Universiteit Utrecht

Virtual Reality on Mobile Devices

Investigating Factors Affecting the Shoebox VR Experience in Motion-based Games

Jannes Beurskens

ICA-3649598

Supervisor

Dr. Wolfgang Hürst

Abstract

Mobile 3D applications, especially games, are gaining popularity and due to the constant increase in graphics capabilities of mobile devices, the creation of more elaborate virtual worlds becomes possible. However, compared to traditional gaming platforms, mobile devices have a small form factor, which imposes new challenges on game designers, but also opens new opportunities.

Shoebox VR offers a novel way of visualizing virtual environments and has already been applied in games. Whereas current games mainly use still scenes, we investigated the use of Shoebox VR in motion-based games. Our previous study demonstrated that Shoebox VR improved subjective user experience, in terms of subjective sense of depth perception, entertainment value and immersion. To further explore how user experience and gameplay are affected by Shoebox VR, this study investigated whether these variables are affected by factors like screen size and user interaction.

By developing games, we were able to show that although screen size did not affect subjective experience, gameplay performance increased with a smartphone-sized device at low speeds using Shoebox VR. In addition, we demonstrated that touch screen interaction affects games using Shoebox VR, as gameplay performance was higher for Shoebox VR compared to standard visualization methods. In contrast, tilting interaction did not appear to affect Shoebox VR, indicating that the tilting controls can be used in combination with Shoebox VR without decreasing user experience or gameplay performance. Finally, our results suggested that Shoebox VR is especially functional and well-used when it is implemented as an essential part of the gameplay.

Our study provided more insight in the usability of Shoebox VR in motion-based games, which contributes to the improvement of visualization methods of 3D environments on mobile devices.

Table of Contents

1. Introduction

1.1 Motivation

Mobile 3D applications, especially games, are gaining popularity and due to the constant increase in graphics capabilities of mobile devices, the creation of more elaborate virtual worlds becomes possible. However, compared to traditional gaming platforms, such as PCs and gaming consoles, mobile devices have a small form factor, which imposes new challenges on game designers. For example, one of these challenges is to create a mobile experience that reaches the same level of immersion that traditional platforms can provide, as research has shown that the user's sense of immersion tends to decrease as the screen size decreases (Hou *et. al.*, 2012). Furthermore, input is more restricted due to the lack of traditional peripherals, such as mice, keyboards and game controllers, thereby limiting, for example, the number of methods available for navigating virtual environments.

Yet, the small form factor combined with special hardware also opens new opportunities. For example, due to the availability of a touch screen, the user has access to touch input and the addition of an accelerometer or a gyroscope allows for input via tilting. The latter also enables new approaches for visualizing 3D virtual environments. One such approach is the Shoebox VR (virtual reality) concept, which allows users to change the perspective from which a virtual world is viewed by changing the tilting angle of the mobile device. Using Shoebox VR developers can create interesting new ways to view and interact with virtual environments as illustrated in figure 1.

Fig. 1: ***Shoebox VR example.*** *A 3D scene representing objects placed in a box. The centre image shows the neutral position, where the user looks directly into the box. The left and right images show how the user can look at the scene from different perspectives by tilting the device. Dynamically changing the perspective according to the phone's tilting angle creates a "3D effect".*

Shoebox VR offers a novel way of visualizing virtual environments and has already been applied in games, for example in puzzle and object search games. However, these games tend to only use still scenes, where the camera remains in one position and thus navigating in the virtual environment is not possible. Therefore, we investigated the usability of Shoebox VR in simple motion-based games. Whereas Van Laar (2014) has focussed on user-controlled motion, i.e. navigation, our research will focus on fixed motion by examining situations in which users do not have full control over the motion of a playable character, as can be illustrated by endless runner games, in which users only have control over direction, but not the speed.

Our previous study (Hürst & Beurskens, 2013) showed that Shoebox VR increased the subjective experience compared to standard visualization in motion-based applications, possibly compensating for the decreased immersion that occurs due to small screen size. However, it is not clear yet how different factors, such as screen size, can affect the effect seen for Shoebox VR in terms both of performance and subjective experience. In addition, it has not been elucidated yet how the implementation of Shoebox VR affects user interaction. For example, it might be possible that the tilting controls that come with Shoebox VR interfere with other forms of interaction, such as touch screen interaction.

Table 1: Definition of terms

Term	**Definition**
Standard visualization	Techniques that are commonly used for the visualization of 3D environments that do not take the user's physical position with respect to the screen into account.
Shoebox VR (Virtual Reality)	Visualization technique for handheld devices that renders an image of a 3D environment from a perspective that is accurate with the user's physical head position. This is done by approximation of the user's head position with respect to the screen, based on the device's tilt angle (Section 2.2).
Subjective experience	We define the subjective experience in this thesis as the combination of entertainment value and immersion.

1.2 Goal and Contributions

The major goal of this study is to further investigate the usability of Shoebox VR in motion-based games, with a particular focus on fixed forward camera motion. Therefore, we determined whether Shoebox VR has any objective or subjective benefits in games based on fixed motion, by measuring different performance indicators and grading the subjective experience, which is defined here as the combination of entertainment value and immersion. In this perspective, the following subjects were addressed:

- Our previous research showed that significant subjective benefit can be provided by Shoebox VR on smartphone-sized displays. Yet, since mobile games are played on devices of various sizes, it is important to know whether these results can be generalized to larger screens. Therefore, we address the following question: How does screen size affect depth perception and subjective experience with Shoebox VR and how does it affect usability of Shoebox VR in a motion-based game? (Section 3.1)
- In order to determine the usefulness of Shoebox VR in mobile games, it is also important to know how user interaction, of which touch screen interaction is most common, is affected by it. Therefore, we address the following question: How does touch screen interaction affect task performance and subjective experience with Shoebox VR in a motion-based game? (Section 3.2)
- Tilting is another from of interaction that is often used for 3D games on mobile devices. Hence, it is important to know if and how this kind of interaction and the tilting actions performed to change perspective in Shoebox VR affect each other. Therefore, we address the following question: How does tilt interaction affect task performance and subjective experience with Shoebox VR in a motion-based game? (Section 3.3)
- In order to further investigate the usefulness of Shoebox VR in games, we also analyse how Shoebox VR is used based on its role as an essential or non-essential gameplay element, since there are indications that the use of Shoebox VR differs between these roles (Section 3.1). Therefore, we state the following question: How does the role of Shoebox VR, either as an essential or non-essential part of the gameplay, affect the interaction with Shoebox VR? (Section 3.2)

In addition to our major contributions, which focus in depth on the effect of Shoebox VR in motion-based games, we made the following more general contributions concerning mobile VR and mobile usability evaluation:

- Contribution to mobile VR (Section 2.3): Despite significant improvements in performance, developing real-time 3D worlds on mobile devices remains a challenging problem. In an effort to create a general method for the visualization of 3D worlds by means of Shoebox VR, a code library was created. The details of this implementation and the challenges are described in this section, including the investigation and evaluation of a general solution for noise-filtering of the accelerometer data.
- Contribution to mobile usability evaluation (Section 2.4): As powerful mobile devices are a relatively new concept and the way they are used is different from more traditional devices, such as PCs, we must consider new and innovative ways to evaluate and study related applications. In order to determine the usability of certain techniques like Shoebox VR in real-life games used by a real-life audience, we examined remote evaluation possibilities, as applied in Experiment C (Section 3.3). Section 2.4 will give a detailed description of our experiences, challenges and suggestions for future research for the usage of such methods.

The remainder of this thesis is split in two parts. The first part describes the theoretical background, containing a related work section (Section 2.1), the Shoebox VR theory (Section 2.2), implementation details (Section 2.3) and evaluation details (Section 2.4). The second part of the thesis describes our investigation of several sub-problems described above, concerning the effect of Shoebox VR in motion-based games. First, the examination of the potential effect of screen size on depth perception and subjective experience is described (Section 3.1), followed by an investigation focusing on touch screen interactions in combination with tilting actions and the role that Shoebox VR plays as either an essential or non-essential gameplay element (Section 3.2). Then we describe the effect of Shoebox VR on tilt interaction (Section 3.3) and the comparison of different noise filter for accelerometers (Section 3.4). Finally, this thesis closes with the conclusions that can be drawn from our findings, their implications on Shoebox VR and mobile VR in general and suggestions for future work (Section 4).

2. Theoretical Background

2.1 Related Work

First, we will describe the research that has been done in an effort to define immersion and how to measure it, as this is an important factor that affects the subjective experience with 3D applications. Secondly, we describe why immersion is a critical issue on mobile devices, because of their form factor. Thirdly, literature on Virtual Reality (VR) is discussed and the possibilities for VR on mobile devices to increase immersion are explored. We also shed a light on previous research performed on Shoebox VR in particular and finally, we will describe research performed on how to evaluate usability issues for mobile applications, in particular remote evaluation, which was used as a motivation to remotely evaluate Shoebox VR usability in Experiment C (Section 3.3).

2.1.1 Immersion

In order to investigate the effects of Shoebox VR on immersion, we looked into existing research on methods for measuring immersion. However, in order to quantify immersion, it needs to be defined first. Even though the term "immersive" is often used as a descriptive term for applications, especially video games, there seems to be no consensus on a formal definition for the term. In fact, there seems to be debate about whether it describes an objective or subjective concept, as Slater *et. al.* (1996; 1999) have described immersion as "a quantifiable description of technology" that "includes the extent to which the computer displays are extensive, surrounding, inclusive, vivid and matching". Using this definition, a system like CAVE (Section 2.1.3) is more immersive than Fish Tank VR (Section 2.1.3) or Shoebox VR (Section 2.2), based on the fact that CAVE surrounds the user with screens and Fish Tank VR and Shoebox VR do not. However, Brown & Cairns (2004) have demonstrated that gamers commonly describe the term as a subjective experience, based on several factors, such as involvement, focus and sense of presence. In agreement with the latter description, immersion is defined as a subjective experience in this thesis.

Several methods for the measurement of immersion, and sense of presence in particular, have been introduced in previous studies. Some studies gathered qualitative data by means of questionnaires (Witmer & Singer, 1998; Lombard *et al.*, 2000), while others tried to gather quantitative data by measuring performance or physiological responses (Cox *et al.,* 2006; Jennett *et. al.* 2008). A detailed summary for many of these different methods and a description of their effectiveness have been provided by Van Baren & IJsselsteijn (2004). We considered the use of these existing methods for our own research, but finally decided that the methods based on physiological responses are too invasive, as it often requires specialized equipment that has a large effect on the used set-up. In addition, the resulting data requires certain expertises to interpret and analyse that we lack at this point. Alternatively, since Browns & Cairns (2004) showed that gamers tend to agree on the definition of immersion, we decided on a single item questionnaire that simply asks about the subject's sense of immersion, instead of the proposed larger, more complicated questionnaires.

2.1.2 Screen size

One of the challenges that mobile developers are presented with when creating 3D applications is to provide users with the same level of immersion as is provided with 3D applications on traditional platforms, such as PCs and gaming consoles. Immersion appears to be affected by screen size, as described by Hou et. al. (2012), who showed that players of video games tend to feel more immersed when playing on a large screen than when playing on a small one. This effect of screen size on immersion especially applies to mobile device users, since these devices commonly have small screens, due to their small form factor. The effective screen size is often reduced further by the implementation of on-screen controls, which requires users to occlude parts of the screen with their hands. In addition, a smaller screen has a smaller physical field of view (FOV), which has been shown to result in a reduced immersive experience (Prothero & Hoffman, 1995). The negative effects of a smaller physical FOV could be compensated for by using a software FOV (fig. 2) that is significantly greater than the physical FOV, but this will result in a distorted and unrealistic image.

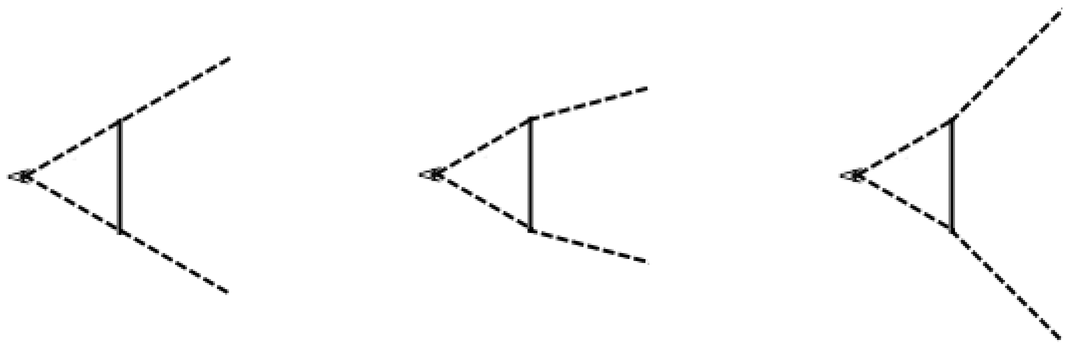

Fig. 2: The eye represents the user's eye position, the solid bar a screen. The left side of screen shows the real or physical FOV, whereas the right side shows software FOV. The left image, shows the case where physical FOV and software FOV are the same, whereas the middle and right image, show situations where the software FOV is smaller than and larger than the physical FOV, respectively.

2.1.3 Virtual Reality

In an effort to determine the effectiveness of mobile devices as platform for Virtual Reality (VR), Hwang et. al. (2006) have shown that the use of motion-based interaction on a mobile device can increase the perceived FOV and sense of immersion to the level of VR platforms with larger screens. Also, it has been shown that VR increases the sense of immersion (Pausch et al., 1997), which suggests that the implementation of a mobile VR technique, such as Shoebox VR, can compensate for decrease in immersion caused by small screen size and small FOV.

Several VR systems were already established before Shoebox VR was introduced, such as CAVE (Cruz-Neira *et. al.*, 1993) and Fish Tank VR (Ware *et. al.*, 1993). However, these systems are expensive and require a large amount of physical space and are therefore not viable solutions for mobile devices. Different adaptations of Fish Tank VR that are less expensive and require less physical space have been described. For example, Lee (2007) presented an adaptation which uses a controller for the Wii gaming console. Lee's approach is inexpensive and far more compact than the original Fish Tank VR set-up, but is not suitable for the use in mobile applications, due to its dependence on additional peripherals. Another possibility is to use the front-facing camera of a mobile device for camera-based head tracking, to implement Fish Tank VR for mobile devices. However, although this approach has been used before, it has two main issues. Firstly, it is computationally expensive to calculate head position from camera images. Secondly, head-tracking is often easily broken by environmental factors. For example, the camera will be unable to track the user's head in a poorly lit environment or the system might get confused when multiple heads are visible to the camera. In addition, the camera will not be able to track the user's head outside the field of view of the camera.

2.1.4 Shoebox VR

Both Fish Tank VR and Shoebox VR create a virtual reality experience by changing the camera perspective based on the user's head orientation with respect to the screen (Section 2.2.1). The difference between Shoebox VR and Fish Tank VR lays in how the change in perspective is accomplished. Fish Tank VR uses head-tracking to determine the user's head position with respect to the screen, whereas Shoebox VR uses the accelerometer to determine the current orientation of the mobile device and assumes a static head position (fig. 3).

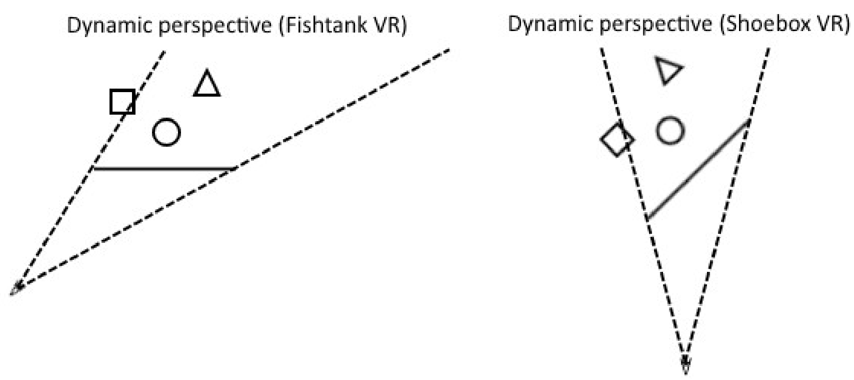

*Fig. 3: **Fish Tank VR and Shoebox VR.** Fish Tank VR determines the user's view perspective based on head-tracking, whereas Shoebox VR determines it based on the device's tilt angle.*

Since the Fish Tank VR approach does not assume the user's head position, but determines it by means of head-tracking, it results in a more realistic image, meaning that the calculated view perspective is more accurate. However, Shoebox VR does not require camera-based head tracking and therefore it does not suffer from the limitations that arise from camera-based head tracking concerning computational power and distortion by environmental factors. Importantly, some of the limitations currently seen for head tracking on mobile devices might be remedied in the near future, as the computational power of mobile devices continue to increase. Moreover, infra-red cameras enable head tracking in poorly lit environments and have already been implemented in the recently released Amazon Fire Phone. However, the addition of infra-red cameras increases the production costs and results in a larger form factor. In addition, the limited view angle of cameras remains an issue. Also, other aforementioned issues remain unsolved, as the system can still be easily confused by certain environmental factors such as additional heads captured by the camera, reducing its reliability. Therefore, we believe that Shoebox VR represents a viable option for realizing VR on mobile devices.

In order to determine how and in which situations Shoebox VR can effectively be implemented, previous studies have been performed. For example, Siewert (2012) investigated Shoebox VR with fixed motion games. However, this has not yet resulted in a general idea on how to effectively implement Shoebox VR for fixed motion games. Furthermore, we investigated the effect of Shoebox VR on depth estimation performance (Hürst & Beurskens, 2013) with fixed motion games. Although the results suggested no increase in depth estimation performance due to Shoebox VR, they did show that Shoebox VR can result in a better subjective user experience when compared to standard visualization. However, we have also observed that the extent of this effect is situational. Several possible variables have been investigated, such as the addition of forward motion of the virtual camera, which has been suggested to give a positive effect on the visual experience with Shoebox VR. Also, results suggested that faster camera speed results in a smaller average change in the device's tilt angle, implying less active use of perspective changing capabilities of Shoebox VR. Furthermore, the type of virtual environment has been classified as a variable that is of no effect so far, even though our original preliminary observations suggested otherwise. However, the effect on subjective Shoebox VR experience by camera motion and environment is merely the tip of the iceberg, as there are many more possible influencing factors that remain unexplored.

2.1.4 Accelerometer noise filtering

One of the factors that is likely to affect the Shoebox VR experience is how the data of the accelerometer is filtered, as the accelerometer is used to determine the device's tilt angle, which allows for the change in perspective. In order to use the accelerometer effectively for Shoebox VR (or any 3D visualization) its data needs to be filtered, as the accelerometer generates a large amount of noise and is extremely sensitive to small motions, such as the natural shaking of hands. Several filters have been suggested in previous research, for example, Liang *et al.* (1991) and Friedmann *et. al.* (1992) demonstrated methods to filter noise from positional and orientation data and counteract lag by using a Kalman filter (Kalman, 1960), whereas Condant (2013) described an algorithm for Total Variation de-noising. In our experiments, we have used similar approaches for noise filtering. Our investigation and experiences on this subject are described in Section 2.2.3.

2.1.5 Mobile usability evaluation

It is common to conduct mobile usability studies in a controlled lab study, as it is relatively easy to avoid unwanted influences that might affect the results in such a setting. However, using this method has disadvantages as well. For example, the way subjects use the tested technique or interface in a controlled environment, such as a lab, might not be representative for how it is used in a real-world situation. An alternative to local controlled lab studies, is having subjects participate in studies remotely.

With the introduction of application store for mobile devices, such as Apple's App Store and Google Play store, it has become relatively easy to distribute a usability test to a large audience. Henze *et al.* (2011; 2012) already have performed a study in this regard, showing that distributing a remote experiment through the Google Play store (which was named Android Market at that point in time) can provide results with high external validity, meaning that the setting in which the experiment is performed, is representative to a real-life situation. This comes at the cost of a lower internal validity, which means that the chance that differences in the results are more likely due to other factors than the tested independent variable. Another factor that cannot be easily controlled for when distributing an experimental application through application stores, is the amount of people that participate in the experiment. Although Henze *et al.* (2011; 2012) described the large number of times downloaded for their apps, they did not describe their process for reaching this amount of attention.

For Experiment C (Section 3.3) we wanted to determine how Shoebox VR was used in a realistic situation and if it would benefit from a set-up that allows for collecting data that has a high external validity. In this light, Experiment C follows a similar set-up as used by Henze *et al.* (2011; 2012), as we released an application in Google Play. Our experiences with this approach are described in Section 2.3.

2.2 Shoebox VR

2.2.1 Theory

To provide a viable VR solution for mobile devices, Hürst & Helder (2011) have introduced the shoebox visualization concept, which was later more aptly renamed to Shoebox VR (Van Laar, 2014). Shoebox VR allows the virtual camera perspective to change based on the user's head position relative to the screen (dynamic perspective) by measuring the device orientation, instead of using the standard visualization technique, which generates an image based on the assumptions that the user is sitting directly in front of the screen at a fixed distance and that the orientation of the display is fixed (static perspective). Using static perspective, the image remains unchanged as user's head position changes with respect to the screen, which essentially results in an incorrect and distorted image. The difference between static perspective and dynamic perspective is illustrated in Fig. 1.

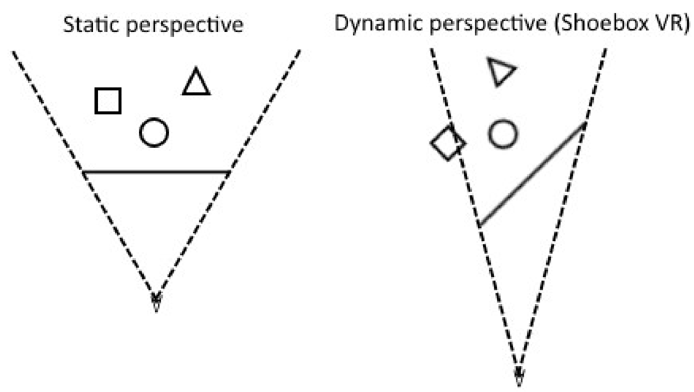

*Fig. 1: **Static versus dynamic perspective.** Standard desktop visualization of a 3D world always shows an image which does not consider the user's head position with respect to the display (static perspective). Shoebox VR will allow the image to change based on the user's head position with respect to the display (dynamic perspective), based on the orientation of the device.*

2.2.2 Implementation

In order to create a general solution for all our Shoebox VR related experiments, a Java library was implemented in collaboration with Marco van Laar (Van Laar, 2014). The implementation is based on the concept of off-axis projection. Standard visualization techniques generally produce an image based on the concept of on-axis projection, which creates an image based on the assumption that the user is focussing on the center of the screen, viewing it at a right angle, whereas off-axis projection uses a view perspective away from this axis, also allowing a view from a different angle as is illustrated in figure 4.

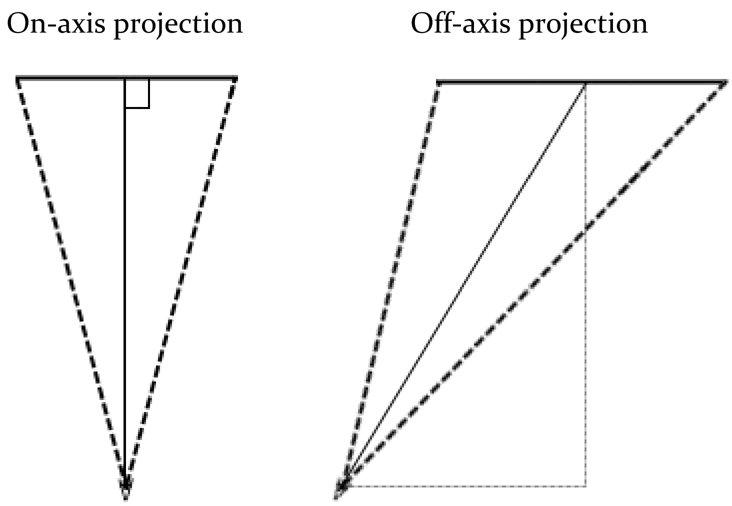

Fig. 4: On-axis projection compared to off-axis projection.

In order to accomplish off-axis projection in our implementation, OpenGL is used to create a skewed view frustum based on the tilt angle of the device as measured by the accelerometer. However, this is not sufficient to create the Shoebox VR effect. By creating a skewed frustum, the user's change in head position with respect to the screen is not yet taken into account. Addressing this issue, the camera position is changed by rotating around a point in front of the camera. The distance of this point is determined by a numeric variable, allowing to change the scale of the rotation. Van Laar later changed this to a translation away from the axis, instead of rotation around a point, because the rotation would change the camera position on the relative z-axis, possibly moving it through or behind objects, as if the user was inside the virtual world, which obviously is incorrect. While Van Laar's argument is valid, we argue that his approach moves the camera farther from the user's point of focus in the virtual environment, while the user's physical head distance from the centre of the screen, does not change when tilting the device. However, at this point this is a difference in opinion as there is no proof that one of these approaches has a more negative effect over the other. In fact, neither approach has shown any immediately clear negative effect on the Shoebox VR experience.

2.2.3 Noise filter implementation

In order to filter the noise from the accelerometer measurements, several different filters have been implemented. In our previous study, accelerometer data was filtered using a simple moving average (SMA) approach. This method has the advantage that the implementation is simple and the computational cost is trivial. However, the drawback of this approach is that it is difficult to find a balance between a smooth signal and a small delay between physical rotation and change in perspective. By increasing the number of measurements that is used for averaging, the resulting image becomes smoother, but the delay increases and sudden quick tilting motions are not registered as such. In contrast, by reducing the number of measurements, the image changes more responsively, but it results in a less stable image. In summary, the SMA approach tends to either result in too much delay or in an unstable image and indeed, our implementation used in our previous research showed a smooth image, but subjects complained about too much delay.

In an attempt to solve the noise problem that occurred in our previous experiments we implemented a Kalman filter, resulting in an image that is both smooth and responsive. This implementation was used with success for Experiment A (Section 3.1) and the noise filter problem was deemed solved. However, for Experiment B (Section 3.2), the Kalman filter did not seem sufficient, as complaints emerged about the image stability.

To tackle this recurring issue we investigated whether a general solution could be found. On that account, a small scale experiment was performed that compared the noise filtering methods that were used so far and compared them with a new candidate, the Total Variation denoiser (Section 3.4). The results of this experiment showed that the Total Variation denoiser is the overall best choice, as it results in the most stable image and the responsiveness is considered to be acceptable. However, the Kalman filter seems to have a slightly better responsiveness. Therefore, the Kalman filter might be a better option in situations where maximum responsiveness is preferred.

2.3 Mobile usability evaluation

In order to determine the real-life usability of Shoebox VR, we investigated the possibilities of remote testing. Based on the research described in the related work section (Section 2.1.5), we implemented the application for experiment C (Section 3.3) to be remotely tested by having users download it through the Google Play store.

The application that was released for Experiment C was installed 37 times. The difference in number of installs compared to Henze *et al.* (2011, 2012) might be due to the fact that they had more resources available to promote their application on a large scale or might be simply due to the fact that the landscape of App stores have changed since then, in the sense that these stores are now far more saturated with large amounts of apps, reducing their discoverability. We used the following means to promote our application:

- Promotion on several (popular) Facebook pages
- Mailing to students
- Mailing to friends and family

While we are satisfied with the number of installs that was reached, releasing an application on an App Store is not necessarily an easy solution for reaching a large number of subjects. Although it has the advantage of creating a more accessible study, it requires a serious amount of effort in polishing the application such that it meets the quality that a general audience demands. In addition, it appeared to be difficult to get the number of installs that is required for the particular study. Finally, the data gained from a remote study should be handled carefully and most likely requires some form of filtering.

Whether remote studies should be used depends on several factors. Henze *et al.* (2011, 20012) mention several issues, such as that the subject sample can not be generalized to the world population, that the acquisition of valuable qualitative data is difficult and that additional care should be taken when analysing the data. Advantages they mention are that remote evaluation results in a high external validity and that they were able to obtain data which would have required far more effort when using local studies.

For the most part, we agree with Henze *et al.* However, they fail to mention the effort it would required to reach the number of subjects they did. Based on our own experience, we speculate that the effort required is quite substantial. This in addition to the effort it requires to polish the application for remote evaluation and process the gathered data. Therefore, we doubt whether the effort required for a remote study is indeed less than for a large scale local study.

In conclusion, if testing usability for realistic situations, remote evaluation through App stores can be a viable option. However, when deciding on whether or not to use remote evaluation, the issues concerning reduced internal validity and the difficulty in generalizing the subject sample to to the population should be considered. For example, if the study focusses on obtaining qualitative data, remote evaluation is not a good option. Also, we want to add that the available resources should also be considered. The effort that is required to polish and market the application and possibly the handling of data is possibly substantial. If time and manpower is limited, constructing a local study might be a better use of resources.

3. Experiments

Our major goal for this thesis is to expand on our insights on which factors have an effect on depth perception and subjective Shoebox VR experience. Firstly, we determined whether the screen size of a device has an effect on depth estimation and subjective user experience when using Shoebox VR, as this would influence future research that makes use of devices of different sizes. Secondly, we wanted to confirm our previous findings on Shoebox VR usage, that suggested that the average change in perspective using Shoebox VR is different in certain situations, by determining the average change in tilt angle, as our previous findings on this subject were based on limited data. Thirdly, we investigated if and how user interaction affects the Shoebox VR experience. The focus here is on touch screen interaction, since this is the most common form of interaction with mobile applications. Finally, based on our findings, we further explored the effect of a different form of interaction, being tilting controls. In addition, the game designed for this experiment was released on Google Play, allowing us to investigate gameplay performance and subjective experience of Shoebox VR compared to standard visualization, as experienced by a realistic audience in an uncontrolled study.

3.1 Experiment A: The Effect of Screen Size

Our previous study (Hürst & Beurskens, 2013) gave us a better understanding of the benefits of Shoebox VR in terms of depth perception performance, subjective sense of depth perception, entertainment value and immersion. We concluded that using Shoebox VR provided no benefit in terms of depth perception performance when compared to standard visualization. However, the subjective experience was significantly improved by Shoebox VR. Users claimed to have a better subjective sense of depth perception and they graded the entertainment value and sense of immersion significantly higher with Shoebox VR than with standard visualization. In addition, increasing the camera speed resulted in a decrease in the average change in tilt angle of the device, which implicated a less active use of Shoebox VR, as movement speed increases. However, this conclusion is based on limited data and therefore requires additional research.

Based on the data gathered in our previous study, we could not investigate the effect of screen size on depth perception and user experience with Shoebox VR, since all subjects performed the experiment with the same device. Therefore, in order to determine the effect of screen size on depth perception performance, subjective sense of depth perception, entertainment value and immersion, this study uses an adapted version of the game we used in our previous study to run on different devices with different screen sizes. In addition, we investigated whether the statement we made about objective depth estimation performance holds true for different screen sizes, confirming that there is no depth estimation performance benefit gained from Shoebox VR. Finally, we investigated in more detail whether camera speed has an effect on Shoebox VR usage, as previous results suggested.

3.1.1 Hypotheses

To investigate the issue of screen size the following question was stated:

How does screen size affect depth perception and subjective experience with Shoebox VR and how does it affect usability of Shoebox VR in a motion-based game?

In order to answer this question in terms of depth perception we tested the following hypotheses:

HA1: Screen size **does not** affect depth perception performance differently for Shoebox VR than for standard visualization.

HA2: Screen size and speed **do not** affect depth perception performance differently for Shoebox VR than for standard visualization.

HA3: Screen size affects subjective sense of depth perception[1] differently with Shoebox VR than with standard visualization.

In order to answer this question in terms of subjective experience we tested the following hypotheses:

HA4: Screen size affects entertainment value[1] differently with Shoebox VR than with standard visualization.

HA5: Screen size affects immersion[1] differently with Shoebox VR than with standard visualization.

In order to investigate Shoebox VR usability, the following hypotheses were tested:

HA6: Screen size affects average change in tilting angle with Shoebox VR.

HA7: Speed affects average change in tilting angle with Shoebox VR.

HA8: Screen size affects average change in titling angle with Shoebox VR differently between screen sizes.

As described in the related work section (Section 2.1.2), screen size affects the subjective experience, specifically immersion. For this reason, we expect that there will be a different effect for subjective experience (**HA3**, **HA4**, **HA5**), but objective performance will be unaffected by screen size (**HA1**, **HA2**). However, even though the change in tilt angle is an objective measure (**HA6**, **HA7**, **HA8**), we hypothesize that this will be different, due to the difference in physical characteristics. For example, we think a larger and heavier device might make it less appealing to tilt the device.

3.1.2 Design: Alignment Game

In order to test our hypotheses, the game used for our previous research, referred to as the alignment game, was adapted. Here, the original design of the experiment is described first,

1 Quantified via averaged user ratings.

followed by the adaptations made for the new experiment. Finally, possible issues are described together with the implementation details that allowed us to avoid them.

The game's scene consists of a tunnel, a blue cube in the centre of the screen and a blue frame, which functions as an additional reference point. The frame is aligned with the blue cube on the z-axis (which points outwards of the display). The tunnel, blue cube and blue border are static and thus do not move. Yellow cubes (targets) appear at the far plane and move along the z-axis towards the screen, as illustrated in Figure 5.

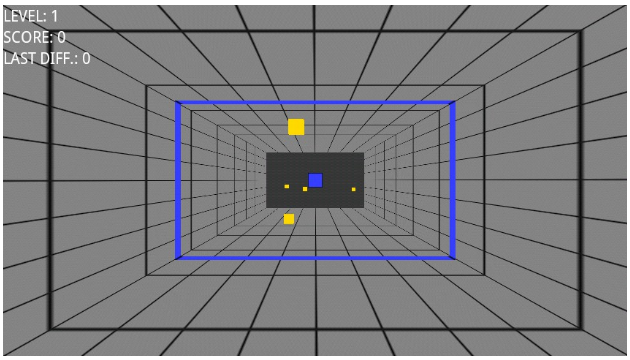

Fig. 5: The virtual environment of the game used in experiment A.

The player's goal is to tap the screen each time a target aligns with the blue cube (and the reference frame) to score points. The game consists of multiple levels and a next level is reached each time 8 targets have passed the virtual camera. With each level increase the targets move faster towards the screen. If players are unable to hit any targets for the duration of 1 level, the game is over. There is an infinite number of levels, as we did not want to assume an upper bound value for speed at which the game is still "playable". This design choice has effect on the way the data should be analysed, which is described in the discussion (Section 3.1.6)

For analysis we logged the following data:

- The depth estimation error on each tap.
- The tilting angle of the device at small time intervals.

In order to investigate the effect of screen size, the following adaptations were made to the original design:

- This experiment was conducted on three devices with different screen sizes (4", 7" and 10") in order to be able to compare the differences in screen size.
- Our previous experiment consisted of three different virtual environments, in order to determine the effect of the type of virtual environment. The three different environment types we defined were "closed", "open" and "none", representing indoor environments, outdoor environments and environments without surface, such as outer space and underwater, respectively. Since our previous study showed that the different environment types did not affect the Shoebox VR experience, only the "closed" environment was used in the current study, because it was considered to be most appealing, in order to control the experiment's complexity.
- In order to gather more data and acquire data that is less biased towards the particular situation of a screen tap, the tilt angle of the device was logged at small time intervals, whereas in our previous experiment this was only logged at each screen tap.
- The far plane will be placed farther from the camera than in our previous study. The reason for this is that subjects commented about the previous experiment, that at very fast speeds they tapped the screen as soon as a target appeared at the far plane, instead of focussing on the moment of alignment. By replacing the far plane, we attempted to avoid this issue as much as possible.

In order to avoid possible negative effects on the validity of our results, the following precautions were taken in the game's implementation:

- The targets are pseudo-randomly placed on the far plane, such that they never collide with the vehicle and so that targets do not often appear in a similar location twice, which is assured by creating a grid (for the x and y axes) and making sure that a yellow cube does not spawn in a section of this grid before yellow cubes have spawned in the other sections.
- The targets have different sizes, so that subjects are unable to determine the moment of alignment based on the target's size.
- To make sure the user cannot get a high score by tapping the screen constantly at a high rate, the user is only allowed to fire once for each target. As soon as the vehicle is passed

halfway towards the next target, one shot is possible for that next target and the opportunity for firing at the previous target has passed.

- The subject gets feedback on the result of a tap. A sound is played when the energy field is fired. If the user cannot fire, because it has already attempted this for the current target, a buzzing sound is played. If the shot is a hit, the target turns green and a distinct sound is played.
- Targets are placed 1 to 3 seconds apart from each other. The actual value is a randomized float within this range, so there is no predictable pattern. By defining the distance as time, instead of spatial units, we made sure the targets do not follow up on each other too quickly when playing the game at higher levels.
- A rotation limit is applied to Shoebox VR, such that the subject will always see the entire contents of the tunnel and will not occlude any targets by tilting. This is necessary for a fair comparison, as occlusion of this type will never occur using the standard visualization.

3.1.3 Procedure

Before playing the game, subjects were asked to fill out a questionnaire concerning their age, gender and their previous experience with video games, mobile games and mobile games that use motion controls. This was followed by an in-game tutorial which explained the Shoebox VR concept and the goal of the game. The tutorial consisted of a set of instructions that demonstrated the controls and goals of the game. After finishing the tutorial, the subjects played the game two times on each device, once using the standard visualization and once using Shoebox VR. Each subject played on three different devices, a 4”, 7” and 10” device.

To control for order effects, the order in which the subjects played on the devices was counterbalanced, such that each possible order in the devices are played on occurred an equal number of times. Afterwards, the subjects were asked to fill out a questionnaire concerning their experience with Shoebox VR on devices with different screen sizes.

3.1.4 Subjects

A number of 24 subjects participated in the experiment. All subjects were Computer Science Master’s students at the University Utrecht and were required to participate as part of a course. During this course they were made aware of what Shoebox VR is and how it works. Details about previous experience of subjects are shown in Table 2 and Table 3. The subjects did not receive payment for their participation.

Table 2: Age, gender and ownership of tablet sizes of the subjects.

Average age	24 (SD 2.5)
Gender	100% male
Owned a mobile device	91,66%
Owned a 4” mobile device	87,5%
Owned a 7” mobile device	8,33%
Owned a 10” mobile device	12,5%

Table 3: Game experience of the subjects

	Video games	Video games on mobile devices	Video games on mobile devices with motion controls
Daily	58,33%	12,5%	12,5%
Weekly	25%	16,67%	12,5%
Monthly	4,17%	0%	0%
Yearly	8,33%	37,5%	37,5%
Never	4,17%	33,33%	37,5%

3.1.5 Results

In order to verify the stated hypotheses, with the goal of answering our earlier stated sub-question “*How does screen size affect depth perception and subjective experience with Shoebox VR and how does it affect usability of Shoebox VR in a motion-based game?*”, the answers given to the questionnaires and the data logged during gameplay were analysed.

Depth estimation: performance

To determine the effect of screen size on depth estimation performance using Shoebox VR, we stated the following hypotheses:

HA1: Screen size **does not** affect depth perception performance differently for Shoebox VR than for standard visualization.

HA2: Screen size and speed **do not** affect depth perception performance differently for Shoebox VR than for standard visualization.

In order to test **HA1**, the average depth estimation errors were calculated for each screen size for the lowest speed and were compared between screen size and visualization type (fig. 6). The data for other speeds was excluded because it was not available for all subjects as not all subjects played all levels and using this data would result in less reliable results. More on this in the discussion (Section 3.1.6).

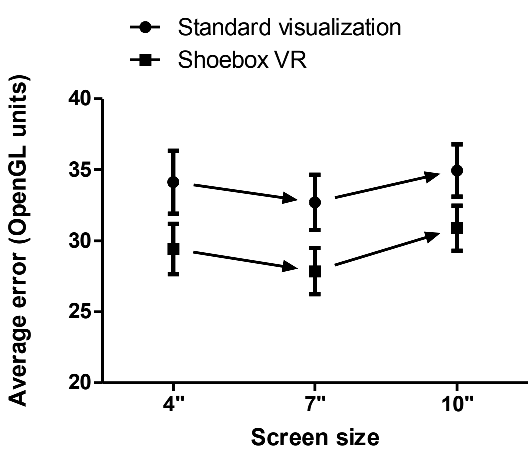

Fig. 6: The average depth estimation error for different screen sizes using different visualization methods. Error bars depict the standard error of the mean.

Statistical analysis was performed by means of a two-way repeated measures ANOVA test, where the screen size and visualization type were the independent variables and the average depth estimation error was the dependent value. The analysis showed that there was no significant effect due to interaction between screen size and visualization type, meaning that there is no difference between Shoebox VR and standard visualization in terms of depth perception performance for different screen sizes. Therefore, we conclude that **HA1** holds true.

In order to verify **HA2**, the data was filtered in such a way that subjects that did not reach level 5 were removed from the data set (fig. 7, 8, 9). This was necessary to avoid gaps in the data pairs. Please take note however that this might effect the validity of the results, as removing data points results in an imperfect order balance. This will be further discussed in the discussion section (Section 3.1.6).

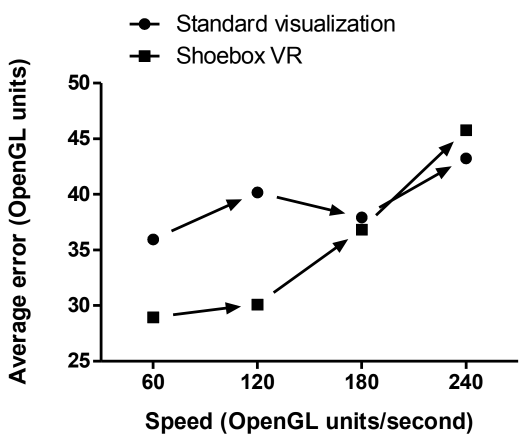

Fig. 7: The average depth estimation error for different visualization types and motion speeds on the 4" device.

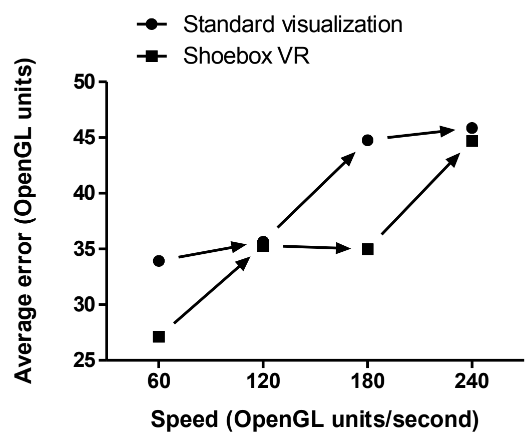

Fig. 8: The average depth estimation error for different visualization types and motion speeds on the 7" device.

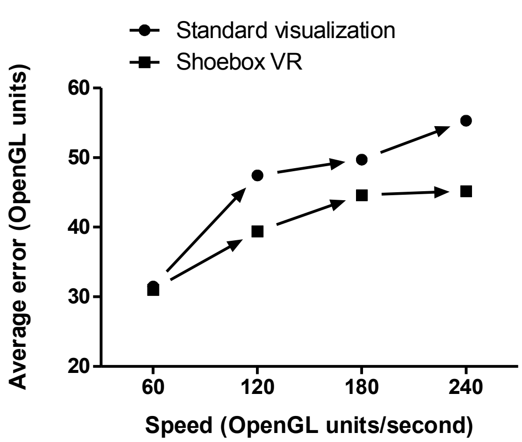

Fig. 9: The average depth estimation error for different visualization types and motion speeds on

the 10" device.

Paired t-tests were performed on average depth estimation error between standard visualization and Shoebox VR for each combination of speed and screen size. Analysis shows that the difference in average estimation error is significant between standard visualization and Shoebox VR for the 4" device for 60 and 120 OpenGL units/s motion speeds. This suggests that for smaller devices, at lower speeds, Shoebox VR has positive effect on depth estimation performance. For higher speeds the difference was not significant and neither were any differences found for the 7" and 10" devices. The p-values are shown in Table 4. These results suggest that Shoebox VR has a positive effect on depth estimation on smaller devices at low speeds, in the sense that the average estimation error is smaller. Therefore, we reject **HA2**.

Table 4: P-values for t-test comparisons between standard visualization and Shoebox VR for each combination of speed and screen size. P-values < 0.05 are considered significant and depicted with a ''.*

Speed	Screen size	P-value
60	4"	0.0332*
60	7"	0.0576
60	10"	0.8086
120	4"	0.0203*
120	7"	0.9436
120	10"	0.2379
180	4"	0.8415
180	7"	0.1080
180	10"	0.2055
240	4"	0.5446
240	7"	0.8537
240	10"	0.1134

Depth estimation: subjective experience

To determine the effect of screen size on the user's subjective sense of depth perception we stated the following hypothesis:

HA3: Screen size affects subjective sense of depth perception differently for Shoebox VR than for standard visualization.

In order to test **HA3**, subjects were asked to rank how difficult they felt it was to estimate depth for each combination of visualization type and screen size, where 1 was the easiest and 6 the hardest (Fig. 10). Statistical analysis was performed by using a two-way repeated-measures ANOVA test, in which visualization type and screen size were the two independent variables and the rank for depth estimation difficulty was the dependent variable. Analysis showed that the interaction between visualization type and screen size was not significant (p=0.7152). However, visualization type showed a significant effect (p=0.0001), on average ranking for depth estimation difficulty, whereas screen size did not (p=0.0617).

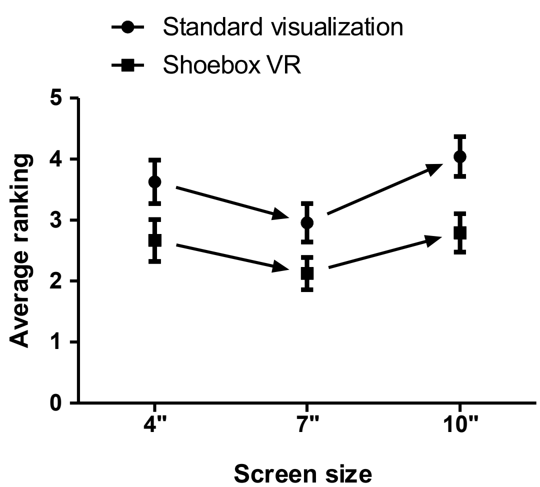

Fig. 10: The average ranking for subjective depth estimation difficulty for different screen sizes using different visualization methods. Ranking goes from easiest (0) to hardest (6). Error bars depict the standard error of the mean.

Based on the analysis, **HA3** was rejected, as the subjective sense of depth perception appeared to be the same between standard visualization and Shoebox VR for different screen sizes. In addition, we conclude that visualization type has an effect on the sense of depth perception, showing that Shoebox VR can be favoured over standard visualization in terms of subjective sense of depth perception, which confirms the findings of our previous study. For screen size a trend can be seen that 7” is favoured over 4”, which in turn is favoured over 10”, but we were unable to confirm statistical significance for this trend.

Entertainment value

To determine the effect of screen size on entertainment value when using Shoebox VR, we stated the following hypothesis:

HA4: Screen size affects entertainment value differently with Shoebox VR than with standard visualization.

In order to test **HA4**, subjects were asked to rank each combination of visualization type and screen size, 1 being the most entertaining and 6 being the least entertaining (Fig. 11). Statistical analysis was performed by means of a two-way repeated-measures ANOVA test, where visualization type and screen size were the two independent variables and entertainment value was the dependent variable. Analysis showed that the interaction between visualization type and screen size was not significant (p=0.6008). However, visualization type showed a significant effect on entertainment value (p=0.0001), whereas screen size did not (p=0.1241).

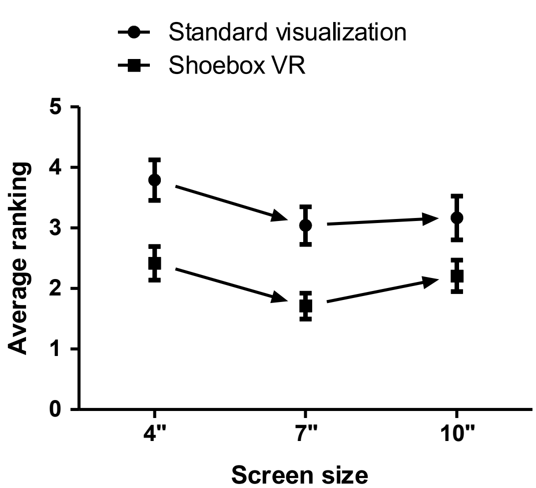

Fig. 11: The average ranking for entertainment value for different screen sizes using different visualization methods. Ranking goes from most entertaining (0) to least entertaining (5). Error bars depict the standard error of the mean.

Based on the analysis, **HA4** was rejected, as the entertainment value appeared to be the same between standard visualization and Shoebox VR for different screen sizes. Also, we conclude that Shoebox VR provided a better entertainment value than standard visualization, which confirms the findings of our previous study. Finally, for the screen size a trend can be seen where 7" is preferred over 10" which is preferred over 4" in terms of entertainment value, but this trend showed no statistical significance.

Immersion

To determine the influence of screen size on immersion when using Shoebox VR, we stated the following hypothesis:

HA5: Screen size affects immersion differently with Shoebox VR than with standard visualization.

In order to verify **HA5**, subjects were asked to rank each combination of visualization type and screen size, 1 being the most immersive and 6 being the least immersive (fig. 12). Statistical analysis was performed by means of a two-way repeated-measures ANOVA test, where visualization type and screen size were the two independent variables and immersion was the dependent variable. Analysis showed that the interaction between visualization type and screen size was not significant ($p=0.9433$) and that visualization type had a significant effect ($p=0.0001$) on immersion, while screen size did not ($p=0.0527$).

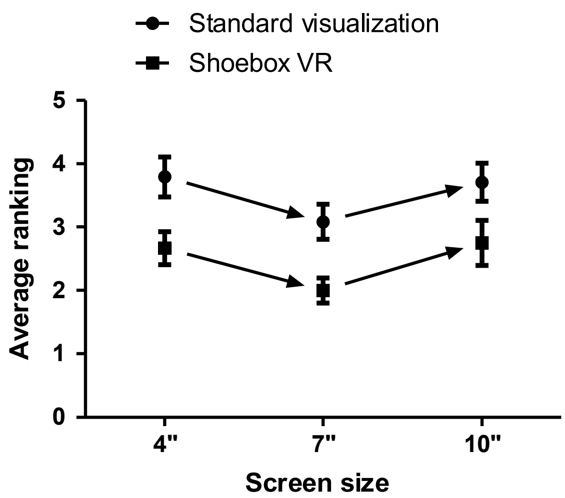

Fig. 12: The average ranking for immersion for different screen sizes using different visualization methods. Ranking goes from most immersive (1) to least immersive (6). Error bars show the standard error of mean.

Based on the analysis, **HA5** was rejected, as the immersion was rated the same for standard visualization as for Shoebox VR using different screen sizes. In addition, we conclude that visualization type can have a significant effect on the sense of immersion, as Shoebox VR increased the sense of immersion, in agreements with our expectations based on our previous results. For screen size, a preference for 7" over 10" and 4", in terms of immersion can be seen, but statistical analysis showed no significant differences.

Average rotation change

To determine the effect of different speed and screen size on average orientation, we stated the following hypotheses:

HA6: Screen size affects average change in tilting angle with Shoebox VR.

HA7: Speed affects average change in tilting angle with Shoebox VR.

HA8: Screen size affects average change in titling angle with Shoebox VR differently between screen sizes.

In order to test **HA6**, the device's tilt angle was recorded during gameplay each tenth of a second, which allowed us to determine the change in tilt angle by calculating the difference between each measurement. The averages of change in tilt angle for each combination of screen size and subject were compared for the lowest speed setting (fig. 13). The data for other speeds was excluded because it contained gaps, as was already described for **HA1**. We will further elaborate on this in the discussion (Section 3.1.6).

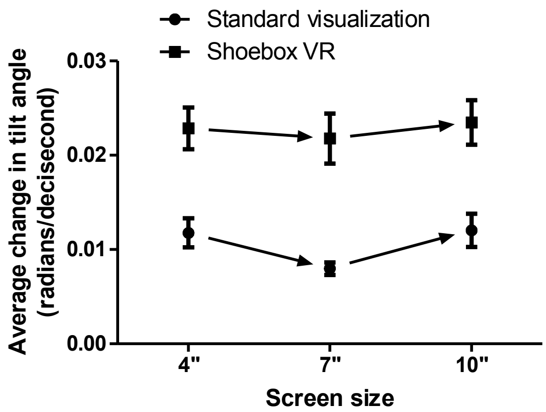

Fig. 13: The average change in tilt angle for different screen sizes using different visualization methods. Standard visualization is displayed as a baseline for comparison.

Statistical analysis was performed by means of a paired t-test for the average change in tilt angle between 4” and 10” for Shoebox VR, showing no significant difference (p=0.8480). Since this is the pair with the largest difference, we can assume that the difference between other possible pairs will not be significant either. Based on this analysis, we conclude that screen size does not have any effect on average change in tilt angle and therefore reject **HA6**.

In order to test **HA7**, the data was filtered in such a way that subjects that did not reach level 5 were removed from the data set to avoid gaps, as was described for **HA2**. Also any data for speed setting larger than 240 OpenGL units/second was excluded (fig 14), for the same reason. More about this in the discussion (Section 3.1.6). Analysis was done by performing a two-way repeated measures ANOVA test, using screen size and speed as the independent variables and the average change in tilt angle as the dependent angle. The analysis showed no statistical significance for interaction between screen size and speed (p=0.9988). In addition, no statistical significance could be determined for the screen size (p=0.1646) or speed (p=0.9398).

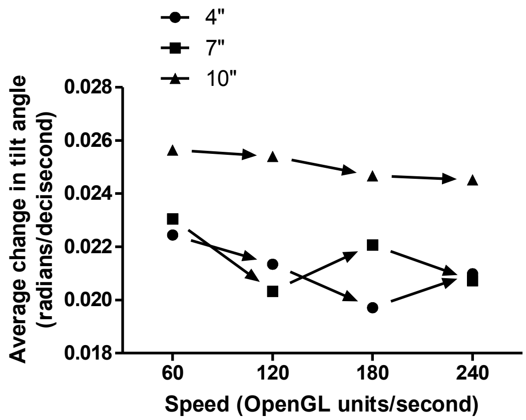

Fig. 14: The average change in orientation per target speed for each screen size.

Based on the analysis results, that showed no significant interaction between the independent variables, we conclude that the change in tilt angle does not differ between different screen sizes for Shoebox VR and therefore we reject **HA7**.

Based on the rejection of both **HA6** and **HA7**, we also reject **HA8**, as we do not see any effect on change in tilt angle for neither screen size nor speed.

3.1.6 Conclusion & Discussion

In order to gain more insight in the effects of screen size on Shoebox VR depth perception and subjective experience, with the goal of contributing to the knowledge of usability of Shoebox VR in motion-based games, we stated the following question:

How does screen size affect depth perception and subjective experience with Shoebox VR and how does it affect usability of Shoebox VR in a motion-based game?

In order to explore this question the following hypotheses were stated:

HA1: Screen size **does not** affect depth perception performance differently for Shoebox VR than for standard visualization.

HA2: Screen size and speed **do not** affect depth perception performance differently for Shoebox VR than for standard visualization.

HA3: Screen size affects subjective sense of depth perception differently with Shoebox VR than with standard visualization.

HA4: Screen size affects entertainment value differently with Shoebox VR than with standard visualization.

HA5: Screen size affects immersion differently with Shoebox VR than with standard visualization.

HA6: Screen size affects average change in tilting angle with Shoebox VR.

HA7: Speed affects average change in tilting angle with Shoebox VR.

HA8: Screen size affects average change in titling angle with Shoebox VR differently between screen sizes.

Based on our analysis, we accept **HA1**, showing that our expectations for the effect of screen size on objective depth perception performance hold true, meaning that different screen sizes have no overall advantage or disadvantage on depth perception performance when using Shoebox VR. However, when looking into different speeds, **HA2** did not hold true, showing that when different motions come into play, a significant difference is seen in depth perception performance at lower speeds for 4", but not for 7" and 10". This suggests that on lower speeds with small screen sizes Shoebox VR provides a depth perception performance benefit.

Furthermore, **HA3**, **HA4**, and **HA5** were rejected, which means that, as in agreement with our previous study, a significant benefit was demonstrated on the subjective depth perception, entertainment value and immersion by using Shoebox VR. However, this benefit was the same for all screen sizes. This suggests that the subjective benefit gained through Shoebox VR is equal for different screen sizes.

Finally, **HA6**, **HA7** and **HA8** were also rejected, which implies that screen size nor speed has an effect on how actively Shoebox VR is used.

In conclusion, our data showed that the subjective experience with Shoebox VR is not affected by screen size. However, depth perception performance with Shoebox VR can be affected by screen size when considering different motion speeds. Slower motions can have a benefit in terms of depth perception when using Shoebox VR on a small screen. Also, the lack of difference in average change of tilt angle indicates that the way that Shoebox VR is used is the same for different screen sizes.

In experiment A, the choice was made to not have a set number of levels that each subject should play. Because of this we were able to determine at which speeds the game was still "playable" and did not have to assume an upper bound for this. However, a drawback of this design choice, is that some of the subjects failed at a level sooner than we expected and thus did not supply data for levels that other subjects did reach. The consequence of this is that the data needed to be filtered, such that there were no missing values. This meant that if we wanted to check data for all subjects we were limited to analysing data for the first level, as was done for **HA1** and **HA6**, as some subjects did not reach higher levels. In order to still be able to look into higher values for speed and the difference between those values, subjects were excluded, leaving a smaller group. In this context, we excluded data from subjects that did not reach level 5 in any of the cases. This boundary was picked as it gave the best balance between a reasonable amount of subjects (n=14) and reasonable number of different values for speed. Also, it should be noted that the order balance is disturbed by the latter filtering method, meaning that learning effects might negatively influence the validity of the data, although the impact of this is hard to determine.

As a result of our data filtering we were unable to effectively analyse the data gathered at higher speeds. Now that the upper bound for a "playable" speed can be approximated based on our results, any future similar experiments could benefit from setting a constant number of levels that should be played by each subject, so that there will be no missing data for higher speeds and thus more insight can be gained for the effect of screen size at higher speeds.

For future work, it can be assumed that the benefit of Shoebox VR on subjective experience and the way Shoebox VR is used is the same for devices of different sizes. However, the effect of motion speed on depth perception performance should be further investigated, as results suggest that the difference in depth perception performance between standard visualization and Shoebox VR decrease as speed increases. In addition, if the maximization of depth perception performance is important, it would useful to approximate the "tipping point" for both speed and screen size, using smaller differences to determine at which screen size and values for speed this effect occurs. In conclusion, by investigating the effects of screen size on applications that use Shoebox VR, we demonstrated that subjective experience is not affected by screen size, but when considering different motion speeds, the objective depth estimation performance is. These insights allow for a more effective implementation of Shoebox VR for motion-based games that should support devices of different sizes.

3.2 Experiment B: Touch Interaction with Shoebox VR

In experiment A, we showed that screen size does not affect subjective experience differently when using Shoebox VR compared to standard visualization. The same is the case for the average change in tilt angle. Although we saw a benefit for depth perception performance on smaller screens at smaller speeds, but this does not seem to generalize to larger screens and higher speeds. Therefore, experiment B focused on the more general benefit of Shoebox VR for subjective experience in games. As interaction is one of the most critical aspects of games and could possibly have an effect on the Shoebox VR experience, this experiment investigated the role of interaction, specifically touch screen interaction as it is the most common form of interaction on mobile devices.

Based on this focus, we stated the following question:

How does touch screen interaction affect task performance and subjective experience with Shoebox VR in a motion-based game?

In order to address this question, we tested the following hypotheses:

HB1: Shoebox VR negatively affects performance involving touch screen interaction.

HB2: Shoebox VR negatively affects entertainment value when used on a touch screen-based application.

HB3: Shoebox VR negatively affects control intuitiveness when used on a touch screen-based application.

HB4: The negative effect of Shoebox VR on performance and subjective experience becomes larger as speed increases.

A negative effect for the task performance is expected, because the controls of an application get more complicated, due to the fact that the implementation of Shoebox VR also introduces a form of tilting controls, possibly to the point that it creates cognitive overload, which can be defined as a situation in which the amount of information and interactions that need to be processed exceeds the amount of working memory that one possesses. This in turn could lead to frustration and thus to a worse subjective experience and the increase in speed could further complicate this issue.

Another factor that might affect the interaction using Shoebox VR, is whether Shoebox VR is an essential or non-essential part of the game. Based on observations of our previous research, we noticed that not all subjects changed camera perspective often. Some subjects reported that they did not change camera perspective, because it was not necessary to do so. To elucidate whether the effect of Shoebox VR on interaction is dependent on the use of Shoebox VR as an essential or non-essential part of the game, we stated the following question:

How does the role of Shoebox VR, meaning either an essential part or non-essential part of the gameplay, effect the interaction with Shoebox VR?

In order to address this question, we stated the following hypothesis:

HB5: The average change in tilt angle is higher when Shoebox VR is implemented as an essential gameplay element than as a non-essential gameplay element.

The predicted difference is based on our observations and comments from previous experiments.

To test our hypotheses a game was created, consisting of three variations, one using standard visualization, one in which Shoebox VR was implemented as an essential part of the game play and one where Shoebox VR was non-essential. This allowed us to compare both the touch interaction using standard visualization with touch interaction using Shoebox VR and the effects of the type of gameplay element on role of Shoebox VR.

3.2.1 Design: Endless runner game

In order to test the stated hypotheses, an endless runner game was created. The endless runner game type is a type that is very popular for mobile devices. The most popular example can be found in the Temple Run series. The goal of these games is to run as far as possible, without falling off the set path or crashing into obstacles, often with the side goal of collecting certain items to score points. The reason we chose this game type to base our experiment on is because this type of game works well on mobile devices. Also, this type of game is easily implemented in such a way that it suits the requirements for the three different variations we needed.

The game's scene consists of an endless road, which is traversed by a blue cube, which represents the player. The road is populated with red and yellow bars, which can be ducked under or dodged over, respectively. The road splits into intersections. One of the intersections is a dead end which is marked by a large blue wall, whereas the other intersections continues on to the next intersection.

The player's goal is to run as far as possible without falling of the set path, running into dead ends or crashing into objects. Objects can be avoided by swiping up or down to either jump over or duck under them, while falling of the road and crashing into dead ends can be avoided by picking a direction to turn before the next intersection by swiping left or right. If the player fails to avoid an object or dead end, or falls of the road, the game is over and the player can try again at the start of the level.

The experiment consists of three different variations of the game, where you cannot, can or must change perspective to determine which of the directions on each intersections is a dead end. The "can't" variation uses traditional visualization, whereas the "can" and "must" variations use Shoebox VR. With the "cannot" and "can" variations the blue walls that represent the dead ends are visible right after turning at an intersection in the default perspective, whereas in the "must" variation the blue walls are never visible using the default perspective, thereby forcing the player to change camera perspective in order to get a good score. Each variation consists of three different levels representing three different speeds, in order to determine the effects of and the relation between speed and variation on Shoebox VR.

3.2.2 Procedure

Participants played all three variations on three different devices, a 4" smart-phone and 7" and 10" tablets. The order in which each subject played on each device, was counterbalanced, such that each possible order was played an equal amount of times. For each variation the subject played the levels in the order from slowest to fastest movement speed. The first level was played each time as part of the tutorial. The player could make attempt for each level as long as the timer allowed. For the first level the timer was set for one minute, whereas for the second and third level the timer was set for 2 minutes. Afterwards, the subjects were interviewed about their experience with Shoebox VR.

3.2.3 Subjects

In the experiment 24 subjects participated, with an average age of 24 (SD±2.4), and of which 4 were female. The majority of the subjects (20) were Computer Science Master's students, that participated as part of a course. The remaining 4 were volunteers from friends and family. None of the subjects were paid for their services.

3.2.4 Results

Interaction performance

In order to determine the effect of Shoebox VR on games that rely on touch screen interaction, the following hypothesis was tested:

HB1: Shoebox VR negatively affects performance involving touch screen interaction.

Performance was measured by the average score (running distance) for each subject grouped by each variation. The "can't" and "can" variations are compared to determine the difference between touch performance for standard visualization and Shoebox VR (fig. 15). The "must" variation is not used for this comparison, as the design for this variation is not the same, due to the slightly different placement of the dead-ends, which could have an effect on performance.

To further investigate the effect of motion speed on performance using the different visualization techniques, the following hypothesis was stated:

HB4: The negative effect of Shoebox VR on performance and subjective experience becomes larger as speed increases.

Subjects played the game at three different speeds. The slowest speed setting was excluded in the comparison, since this was used for the tutorial level. Comparing standard visualization and Shoebox VR for the different speed settings, showed that the average gameplay performance decreases as the speed increases. In addition, a trend is seen that the difference in gameplay performance between the visualization techniques gets smaller at faster speeds.

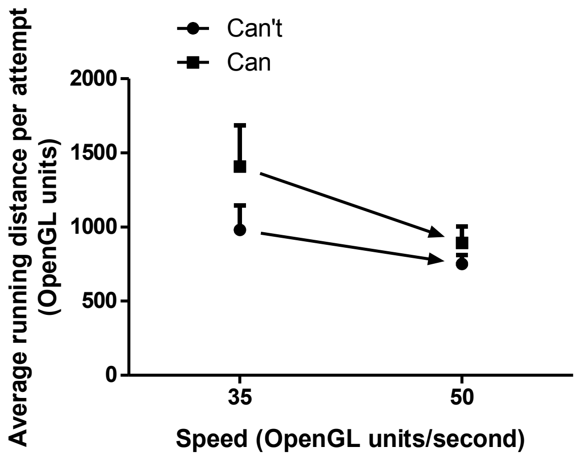

Fig. 15: The average running distance per attempt for "can't" and "can" variations, for different speed settings. The error bars represent the standard error of the mean. The arrows illustrate the downwards trend.

A two-way repeated measures ANOVA test, showed that the effect of visualization type on gameplay performance is significant (p=0.0226). Importantly, but the speed (p=0.0859) as well as the interaction between speed and visualization type (p=0.2489) do not significantly affect gameplay performance, indicating that the implementation of Shoebox VR can affect gameplay performance. In addition, a trend can be observed that an increase in speed negatively affects gameplay performance as well as a trend that shows that the difference in gameplay performance for the two visualization techniques gets smaller as the speed increases, but neither of these differences showed statistical significance. Based on these findings we reject **HB1** and **HB4**.

Entertainment value

To determine the effect of Shoebox VR on subjective experience with touch-based games, the following hypothesis was stated:

HB2: Shoebox VR negatively affects entertainment value when used on a touch screen-based application.

To determine the effect of the implementation of Shoebox VR on the entertainment value of touch-based games, the average grades, were compared between the different variations (fig. 16). Grades were based on a 7-point scale, where -3 represented very boring and +3 represented very entertaining,

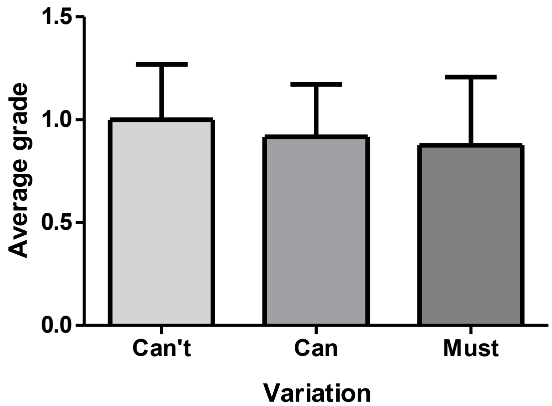

Fig. 16: The grade for average entertainment value for each variation. The error bars represent the standard error of the mean.

The effect of the implementation of Shoebox VR in a touch-based application is based on a comparison between the grades for the “can't” and “can” variations, representing standard visualization and Shoebox VR, respectively. A paired t-test was performed on these variations, showing no significant difference (p=0.7920), suggesting that there is no effect of Shoebox VR on the entertainment value. In addition, the effect of the role of Shoebox VR as either an essential or non-essential gameplay element is determined based on the comparison for grades between the “can” and “must” variations, representing Shoebox VR as a non-essential gameplay element and an essential gameplay element, respectively. A paired t-test revealed no significant difference (p=0.8852), which suggests that Shoebox VR does not effect entertainment value differently as a non-essential gameplay element than as an essential element. Based on these results, we reject **HB2**.

Controls intuitiveness

Subjects graded the intuitiveness of the controls for each variation. To determine the effect of Shoebox VR on touch-based games, the following hypothesis is tested:

HB3: Shoebox VR negatively affects control intuitiveness when used on a touch screen-based application.

Comparing the average grades given for intuitiveness for each variation (fig. 17), demonstrates the effect of Shoebox VR and the role as a gameplay element of Shoebox VR on control intuitiveness.

Grades were based on a 7-point scale, where -3 represented very unintuitive and +3 represented very intuitive,

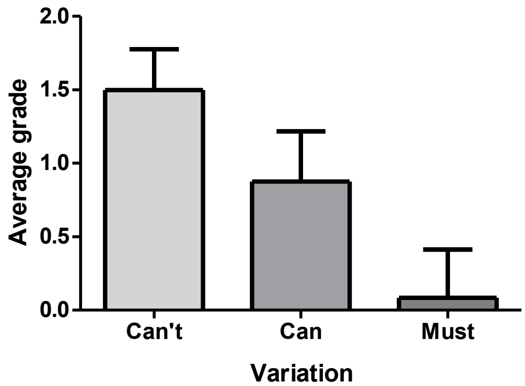

Fig. 17: The average grades for each variation, Error bars represent the standard error of th e mean.

By means of t-tests, we observe that there is a significant difference (p=0.0437) between the “can” and “must” variations. We also see a difference, between the “can't” and “can” variations, where the average grade is lower for “can”, but no statistical significance for this difference was detected (p=0.0698). This suggests that the controls are considered to be less intuitive if Shoebox VR is implemented as a required gameplay element, but not as an optional gameplay element. Therefore, **HB3** holds true.

Perspective change usage

The change in tilt angle was compared between different variations in order to determine if the way Shoebox VR is implemented, affects how it is used, thereby verifying the following hypothesis:

HB5: The average change in tilt angle is higher when Shoebox VR is implemented as an essential gameplay element than as a non-essential gameplay element.

Comparing the three variations (fig. 18), demonstrated that how actively the perspective is changed by the user is based on whether this is required by the game's design or not. The “can't” variation represents a base line, since change in tilt does not change perspective in this case, “can” represents the implementation with optional change in perspective, whereas with the “must” variation the change of perspective is required to perform well.

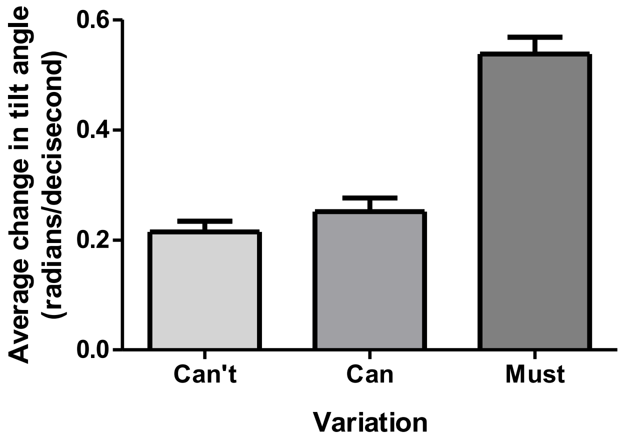

Fig. 18: The average change in tilt angle for each variation. Error bars depict the average error of the mean.

By means of t-tests, a significant difference is observed ($p<0.0001$) between the "can" and "must" variations, while there is no significant difference ($p=0.1236$) between the "can't" and "can" variations. This suggests that Shoebox VR must be implemented as an essential part of the gameplay to stimulate users to actively change perspective. Therefore, **HB5** is accepted.

3.2.5 Conclusions & Discussion

In order to gain more insight in the effect of touch screen interaction on motion-based games with touch screen interaction using Shoebox VR, the following hypotheses were stated:

HB1: Shoebox VR negatively affects performance involving touch screen interaction.

HB2: Shoebox VR negatively affects entertainment value when used on a touch screen-based application.

HB3: Shoebox VR negatively affects control intuitiveness when used on a touch screen-based application.

HB4: The negative effect of Shoebox VR on performance and subjective experience becomes larger as speed increases.

HB5: The average change in tilt angle is higher when Shoebox VR is implemented as an essential gameplay element than as a non-essential gameplay element.

Based on the analysis, **HB1** is rejected. We observe no negative effect on performance due to Shoebox VR. A negative effect was expected due to the complicated controls. Instead, the opposite is observed, as the performance was significantly better with Shoebox VR. This is possibly due to the ability to check again for the dead-ends. The results suggest that this added ability outweighs the negative effect of complex controls, in terms of performance.

HB3 hold true, as there is a significant decrease in control intuitiveness with Shoebox VR. Subjects often commented that the touch controls were confusing in combination with the tilting controls of Shoebox VR, this is likely the cause of the decrease in the rating for control intuitiveness.

HB5 also holds true, showing that the perspective changing capabilities introduced by Shoebox VR are only actively used if this is required to perform well in the game. Shoebox VR as an optional feature, is largely ignored.

HB2 and **HB4** were rejected, suggesting that Shoebox VR does not affect the entertainment value of motion-based games with touch screen controls in a negative way and that the speed is no factor of influence on either the objective or subjective experience with Shoebox VR. This indicates that Shoebox VR can be implemented in games without affecting entertainment value or experience.

Unfortunately, we were unable to analyse the immersion for motion-based games with touch screen controls, as these were not properly stored due to a bug in the application. It would have been particularly interesting to study immersion, as users commented that the user accelerometer noise filter resulted in an unstable image, and to determine whether this "jitter" effect would break immersion.

3.3 Experiment C: Tilt Interaction with Shoebox VR

To determine the effect of Shoebox VR on motion-based games that use tilt interaction, the following question was stated:

How does tilt interaction affect task performance and subjective experience with Shoebox VR in a motion-based game?

To investigate the influence of tilt interaction on performance, we stated the following hypotheses:

HC1: Shoebox VR negatively affects tilt interaction performance.

HC2: Placement pattern **does not** affect tilt interaction performance differently for standard visualization than for Shoebox VR.

HC3: Object population density **does not** affect tilt interaction performance differently for standard visualization than for Shoebox VR.

To investigate the influence of tilt interaction on subjective experience, we stated these hypotheses:

HC4: Shoebox VR has a negative effect on entertainment value in motion-based games with tilt controls.

HC5: Shoebox VR affects entertainment value differently depending on the object placement pattern.

HC6: Shoebox VR affects entertainment value differently depending on the object population density.

HC7: Shoebox VR affects visual experience differently depending on the object placement pattern.

HC8: Shoebox VR affects visual experience differently depending on the object population density.

For **HC1** a negative influence is expected, due to the fact that the implementation of Shoebox VR introduces tilting controls and adding additional tilt controls might become confusing. As for

HC2, **HC3**, no general performance benefit has been observed before and we do not expect that to change due to the addition of tilting controls. In contrast, it has been demonstrated that Shoebox VR can positively effect subjective experience and preliminary observations suggest that the type and number of reference points can have an influence on this effect, which is reflected in **H5**, **H6**, **H7** and **H8**.

3.3.1 Design

To verify the stated hypotheses concerning the effect of tilting interaction when using Shoebox VR, we designed a third person dodging game. In this game, a spaceship model moves forward through a tunnel at a constant pace. The player needs to navigate the camera to avoid mines (red orbs), which temporarily slow down the ship, and collect power-ups, which temporarily increase the ship's speed. The player's goal is to reach the end of a level as quick as possible.

The game consists of three different levels, each adhering to a different object placement pattern, annotated as "block", "circle" and "random", consisting of block shaped object groups, circle shaped object groups and randomly placed objects, respectively. An example for the block shaped group would be a cube volume filled with evenly spaced objects. Whereas the circle shaped groups could for example consist of objects placed in a spiral or cylinder. All levels are randomly generated based on predetermined parts, consisting of multiple orbs, except for the "random" levels, where the position for each orb is random generated. Levels are generated once and stored as part of the application, so that all players play the exact same levels, allowing for the direct comparison of data collected for each level.

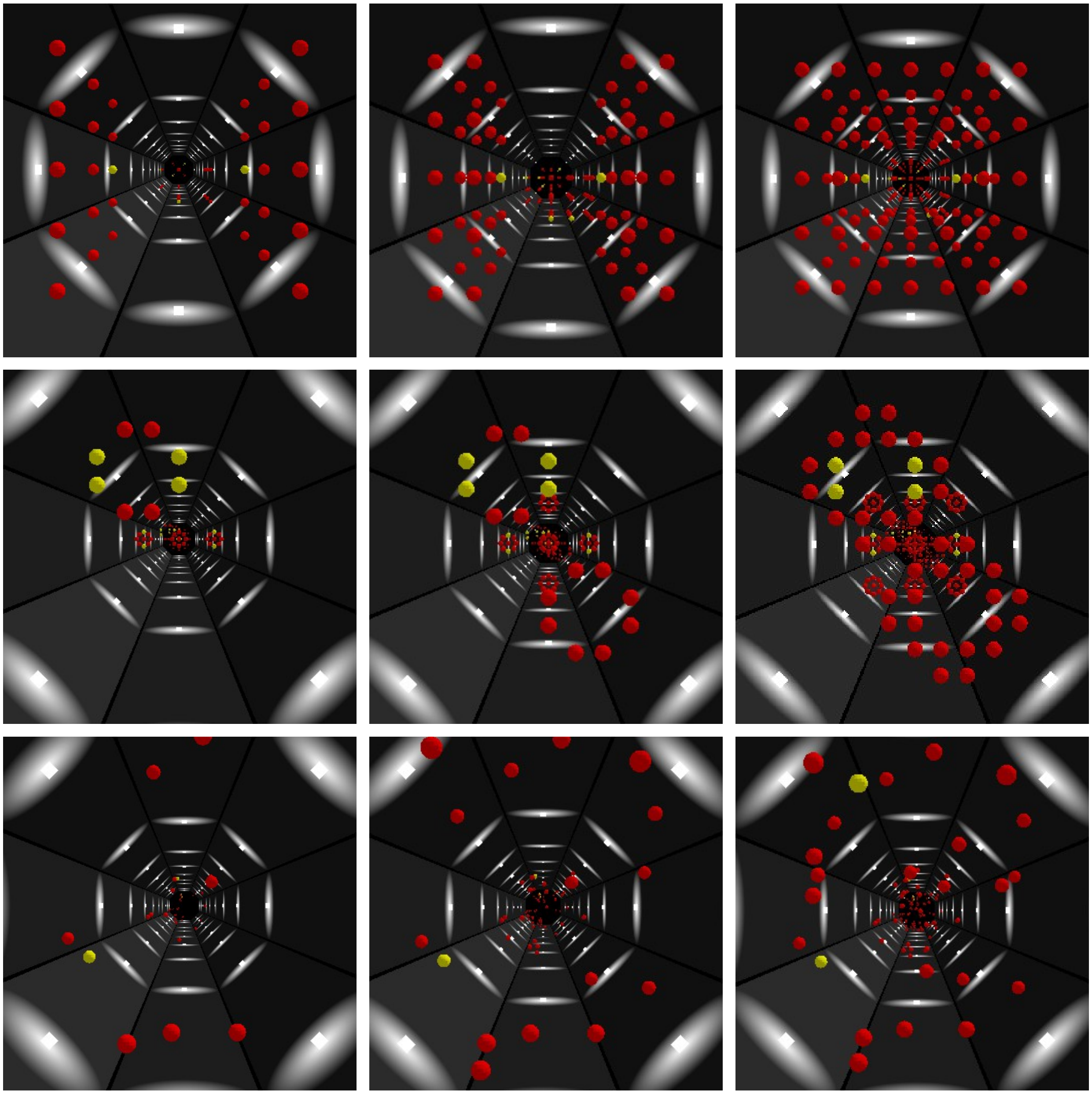

Fig. 19: Screen shots for each combination of placement pattern and object population density. Each row shows a different placement pattern and each column shows a different density.

There are three levels for each variation, each representing a different difficulty level, through different object population densities. The different levels are illustrated in figure 19. A level can be played in standard visualization mode or Shoebox VR mode. Once players completed a level for the first time, they were asked to fill in a questionnaire about entertainment value and visual experience.

The application collects data which is stored on a remote server. The first time the application connects to the server, it sends the device model on which the application is installed. The server sends a reply containing the device's assigned device ID. The device ID gets stored on the device and is sent along with each request to the server, so the data can be grouped for each device. The device ID gives an indication about which game was played by which subject, but obviously we have no guaranty about the uniqueness of users as multiple users can play the game on the same device.

If the device cannot connect to the server at any time, for example due to the lack of an active internet connection, the data is temporarily stored locally. This data is send to the server on the first next successful attempt to connect to the server.

The application will gather the following data and send it to the server:

- The ID of the played case.
- The completion time for each played game.
- Questionnaire results.

The gathered data shows the gameplay performance for each visualization type and will give insight in subjective experience in terms of entertainment value and visual experience with respect to placement patterns and object population density. This allowed us to determine which are the preferred values for these variables. The questionnaire results give insight in whether these preferences are related to either entertainment value or visual appeal.

3.3.2 Procedure

At first start-up, users will be asked permission to send anonymous data to our server. After accepting the terms, the game is explained by a series of explanatory screen shots, after which the main menu is shown. In the main users can pick a level, each representing a different combination of placement pattern, object population density and visualization type. By completing a level within the time limit (1 minute), the player unlocks levels with a higher object population density.

The order in which placement patterns and view modes are shown is randomized, to minimize order effects. Obviously, this will not result in a perfect balance, but due to the nature of this remote study, we cannot enforce subjects to play each level or play levels in a specific order.

3.3.3 Subjects

Because the game was distributed using the Google Play store, the number of gameplay disruptions were kept to a minimum, in order to prevent players getting frustrated. For this reason, we did not obtain knowledge of the previous experience and age of the participants. However, we can assume that these users have at least some experience with mobile devices. A number of 36 subjects installed our application, and played a total of 249 games.

3.3.4 Results

The data for this experiment was gathered remotely and contained gaps as not all of subjects played all levels. However, we were able to calculate the average completion times for each level, based on the averages for each player that played that level. We did not perform any form of data filtering as no anomalies, such as outliers, were observed.

Interaction performance

To determine if there is a difference in gameplay performance for a game with tilt-based controls between standard visualization and Shoebox VR, we stated the following hypothesis:

HC1: Shoebox VR negatively affects tilt interaction performance.

Comparing the average level completion time between standard visualization and Shoebox VR, showed a difference in objective gameplay performance between the visualization types (fig. 20).

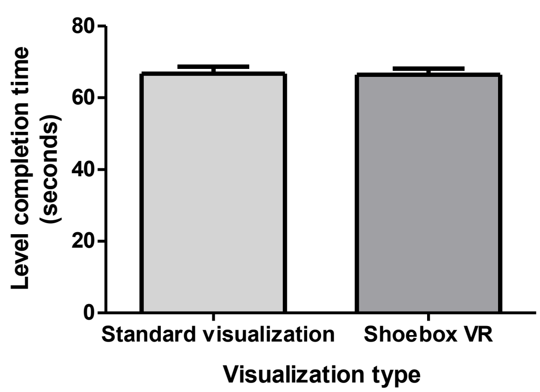

Fig. 20: The average level completion time for both standard visualization and Shoebox VR. The error bars represent the standard error of the mean.

A t-test comparing the average completion times between standard visualization and Shoebox VR, showed no significant difference (p=0.9091), suggesting that Shoebox VR does not effect interaction performance for motion-based games with tilt controls. Therefore, **HC1** is rejected.

To discover the effect of different placement patterns on gameplay performance, we stated the following hypothesis:

HC2: Placement pattern **does not** affect tilt interaction performance differently for standard visualization than for Shoebox VR.

Comparing the difference in average level completion time between standard visualization and Shoebox VR for each placement pattern, demonstrated how placement pattern affects Shoebox VR differently than it affects standard visualization (fig. 21).

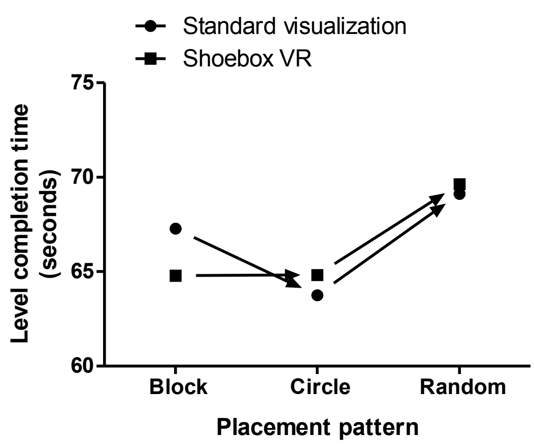

Fig. 21: The average level completion time for standard visualization and Shoebox VR per placement pattern.

A two-way ANOVA test reveals that the effect of placement pattern (p=0.3287) as well as the effect of visualization type (p=0.9131) are not statistically significant. Furthermore, any interaction between placement pattern and visualization type is also not significant (p=0.8485). This analysis showed that placement pattern does not affect Shoebox VR differently than standard visualization in terms of gameplay performance. In fact, Shoebox VR showed no significant effect of placement pattern on gameplay performance at all. Based on this, we conclude that **HC2** holds true.

Finally, in order to determine the effect of different population densities when using Shoebox VR in tilt controlled motion-based games, we stated the following hypothesis:

HC3: Object population density **does not** affect tilt interaction performance differently for standard visualization than for Shoebox VR.

Comparison of the average level completion times for different densities between standard visualization and Shoebox VR shows whether the object population density affects gameplay performance using Shoebox VR (fig. 22).

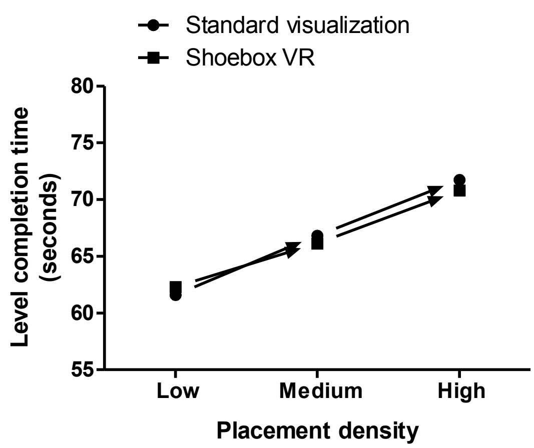

Fig. 22: The average level completion time for standard visualization and Shoebox VR per object population density.

Two-way ANOVA analysis shows that the effect of object population density on level completion speed is significant (p=0.0092). However, the effect of visualization type is not (p=0.8842). Also, any interaction observed between population density and visualization type is not significant (p=0.9376). Thus, the analysis showed that increasing density affects gameplay performance, which is to be expected as this makes the game more difficult, but that this increase was not significantly different for Shoebox VR compared to standard visualization. Therefore, we conclude that **HC3** holds true.

Entertainment value

In order to show the effect of Shoebox VR on motion-based games with tilt controls, we stated the following hypothesis:

HC4: Shoebox VR has a negative effect on entertainment value in motion-based games with tilt controls.

By comparing the average entertainment rating for both standard visualization and Shoebox VR (fig. 23), the effect of Shoebox VR on games with tilt controls could be determined.

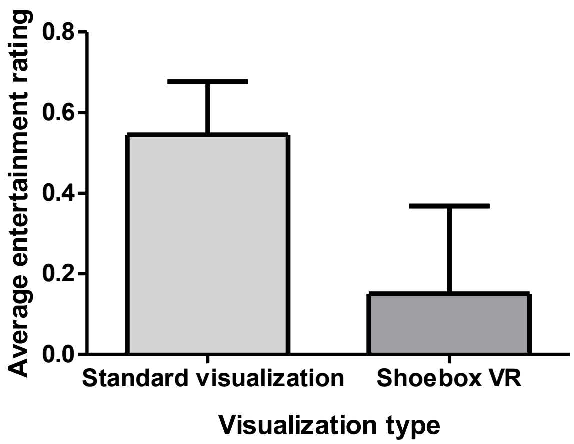

Fig. 23: The average entertainment rating for standard visualization and Shoebox VR. The error bars represent the standard error of the mean.

By performing a t-test on the average entertainment rating between standard visualization and Shoebox VR, we observe that the difference in rating between these visualization types is not significant ($p=0.1415$). Based on this result, we reject **HC4**.

To explore the effect of different placement patterns on the entertainment value when using Shoebox VR, we stated the following hypothesis:

HC5: Shoebox VR affects entertainment value differently depending on the object placement pattern.

The average rating for entertainment value is compared for each placement pattern between standard visualization and Shoebox VR (fig. 24), in order to determine any effect of object placement patterns on Shoebox VR.

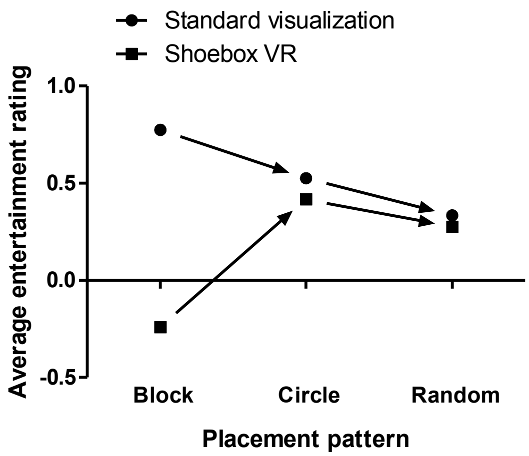

Fig. 24. The average entertainment rating for standard visualization and Shoebox VR for each placement pattern.

A two-way ANOVA test showed no significant effects for either visualization type (p=0.1562) or placement pattern (p=0.7961). Also, the any interaction effect between visualization type and placement pattern was statistically not significant (p=0.2792). This result suggests that the effect of Shoebox VR on motion-based games with tilt control does not depend on placement patterns and therefore we reject **HC5**.

In order to determine the effect of different object population densities on entertainment value when using Shoebox VR, the following hypothesis was stated:

HC6: Shoebox VR affects entertainment value differently depending on the object population density.

The differences in average entertainment rating for each value for population density between standard visualization and Shoebox VR were compared (fig. 25), to determine whether the effect of Shoebox VR on entertainment value is different depending on object population density.

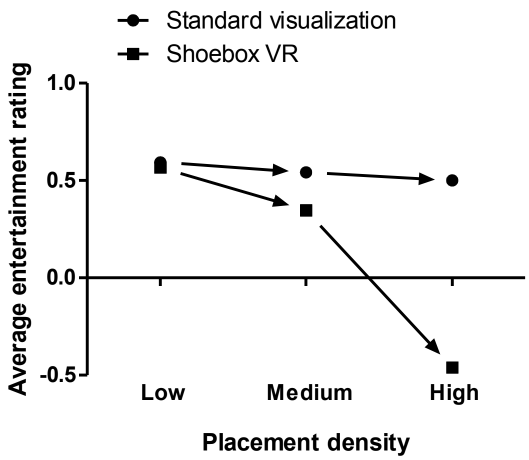

Fig. 25. The average entertainment rating for standard visualization and Shoebox VR for each object population density.

Statistical analysis by means of a two-way ANOVA showed that the effects of both visualization type ($p=0.1150$) and object population density ($p=0.1630$) were not significant. Furthermore, no significant effects of interaction between visualization and population density were found ($p=0.2541$). Based on these results, **HC5** is rejected, as they suggest that there is no difference in average entertainment rating using Shoebox VR between different density populations.

Visual experience

To determine the influence of Shoebox VR on visual experience of motion-based games with tilt controls, the following hypothesis was stated:

HC6: Shoebox VR has a negative effect on visual experience in motion-based games with tilt controls.

To test the hypothesis, ratings for visual experience are compared between standard visualization and Shoebox VR (fig. 26). Ratings are given on a 7-point scale, ranging from -3 and +3, where -3 represents very unappealing and +3 represents very appealing.

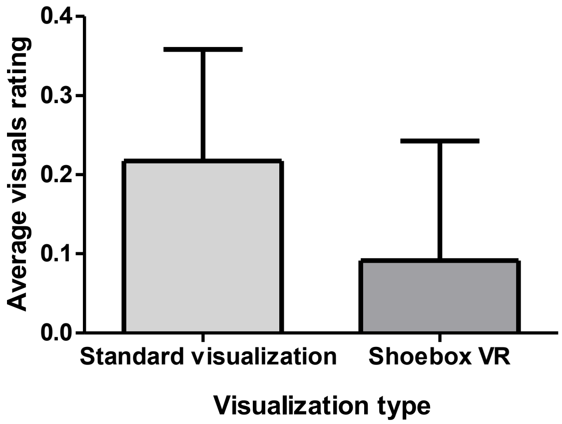

Fig. 26: The average visual experience rating for standard visualization and Shoebox VR. Error bars represent the standard error of the mean.

Performing a t-test revealed no statistically significant difference between the average visual experience rating for standard visualization and for Shoebox VR, suggesting that Shoebox VR does not have a significant effect on visual experience with motion-based games using tilt controls. Therefore, we reject **HC6**.

The following hypothesis was stated, to determine the effect of Shoebox VR on visual experience using different placement patterns:

HC7: Shoebox VR affects visual experience differently depending on the object placement pattern.

The average ratings for visual experience for standard visualization and different placement patterns are compared to determine an effect of Shoebox VR with respect to placement patterns (fig. 27).

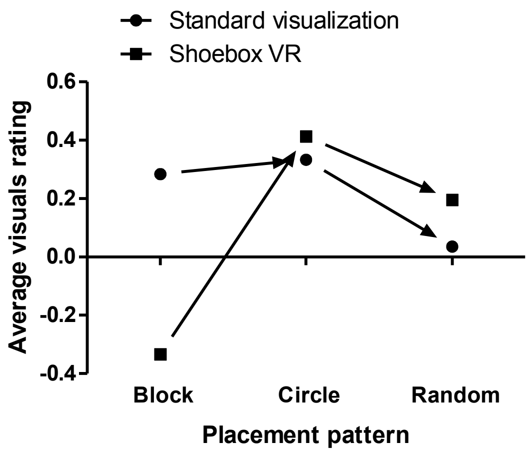

Fig. 27: The average visual experience for standard visualization and Shoebox VR for each placement pattern.

A two-way ANOVA analysis did not show statistically significant effects for both visualization type (p=0.5281) and placement pattern (p=0.2751) on visual experience rating. In addition, the effect of interaction between visualization type and placement pattern is not significant (p=0.2394). These result suggest that Shoebox VR affects visual experience the same for difference placement patterns. Based on this **HC7** is rejected.

In order to determine the effect of Shoebox VR on visual experience using different population densities, we stated the following hypothesis:

HC8: Shoebox VR affects visual experience differently depending on the object population density.

For each population density and visualization type pair the average visual experience rating was compared (fig. 28), to test if Shoebox VR affects visual experience differently for different object population densities.

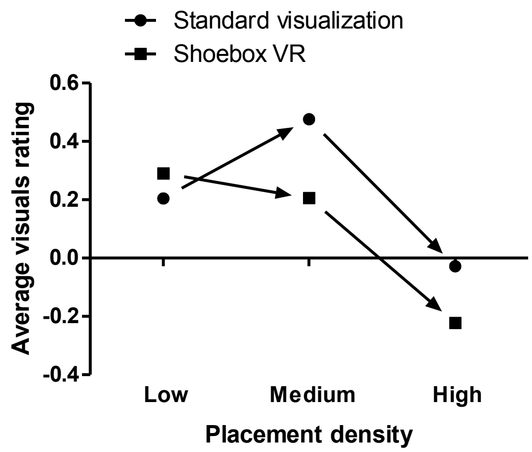

Fig. 28: The average visual experience for standard visualization and Shoebox VR for each population density.

A two-way ANOVA test shows no significant effects for visualization type (p=0.5469) or population density (p=0.1835), nor is the effect of interaction between visualization type and population density significant (p=0.7584). Based on these results, we reject **HC8**.

3.4.5 Conclusions & Discussion

In order to gain more insight in the effect of touch interaction motion-based games using Shoebox VR, we stated the following hypotheses:

HC1: Shoebox VR negatively affects tilt interaction performance.

HC2: Placement pattern **does not affect** tilt interaction performance differently for standard visualization than for Shoebox VR.

HC3: Object population density **does not affect** tilt interaction performance differently for standard visualization than for Shoebox VR.

HC4: Shoebox VR has a negative effect on entertainment value in motion-based games with tilt controls.

HC5: Shoebox VR affects entertainment value differently depending on the object placement pattern.

HC6: Shoebox VR affects entertainment value differently depending on the object population density.

HC7: Shoebox VR affects visual experience differently depending on the object placement pattern.

HC8: Shoebox VR affects visual experience differently depending on the object population density.

Based on the analysis, **HC2** and **HC3** hold, while the other hypotheses are rejected. This suggests that tilt interaction does not necessarily have a negative effect on the gameplay performance and subjective experience with motion-based games using Shoebox VR. Furthermore, the placement pattern and object population density seem to have no effect on gameplay performance or subjective experience using Shoebox VR. While this means that visualization with Shoebox VR is possible without significantly disrupting gameplay performance or subjective experience, one could argue that it is not useful to implement Shoebox VR if it does not add to the experience. Therefore, it would be useful to further investigate scenarios where Shoebox VR could add certain benefits.

3.4 Experiment D: Filter evaluation

Since Shoebox VR makes use of the accelerometer, which provides data with high noise levels, it was necessary to filter this data. To do so, we used the Kalman filter, as this is an established method for the filtering of accelerometer data. However, participants of experiment B commonly commented that the image was "shaky" or "unstable", despite the use of Kalman filtering. As our goal is to implement Shoebox VR as robust as possible in order to get representative results, this experiment compares three different methods for filtering noise from the accelerometer data to see whether the filter type affects the results. As a baseline to compare the noise filters to, the use of no filter was included in this experiment.

In order to determine the user experience with each filter, we conducted a small experiment, where subjects were able to compare these filters. Also, the filter preference for a still scene or a moving scene was compared, as it seemed plausible that these might differ. For example, a moving scene might require the most responsive filter possible, whereas for a still scene the most stable image might be preferred.

3.4.1 Design and procedure

The subjects were presented with a simple scene, showing a large collection of spheres. The subjects were able to switch between the filters at will as they were viewing the scene by tapping on screen buttons. The subjects were asked to list their preferences for filters, by depicting which filter was the most preferred and which filter was the least preferred, depending on two criteria, responsiveness and image stability. This was done both for the still scene, where the spheres did not move, and the moving scene, where the spheres were moving towards the camera.

3.4.2 Subjects

A small group of 6 people participated. We considered using a small number of subjects justifiable in this case, as the experiment is used to backup objective facts about the noise filters with subjective experience. Of the 6 subjects, 5 were Computer Scene Master's students and male and 1 subject was a female volunteer.

3.4.3 Results

In order to determine which of the test filters results in the best responsiveness and which filter results in the most stable image and whether the addition of motion has an effect on this (fig. 29, 30), two-way repeated measures ANOVA tests were performed, with filter type and movement as independent variables and ranking for responsiveness or image stability as dependent value. For responsiveness, the interaction between movement and filter type (p=0.2371) and the effect of movement individually (p=1.000) were not significant. However, the effect of filter type on responsiveness was significant (p=0.0152), showing that the type of filter affects the responsiveness perceived by the user, as expected, whether the scene is moving or not not.

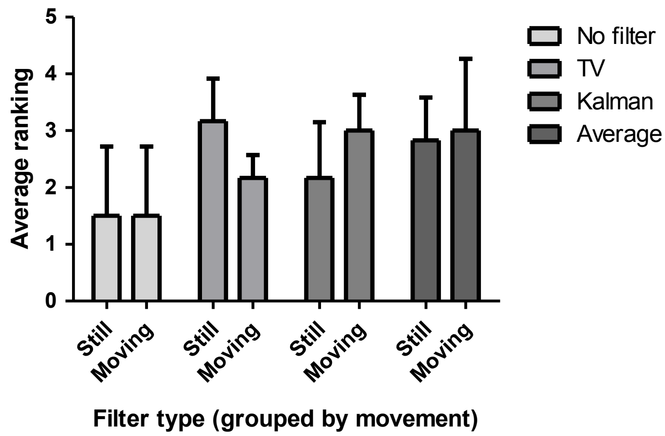

Fig. 29: Average responsiveness ranking for each filter type grouped movement. (TV = Total Variation filter). A rank was assigned to each filter (or lack there of) for both cases, where 1 represents the best responsiveness and 4 represents the worst responsiveness.

For image stability, the interaction between movement and filter type (p=0,1806) and the effect of movement individually on stability (p=0,7342) were not significant. However, the filter type showed a significant effect on image stability ranking (p<0.0001), showing that the type of filter affects the image stability perceived by the user, independent of whether the scene moves or not. These results meet our expectations.

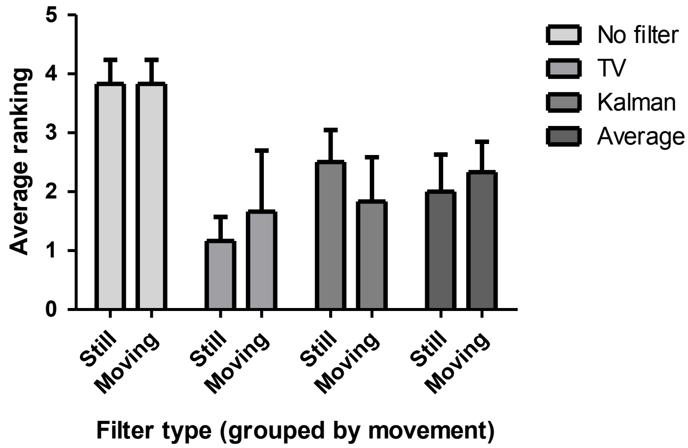

Fig. 30: Average stability ranking for each filter type grouped movement. A rank was assigned to each filter (or lack there of) for both cases, where 1 represents the best stability and 4 represents the worst stability.

In order to determine the significance in difference between each filter type, an average ranking was calculated for each subject based on the rankings for the moving scene and the non-moving scene (fig. 31, 32). On these rankings paired t-tests were performed between each pair of filter types, for both responsiveness and image stability.

For responsiveness, the differences between TV versus Kalman (p=0.7711), TV versus Average (p=0.3939) and Kalman versus Average (p=0.1852) were not significant, showing that there is no significant difference in responsiveness between any of the tested filters.

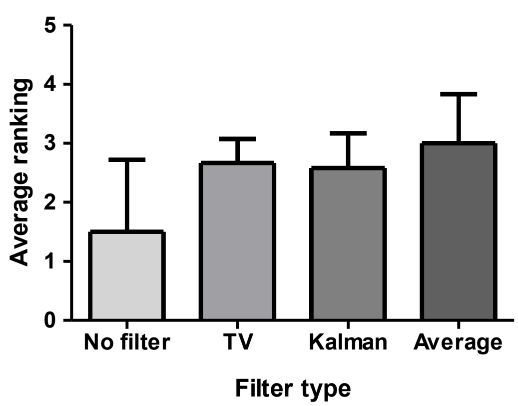

Fig. 31: Average responsiveness ranking for each filter type based on average of still and moving scene rankings, where 1 represents the best responsiveness and 4 represents the worst responsiveness.

For image stability the differences between TV versus Kalman (p=0.0301) and TV versus Average (p=0.0446) were significant, indicating that the TV filter has an advantage over Kalman and Average filter in terms of image stability. As expected, all filters provided better image stability than when no filter was used. The difference between Kalman and Average filters (p=1.0000) is not significant.

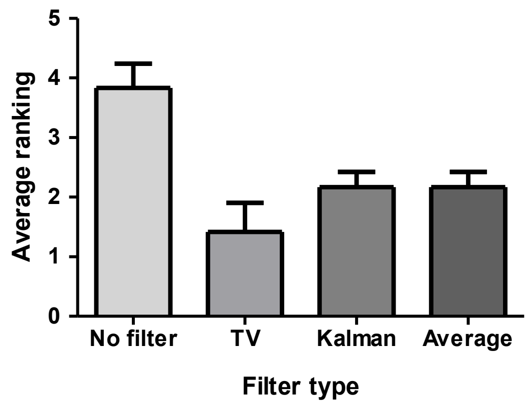

Figure 32: Average stability rating for each filter type grouped based on average of still and moving scene rankings, where 1 represents the best stability and 4 represents the worst stability.

3.4.4 Conclusion

In order to determine the effect of different noise filters on subjective experience of Shoebox VR, subjects were asked to rank three different filters based on responsiveness and image stability, for both a still and a moving scene. By comparing the questionnaire results we determined that users consider the Total Variation filter to produce the most stable image. The three filters did not show a significant difference in responsiveness is not significant. Therefore, we conclude that the Total Variation filter is the most appropriate technique to filter accelerometer data for Shoebox VR visualization.

4. Conclusions & Discussion

In the introduction we stated the following questions:

- How does screen size affect depth perception and subjective experience with Shoebox VR and how does it affect usability of Shoebox VR in a motion-based game?
- How does touch screen interaction affect task performance and subjective experience with Shoebox VR in a motion-based game?
- How does tilt interaction affect task performance and subjective experience with Shoebox VR in a motion-based game?

We performed experiments A, B and C in order to determine answers for these questions. Based on our results, we concluded that screen size has no effect on subjective experience with motion-based games using Shoebox VR. However, we did see a significant difference in depth perception performance with slower values for speed on a small screen, showing that Shoebox VR can provide a benefit in depth perception performance under these conditions. This means that Shoebox VR can be used effectively on devices on different screen sizes, without reducing subjective benefit that it can provide. Moreover, this proves that situations exist where Shoebox VR can be beneficial in terms of objective depth perception performance. By further exploring and clearly defining the relation between depth perception performance, screen size and motion speed, certain games could improve their usability by applying Shoebox VR.

The knowledge gained about the effect of Shoebox VR on touch screen interaction shows that Shoebox VR can have a positive effect on gameplay performance, but we did not observe a significant difference in entertainment value. However, the control intuitiveness was significantly lower with Shoebox VR. It is reasonable to assume that this is due to the complicated touch control schemes, as the there have not been comments about this issue with previous experiments, which were based on simple touch screen input. It might be useful to determine what level of complexity the touch screen controls can have before Shoebox VR starts to have a negative effect on control intuitiveness.

Furthermore, we observed that Shoebox VR is only actively used when used as an essential part of the gameplay. This shows us that it is particularly useful to implement Shoebox VR in fixed motion-based games, if its usage is enforced by the gameplay.

Concerning tilting interaction, the results suggest that Shoebox VR can be implemented in these type of games without affecting gameplay performance and subject experience. Obviously, this should be done with care as Shoebox VR itself is already controlled by tilt controls, combining this with additional tilt controls, could result in a counter-intuitive control scheme. However, since it is yet unknown what kind of tilt controls would result in such a scheme further research is required.

Concluding, this report provided new insights into the usability of Shoebox VR in motion-based games. Most importantly, we showed that care should be taken when implementing touch screen controls for motion-based games using Shoebox VR and that Shoebox VR is most effective when used as an essential part of the gameplay experience. These insights will allow for more successful development of mobile VR games, that will provide new immersive and entertaining experiences.

5. Acknowledgements

First of all, I'd like to thank Dr. Wolfgang Hürst for giving me the opportunity to perform this study under his supervision and for his guidance during the project. I'd also like to acknowledge Marco van Laar, for our collaboration on the mobile VR library. Furthermore, I'd like to thank all the participants of the experiments for their cooperation and willingness to give me useful feedback.

6. References

Brown, E., Cairns, P. (2004). A grounded investigation of game immersion. C*HI'04 Extended abstracts on human factors in computing systems. ACM*: 1297-1300.

Condat, L. (2013). A direct algorithm for 1D total variation denoising. *IEEE Signal Processing Letters*; 20: 1054-1057.

Cox, A.L., Cairns, P., Berthouze, N., Jennett, C. (2006). The use of eyetracking for measuring immersion. *Workshop on What have eye movements told us so far, and what is next? at Proceedings of CogSci2006, the Twenty-Eighth Annual Meeting of the Cognitive Science Society*; July 26-29.

Cruz-Neira, C., Sandin, D.J., and DeFanti, T.A. (1993). Surround-screen projection-based virtual reality: the design and implementation of the CAVE. *Proceedings of the 20th annual conference on Computer graphics and interactive techniques. ACM*: 135-142.

Friedmann, M., Starner, T., Pentland, A. (1992). Device synchronization using an optimal linear filter. *Proceedings of the 1992 symposium on Interactive 3D graphics*. *ACM*; 57-62.

Jennett, C., Cox, A.L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A. (2008). Measuring and defining the experience of immersion in games. *International journal of human-computer studies*: 66; 641-661.

Henze, N. (2012). Hit it!: an apparatus for upscaling mobile HCI studies. *CHI'12 Extended abstracts on human factors in computing systems. ACM*: 1333-1338.

Henze, N., Pielot, M., Poppinga, B. (2011). My app is an experiment: experience from user studies in mobile app stores. *International journal of mobile human computer interaction*; 3: 71-91.

Hou, J., Yujung, N., Peng, W., Lee, K.M. (2012). Effects of screen size, viewing angle, and players' immersion tendencies on game experience. *Computers in Human Behavior*; 28: 617-623.

Hürst, W., Beurskens, J. (2013). Shoebox VR: the influence of motion and environment. *Experimentation Project, Utrecht University.*

Hürst, W., Helder, M. (2011). Mobile 3D graphics and virtual reality interaction. *Proceedings of the 8th international conference on advances in computer entertainment technology. ACM.*

Hwang, J., Jung, J., Kim, G.J. (2006). Hand-held virtual reality: a feasibility study. *Proceedings of the ACM symposium on virtual reality software and technology*. *ACM*: 356-363.

Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Journal of fluids engineering*; 82*:* 35-45.

Lee, J. (2007). Head tracking for desktop VR displays using the Wii remote. Available at http://johnnylee. net/projects/wii.

Liang, J., Shaw, C., Green, M. (1991). On temporal-spatial realism in the virtual reality environment. *Proceedings of the 4th annual ACM symposium on User interface software and technology*. *ACM*: 1-7.

Lombard, M., Ditton, T.B., Crane, D., Davis, B., Gil-Egui, G., Horvath, K., Rossman, J. (2000). Measuring presence: a literature-based approach to the development of a standardized paper-and-pencil instrument. *Project abstract submitted for presentation at Presence 2000: the third international workshop on presence, Delft.*

Pausch, R., Proffitt, D., Williams, G. (1997). Quantifying immersion in virtual reality. *Proceedings of the 24th annual conference on computer graphics and interactive techniques*. *ACM.*

Prothero, J. D., Hoffman, H.G. (1995). Widening the field-of-view increases the sense of presence in immersive virtual environments. *Human interface technology laboratory technical report*; TR-95-2.

Siewert, J. (2012). Experimentation project report. *Experimentation Project, Utrecht University.*

Slater, M. (1999). Measuring presence: a response to the Witmer and Singer presence questionnaire. *Presence: teleoperators and virtual environments*; 8: 560-565.

Slater, M., Linakis, V., Usoh, M., Kooper, R., Street, G. (1996). Immersion, presence, and performance in virtual environments: an experiment with tri-dimensional chess. *ACM virtual reality software and technology*: 163-172.

Van Baren, J., Ijsselsteijn, W. (2004). Measuring presence: a guide to current measurement approaches. *Deliverable of the OmniPres project*: IST-2001-39237.

Van Laar, M. (2014). Opening up the box: a comprehensive evaluation of shoebox virtual reality and its function in open environments. *MA thesis, Utrecht University.*

Ware, C., Arthur, K., Booth, K.S. (1993). Fish tank virtual reality. *Proceedings of the INTERACT'93 and CHI'93 conference on human factors in computing systems*. *ACM:* 37-42.

Witmer, B.G., and Singer, M.J. (1998). Measuring presence in virtual environments: a presence questionnaire. *Presence: teleoperators and virtual environments*; 7: 225-240.