). The different systems have different requirements on the UI and interaction design.

Interaction in the NO_FB condition is possible. However, this interaction mode puts strong limitations on the interface design, since the touches have a large distribution and a steady drift over time. Further, the distribution of touches gets larger over time, as it was shown in studies with pointing to targets in 3D space before [6]. Therefore not presenting any feedback might only be an option for short-term interactions with a separation of the targets based on directions relative to the user, like left and right or up and downwards.

The TOUCHPOINT_FB helped to improve accuracy as we expected from similar studies in different interaction scenarios before [1,3,12,17]. Therefore existing surfaces with touch detection should be included in the interaction whenever possible. Examples could be but are not limited to tablets, touchscreens or other devices that include tracking like mobile projection system -e.g., Sony Xperia touch-. However, in contrast to our assumption based on the results of MacKenzie [21], the resulting mean deviation of 2.7 cm for the touches is much higher than in a RW scenario with 0.7 cm of deviation.

Further, we could show that the representation of the touchpoint does not only improve the accuracy in the course of the interaction but further prevents a drift of the touches centroid over time. This leads to a reliable touch input for the interaction design.

When accepting 5% touches that miss the desired button, it is possible to fit nine bounding boxes lateral to the users forward direction and five bounding boxes on the used screen along the users forward direction. That leads to 45 hit zones in total for the on-screen touch feedback. This would be enough to design a usable keyboard for text input.

HAND_FB by the usage of a head-mounted finger tracker enables interaction within the limitations of the used finger tracker. In our system, the finger tracker helped to increase the touch accuracy. This is highly dependent on the used head-mounted finger tracking system. Therefore not the definition of buttons sizes but the finding for interaction design is the interesting outcome for this condition in our study.

We could show a drift in the centroid of touches that most likely was introduced due to our manual calibration process. Although we checked the calibration on different targets before the study, the calibration might have been lost during the study by movements of the HMD relative to the users head. Therefore the hand-eye coordination was affected. This is less a problem in purely VR systems as the user does not recognize the shift. But it is a problem for detecting precisely the touch on a target positioned relative to the real-world surface as it changes the calibration. We therefore suggest to include an algorithm that continually recalibrates the finger tracker, e.g., by a technique introduced by Buschek and colleagues [8].

Surprisingly the distribution of touches was not much smaller in the HAND_FB then in the TOUCHPOINT_FB condition. The main reason for that is occlusion of the finger by the user's hand when pointing at a surface. Therefore no absolute tracking of the finger is possible for the tracking system. Therefore the underlying algorithms of the tracking systems assume a finger position as it was trained. Most likely this is the decision between a straight and a bent finger. As a result, we suggest using buttons that have a portrait format. This would lead to a UI that has more buttons in the horizontal direction than in the vertical direction relative to the user. To prevent the occlusion it might help to position buttons in a way that leads to a hand position with the back of the hand perpendicular to the users head.

Further, the head-mounted finger tracker needs to be carefully calibrated to match the user's fingers precisely. Otherwise, the visual position of the finger, the physical feedback of the touch surface and the position of the target relative to the touch surface will not match, which results into a confusion of the user. Also, incomprehensible reactions of the systems can happen as the users see their fingers touching the correct button, but the physical finger hits the button next to it that is positioned relative to the physical touch surface. We calibrated our system in a somewhat naive way, by asking the users about their perception of the finger position and moving it if a mismatch was reported. A calibration process with about 10 touches to known targets to find the delta between touch position and target position could be used to improve this calibration.

Adaptation of the UI for the User Over Time. The TOUCHPOINT_FB helped, as expected from related work, to improve the touch accuracy over the trials. The maximum accuracy is reached after about 18 touches, where the user had touched each target two times. As a result, the user should either be trained to use the system, or the size of targets should adapt over time. The last concept could be implemented through bigger buttons at the beginning of the interaction that shrink over time. But this also means that there is less space for targets on the screen at the beginning of the interaction. Taking this into account on could imagine the presentation of a login screen with login buttons distributed over the whole display or a game during the start of the system. This game could be designed as a multi-level game that offers more and smaller buttons in the second step of the procedure.

Pay Attention When Using Head-Mounted Finger Tracking. As we reported above, the user's hand in a pointing task has a high potential to introduce inaccuracies by shadowing the finger. Therefore in systems with head-mounted finger tracking, the buttons should be positioned in a way that the user can touch them with the back of his/her hand perpendicular to his/her line of sight.

Position of Buttons has a Minor Influence on Accuracy. The TOUCHPOINT_FB condition did not show lower accuracy in the border areas of the field of view, e.g., due to lens distortion in the HMD, as we expected from related work [26]. The inaccuracy might not be present in comparison to earlier experiments due to the significantly better visual systems which are known for improved size and distance estimation in consumer grade HMDs [10,19]. Also, the higher demand on the natural motion when reaching to a target further away [21] in combination with the missing visual feedback during the pointing task therefore seemed not to influence the users pointing accuracy. This also means that it is possible to transfer the findings made in our study to other touch screen systems. We assume the results to be consistent for systems with small deviations in size, position, and tilt of the touch surface, as long as the interaction is in the arms reach distance for the user. However, there might be extreme positions in which the human motor system comes to a boundary of the ability to move precisely, like the extreme case of leaning backward and pushing a button above or in the back of the user.

5 Future Work

The system setup used in this work is a minimal setup for the least complicated UI with static targets. More complex UIs might have moving or moveable UI elements, which might lead to different results for button sizes which should be researched in the future. During the everyday use of a substitutional environment, the user might interact with surfaces at particular points in time during the room-scale virtual reality session. Therefore another open question is, if the user can remember the visuomotor system adaptation to possible drift in the conditions TOUCHPOINT_FB and HAND_FB and if s/he can transfer this from one touch surface to another touch surface. The position of the touch surface and the relative angle to the user is another factor that needs to be researched.

References

1. Altenhoff, B., Napieralski, P., Long, L., Bertrand, J., Pagano, C., Babu, S., Davis, T.: Effects of calibration to visual and haptic feedback on near-field depth perception in an immersive virtual environment. In: Proceedings of the ACM Symposium on Applied Perception (SAP 2012), pp. 71–78. ACM, New York, NY, USA (2012)
2. Azmandian, M., Hancock, M., Benko, H., Ofek, E., Wilson, A.: Haptic retargeting: dynamic repurposing of passive haptics for enhanced virtual reality experiences. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI 2016), pp. 1968–1979. ACM, New York, NY, USA (2016)
3. Bingham, G.: Calibration of distance and size does not calibrate shape information: comparison of dynamic monocular and static and dynamic binocular vision. Ecol. Psychol. **17**(2), 55–74 (2005)
4. Bingham, G., Bradley, A., Bailey, M., Vinner, R.: Accommodation, occlusion, and disparity matching are used to guide reaching: a comparison of actual versus virtual environments. J. Exper. Psychol. Hum. Percept. Perform. **27**(6), 1314–1334 (2001)
5. Bingham, G., Crowell, J., Todd, J.: Distortions of distance and shape are not produced by a single continuous transformation of reach space. Percept. Psychophys. **66**(1), 152–169 (2004)
6. Bingham, G., Zaal, F., Robin, D., Shull, A.: Distortions in definite distance and shape perception as measured by reaching without and with haptic feedback. J. Exper. Psychol. Hum. Percept. Perform. **26**(4), 1436–1460 (2000)
7. Bruder, G., Steinicke, F., Sturzlinger, W.: To touch or not to touch?: Comparing 2D touch and 3D mid-air interaction on stereoscopic tabletop surfaces. In: Proceedings of the 1st Symposium on Spatial User Interaction (SUI 2013), pp. 9–16. ACM, New York, NY, USA (2013)
8. Buschek, D., Alt, F.: Touchml: A machine learning toolkit for modelling spatial touch targeting behaviour. In: Proceedings of the 20th International Conference on Intelligent User Interfaces (IUI 2015). ACM, New York, NY, USA (2015)
9. Cheng, L.P., Ofek, E., Holz, C., Benko, H., Wilson, A.: Sparse haptic proxy: Touch feedback in virtual environments using a general passive prop. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17). pp. 3718–3728 (2017)
10. Creem-Regehr, S., Stefanucci, J., Thompson, W., Nash, N., McCardell, M.: Egocentric distance perception in the oculus rift (dk2). In: Proceedings of the ACM SIGGRAPH Symposium on Applied Perception (SAP 2015), pp. 47–50. ACM, New York, NY, USA (2015)
11. Desai, P., Desai, P., Ajmera, K., Mehta, K.: A review paper on oculus rift-a virtual reality headset. CoRR abs/1408.1173 (2014)
12. Ebrahimi, E., Altenhoff, B., Hartman, L., Jones, A., Babu, S., Pagano, C., Davis, T.: Effects of visual and proprioceptive information in visuo-motor calibration during a closed-loop physical reach task in immersive virtual environments. In: Proceedings of the ACM Symposium on Applied Perception (SAP 2014), pp. 103–110. ACM, New York, NY, USA (2014)
13. Ebrahimi, E., Altenhoff, B., Pagano, C., Babu, S.: Carryover effects of calibration to visual and proprioceptive information on near field distance judgments in 3d user interaction. In: 2015 IEEE Symposium on 3D User Interfaces (3DUI), pp. 97–104 (2015)
14. Harris, C.: Perceptual adaptation to inverted, reversed, and displaced vision. Psychol. Rev. **72**(6), 419–444 (1965)

15. Held, R.: Plasticity in sensory-motor systems. Sci. Am. **213**(5), 84–94 (1965)
16. Jankowski, J., Hachet, M.: Advances in interaction with 3D environments. Comput. Graph. Forum **34**(1), 152–190 (2015)
17. Kelly, J., Hammel, W., Siegel, Z., Sjolund, L.: Recalibration of perceived distance in virtual environments occurs rapidly and transfers asymmetrically across scale. IEEE Trans. Vis. Comput. Graph. **20**(4), 588–595 (2014)
18. Kennedy, R., Lane, N., Berbaum, K., Lilienthal, M.: Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol. **3**(3), 203–220 (1993)
19. Li, J.: The benefit of being physically present. Int. J. Hum Comput Stud. **77**(C), 23–37 (2015)
20. Lillicrap, T., Moreno-Briseno, P., Diaz, R., Tweed, D., Troje, N., Fernandez-Ruiz, J.: Adapting to inversion of the visual field: a new twist on an old problem. Exper. Brain Res. **228**(3), 327–339 (2013)
21. MacKenzie, C., Iberall, T.: The Grasping Hand. In: Advances in Psychology, vol. 104. North-Holland, Amsterdam and New York (1994)
22. Naceri, A., Chellali, R.: Depth perception within peripersonal space using head-mounted display. Presence: Teleoperators Virtual Environ. **20**(3), 254–272 (2011)
23. Napieralski, P., Altenhoff, B., Bertrand, J., Long, L., Babu, S., Pagano, C., Kern, J., Davis, T.: Near-field distance perception in real and virtual environments using both verbal and action responses. ACM Trans. Appl. Percept. **8**(3), 18:1–18:19 (2011)
24. Nilsson, N., Peck, T., Bruder, G., Hodgson, E., Serafin, S., Suma, E., Whitton, M., Steinicke, F.: 15 years of research on redirected walking in immersive virtual environments. IEEE Comput. Graph. Appl. **38**, 44–56 (2018)
25. Pai, Y., Kunze, K.: Armswing: Using arm swings for accessible and immersive navigation in AR/VR spaces. In: Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia (MUM 2017), pp. 189–198. ACM, New York, NY, USA (2017)
26. Renner, R., Velichkovsky, B., Helmert, J.: The perception of egocentric distances in virtual environments - a review. ACM Comput. Surv. **46**(2), 1–40 (2013)
27. Simeone, A., Velloso, E., Gellersen, H.: Substitutional reality: Using the physical environment to design virtual reality experiences. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI 2015), pp. 3307–3316. ACM, New York, NY, USA (2015)
28. Slater, M.: Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos. Trans. R. Soc. London B: Biol. Sci. **364**(1535), 3549–3557 (2009)
29. Sprague, D., Po, B., Booth, K.: The importance of accurate VR head registration on skilled motor performance. In: Proceedings of Graphics Interface 2006 (GI 2006), pp. 131–137. Canadian Information Processing Society, Toronto, Ont., Canada, Canada (2006)
30. Sra, M., Garrido-Jurado, S., Schmandt, C., Maes, P.: Procedurally generated virtual reality from 3D reconstructed physical space. In: Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology (VRST 2016), pp. 191–200 (2016)

31. Valkov, D., Giesler, A., Hinrichs, K.: Evaluation of depth perception for touch interaction with stereoscopic rendered objects. In: Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces (ITS 2012), pp. 21–30. ACM, New York, NY, USA (2012)
32. Valkov, D., Steinicke, F., Bruder, G., Hinrichs, K.: 2d touching of 3d stereoscopic objects. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2011), pp. 1353–1362. ACM, New York, NY, USA (2011)