Utrecht University

Master Thesis

Influence of the Shoebox Model and its Parameters

on the 3d Perception and Experience of Users

Author:
Martin de Jong, BSc

Student Number:
3117340

Supervisor:
Dr. Wolfgang Hürst

Thesis Number:
ICA-3117340

August 29, 2012

Abstract

In recent years the adoption of mobile devices such as smartphones and tablets has increased dramatically. These devices allow for new 3d visualization concepts, such as the shoebox model. This concept uses perspective correction to the image on screen when the device tilted to enhance the 3d effect of the image. The field of view and the shearing factor are identified as the main parameters that determine the look of the shoebox effect. This report researches the effect of the shoebox model, on users 3d perception and experience.

To examine the effects, several user studies are performed. Each of these user studies focuses on a different task involving some form of depth perception (used as an indication of 3d perception). The participants in these studies are asked to fill out a questionnaire to gain insight in the subjective experience of users when using the concept.

The results of these studies suggest that the shoebox model has a significant positive influence on a user's depth perception. The influence of the different parameters on the depth perception is not significant. Application developers are therefore recommended to choose a value for these parameters that results in the most visually appealing effect or is most practical for its intended application.

The user reaction to the shoebox model is moderately positive. Users prefer it to the use of auto stereoscopic screens, which they noted offer a better 3d effect but are more cumbersome to use and cause fatigue of the eyes.

Contents

Chapter 1

Introduction

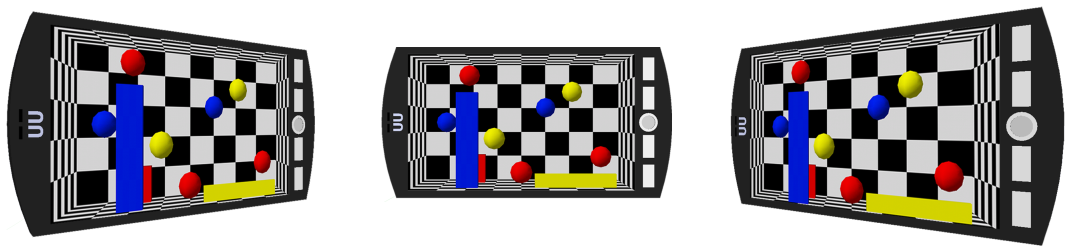
1.1 The Shoebox Model

Recent years saw a big rise in the adoption of mobile devices such as smartphones and tablets. The portable nature of mobile devices makes them distinctly different from traditional desktop environments and introduces new problems and opportunities. Mobile devices have several limitations compared to desktop environments, mainly because of their low power supply, low computational power, small physical display size and limited input modalities. However, the area of research in mobile graphics has been taking a big step in recent years due to technological advances [4]. These advances overcome a lot of previous limitations and allow for more detailed 3d graphics.

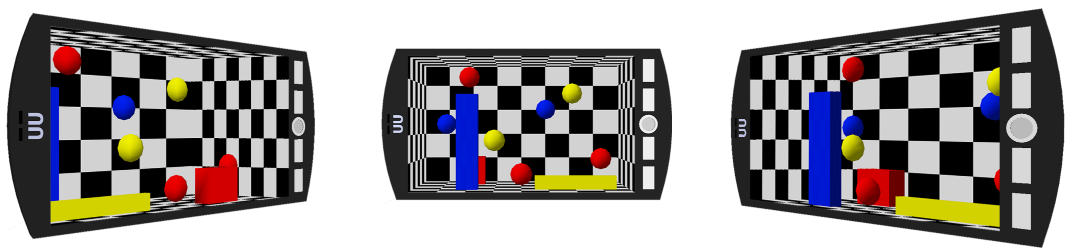
Contemporary mobile devices, in particular the smart-phone, are not just equipped with the traditional input and output methods. They come equipped with a wide variety of low cost sensors, such as cameras, compasses and inertial sensors such as accelerometers or gyroscopes. These sensors allow the mobile device to be spatially aware. This introduces new possibilities for different 3d concepts that were not possible on traditional stationary hardware.

One of these new 3d concepts is the *shoebox model* [13]. It is a novel way to display 3d graphics on mobile devices. It uses the orientation of the device, deduced from the output of the devices' inertial sensors, as input for its 3d visualisation. The shoebox model uses this information to emulate a 3d space attached behind the device. If a person holds a box in his hands and rotates the box, his perspective of the inside of the box changes. This does not happen when a user rotates a device with traditional 3d graphics (see figure 1.1(a)). So a perspective correction needs to be applied to the image to gain this same effect. This perspective correction must correspond to the orientation of the device. Now the user can look into the virtual space at different angles as the device is tilted (see figure 1.1(b)). The name of the effect comes from its comparison to a peep-hole shoebox. In this comparison the 'peep-hole' consists of the entire screen of the device.

It draws comparison to the Fish tank model [24], in which the perspective of an image on a static screen is corrected to account for the position of the head of the user. This is more complex than the fixed head position of the shoebox model, since some form of head tracking needs to be applied to account for the position of the user's head. This head tracking can also be applied to a mobile device, but there are some technological problems involved. Most importantly, the mobile device would need to be equipped with a front facing camera, which



(a) The effect of tilting of the device on the image in traditional computer graphics (without perspective correction).



(b) The effect of tilting of the device on the image in the shoebox model (with perspective correction).

Figure 1.1: A mobile device displaying 3d graphics without (a) and with (b)perspective correction.

a lot of mobile devices lack. It is expected though, in the future, almost all mobile devices will be equipped with a front facing camera. Still, even if front facing cameras would be widely adopted, problems remain. Tracking the position of one's head requires complex algorithms. Varying lighting conditions make it difficult to correctly calibrate such an algorithm. Another problem is that most front-facing cameras on mobile phones have a limited field of view meant for video communication. This means a user would quickly be outside of the range of the camera, and thus would not be tracked.

1.2 Motivation

The shoebox model could offer several advantages over the traditional way of displaying 3d graphics. Because users can look at scene from different angles more depth information of the scene can be acquired. This suggests a better 3d perception of users. The changing perspective of the shoebox model more closely resembles the real experience of looking at objects in three-dimensional space. It can therefore be assumed that the shoebox model will give users a greater sense of realism. This will most likely make the presented graphics more visually appealing to users.

If these intuitive assumptions are true, the shoebox model can offer potential benefits in certain application areas. For example, the extra depth information can be beneficial to *data visualisation* applications. Three-dimensional charts can be much more intuitive and easy to read if extra depth information is present in the visualization. Geographical data can also benefit from the addition of extra depth information, for example the images of mountain ranges or rock formations. In the medical area, it could help patients and doctors better interpret medical images, such as MRI scans and CT X-Ray images. The additional depth information and assumed visual appeal make this concept also fitted for a *3d model viewer* application. For example, precious objects in a museum could be on virtual display using the shoebox model. This way people not present at the museum can still view the object as if they are presented to them in a glass cabinet. Web-shops can also use it in a similar way to show objects to customers. This way, it is possible for customers to get a better impression of the object of their interest, before buying it. Furthermore, all of the assumed advantages make the shoebox model beneficial for use in *games*. Added depth information can be useful for certain tasks users need to perform in games, for example judging the distance at which enemies are present. A greater sense of realism could get users more immersed in the game. And since games are mostly applied for entertainment purposes, the visual appeal of the game is very important. Some games (such as "a 3d Tilt Labyrinth" [21]) already make use of the shoebox model. However, the number of games that make use of this concept is very low.

The intuitive assumptions made will need to be confirmed to know if use in the proposed application areas is feasible. Furthermore it should be studied how all involved parameters of the shoebox model affect these intuitive assumptions, to know how to maximize the benefits of the shoebox model.

1.3 Goals, Research questions and Objectives

The goal of this thesis is to confirm the possible potential benefit on 3d perception, realism and enjoyment of users resulting from use of the shoebox model and investigate the influence of the different parameters of the shoebox model on these potential benefits (see section 1.2).

To reach this goal the following research questions have to be answered:

1. Does the shoebox model achieve a better 3d perception of users compared to the conventional 3d visualisation?
2. How does 3d perception in the shoebox model depend on the involved parameters?
3. Is the shoebox model beneficial to a users sense of realism?
4. How does the realism in the shoebox model depend on the involved parameters?
5. Is the shoebox model beneficial to an application's visual appeal?

Each of these research questions is divided into several objectives.

1. Does the shoebox model achieve a better 3d perception of users compared to the conventional 3d visualisation?

To investigate this the depth perception of users is chosen as an indication of 3d perception, since it is the easiest to measure. This results in the following objective:

> **Obj1:** Compare the shoebox model and the conventional 3d visualisation in terms of depth perception.

2. How does 3d perception in the shoebox model depend on the involved parameters?

To answer this question, first the involving parameters must be known:

> **Obj2a:** Identify the different parameters that influence the look of the shoebox model.

After that the influence of these parameters on the depth perception can be measured:

> **Obj2b:** Investigate the influence of the different parameters on the depth perception of users.

The effectiveness of the shoebox model most likely also depends on the nature of the scene. The shift in perception is more noticeable if clear 'reference planes' are present (see section 5.2.2).

> **Obj2c:** Investigate if the presence of clear reference planes in the scene is important for the users depth perception.

3. Is the shoebox model beneficial to a users sense of realism?

To research the users' sense of realism using the effect, their opinion on its realism will be asked in a questionnaire.

> **Obj3:** Find out the opinion of user regarding the realism of the effect by letting them fill out a questionnaire.

4. How does the sense of realism of users in the shoebox model depend on the involved parameters?

To investigate this the sense of realism of users will need to be asked at different values for each parameter.

> **Obj4:** Investigate the influence of the different parameters on the users sense of realism.

5. Is the shoebox model beneficial to an application's visual appeal?

The visual appeal of the shoebox model is indicated by the users wish to see the shoebox model used in different applications.

> **Obj5a:** Find out if users want to see this concept applied more often.

Since the shoebox model is an approach to add extra depth information, it is relevant to compare it to other approaches that strive towards the same goal.

> **Obj5b:** Find out how users think this concept compares to auto stereoscopic screens.

Chapter 2

Related work

2.1 Mobile 3d Concepts

As mentioned in chapter 1, the new technologies present in current day smartphones allow for new 3d concepts.

2.1.1 Fixed World Model

A well known 3d concept on mobile devices is the *fixed world concept* [10]. In this concept users view a virtual world that is aligned with the real world through the screen of their mobile device.

The first research on this concept stems from the early nineties [6] [7]. Before the existence smartphones, Fitzmaurice did research into the concept of a portable spatially aware palmtop that could act as a window to a 3d world. He envisioned a world where physical objects contained virtual information that could be viewed by this palmtop. To test this he built a prototype called Chameleon, which consisted of a palmtop with position and orientation sensors attached that got its images from a camera feed of a desktop computer. The conclusion of his research was that this approach results in similar depth perception as a large static display.

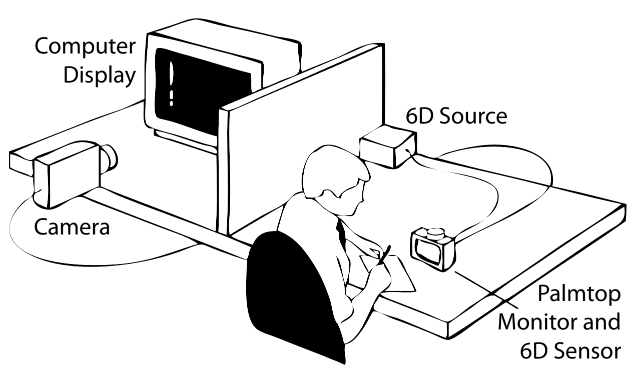


Figure 2.1: A sketch of the Chameleon prototype by Fitzmaurice et al. [7]

Baillie et al. thought of a different application, that allows users in tourist areas to 'time travel' using their smartphone as a window to the past [2]. Three-dimensional models are shown on screen depicting the buildings in their old shape. The fixed world model can be combined with a camera feed to generate an effect called *augmented reality*.

The fixed world model is shown to have several advantages over other types of presenting 360°data. The dynamic motion based scrolling of the viewpoint is proven to increase a users perceived field of view (see section 3.3) by as much as 50% and thereby increasing a users immersion [14]. Users have found this concept visually appealing [12]. However, the need for rotating movement make it not suitable for all application areas [12].

The fixed world model will not be further explored in this thesis.

2.1.2 Shoebox Model and Related Concepts

Another method to realize enhanced 3d graphics on mobile devices is to apply perspective correction to the 3d graphics, to accommodate for the head position of the user. This way, the illusion of an actual space behind (and sometimes in front) of the screen is realized.

The *fishtank model* is an example of a 3d concept using perspective correction [24]. This model is applied to stationary screens and uses *head tracking* to gain information on the user's head position. The position of the users head is then coupled to the amount of perspective correction (head coupled perspective correction). It can also be combined with *stereoscopy* using binocular glasses. Early research suggested that head-coupled perspective correction is more important for giving a strong impression of three-dimensionality than stereoscopy [24]. The fishtank model is also applied to horizontal displays. This way a workbench showing 3d graphics can be created [23]. Stavness et al. even extended the standard fishtank model to a cubic display using 5 LCD screens [16] [22]. The cubic display is hand held and uses a real-time physics engine to simulate having real objects inside a physical box. The user can hold and manipulate this box and indirectly the objects inside it. Head tracking used in the fishtank model requires complicated algorithms and the results are highly dependent on lighting conditions. This can be avoided by using markers on the user's head such as the low-tech implementation using a Wiimote controller and LED markers on glasses by Lee [17]. The fishtank model is successfully applied to mobile devices equipped with a front-facing camera such as the Ipad and Iphone [8].

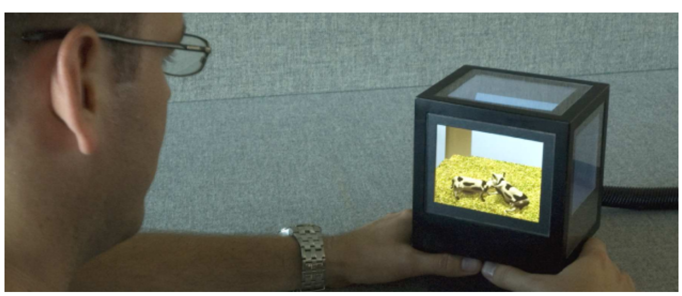


Figure 2.2: The pCubee system by Stavness et al. [16] [22]

The *shoebox model* uses a different approach to the determine the amount of perspective correction [13]. It uses the orientation of the device as a method to approximate the position of the head. So it couples orientation of the device to the amount of perspective correction (device coupled perspective correction). Because of the relative novelty of the shoebox model, the amount of research on the topic is low. The research of Helder [10] suggests that the shoebox model increases users ability to search and locate an object in a 3d space. Also participants in this study generally favoured the shoebox model over the traditional 3d visualisation.

2.2 3D Perception of Computer Graphics

An important goal in 3d graphics is to give users the idea a two-dimensional image has three dimensions. Traditionally, this is achieved by adding enough *pictoral depth cues* to the image [19]. They are two-dimensional cues in the image that suggest depth, such as linear perspective and shading. Research has been done to maximize the effect of these cues in a single image by selecting and applying the most important cues to a scene [5]. However placement and distance of objects have proved more important to 3d perception than rendering effects [5].

Most research concerning 3d perception and computer graphics is done using head mounted displays (HMD) and large screen immersive displays (LSID). One consistent observation is that users consistently underestimate distance in purely virtual environments [18]. However distance perception in augmented reality does not get underestimated [15].

The effect on user 3d perception in the shoebox model has not been researched yet. By adding movement the shoebox model adds a new motion based depth cue. The effect of this addition is researched in this report.

Chapter 3

The Shoebox model

3.1 Design Space

To better describe the new 3d concepts possible on mobile devices, such as the shoebox model, Helder [10] proposes a design space which can be used to define and research new concepts. This design space defines the concepts as a mapping between different coordinate spaces. The three main coordinates spaces are: the real world (RW), the screen world (SW) and the virtual world (VR).

When one coordinate space is mapped to another, it means that their origins are made to coincide and their axes are made parallel to each other. This has the effect of attaching the rotations of one space to the other space. The mapping of one space to another is a symmetric relationship and thus a mapping from A to B is the same as a mapping from B to A. If two spaces are mapped to each other the unattached one can be seen as the remaining degree of freedom.

In the *traditional concept* (also referred to as the *monitor view*) the screen world is mapped to the real world ($SW \Leftrightarrow RW$). The virtual world remains the only degree of freedom. This results in a fixed screen in the real world on which a virtual scene is viewed. The user has only direct control over the VW (moving the VW relative to the SW and the RW). This is the concept used for most traditional computer-games.

In the *fixed world concept* the virtual world is mapped to the real world ($VW \Leftrightarrow RW$). The remaining degree of freedom is the screen world. This results in a concept in which the user can move the screen around to look into the virtual world which is aligned with the real world (moving the SW relative to the RW and the VW) .

In the *shoebox model* concept the screen world is mapped to the virtual world ($SW \Leftrightarrow VW$). The remaining degree of freedom is the real world. So in this case the user can only move the position of the device in the real world (moving the RW relative to the SW and the VW). This results in the intended effect that a space seems to exist behind the screen of the device.

It is possible to extend the design space, by adding a fourth coordinate space, to represent the position of the head. This allows for head tracking to be modelled using this design space.

3.2 Technological

To accomplish the intended effect of the shoebox model, there needs to be a perspective correction applied to the image that is projected on the screen of the device. This perspective correction needs to take into account the orientation of the device.

3.2.1 Shearing Method

The simplest way to accomplish the perspective correction in the shoebox model is to use the *shearing method*. In the shearing method, a normal projection is used to display the scene [9]. Instead of altering the projection, a translation is applied to every vertex in the scene. Vertices further away from the screen get a larger translation. The amount of translation is altered to correspond to the rotation of the device.

3.2.2 Off-Axis Projection

The 'correct' way of implementing perspective correction in the shoebox model is using *off-axis projection*. Off-axis projection is a way of *perspective projection* in which the centre axis of the projection plane is not aligned with the camera position [3]. This is illustrated in figure 3.1.

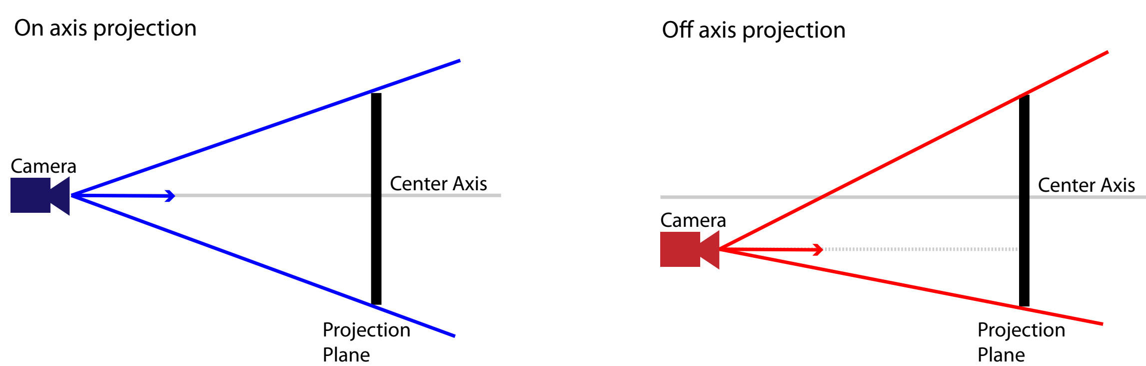


Figure 3.1: A schematic drawing of on-axis perspective projection (left) and off-axis perspective projection (right).

To create the shoebox illusion, the amount of offset of the camera from the centre axis of the projection plane must correspond to the angle of the orientation of the device. If a person looks at the off-axis projection from the front, the image looks distorted. But if a person looks at the off-axis projection from an angle it looks correct.

3.3 Parametric Influences

For the shoebox approach different parameters influence how the user will perceive the effect.

The following parameters are identified as the primary parameters that influence how the shoebox model operates:

shearing factor
This parameter influences the amount at which projection corrected at a certain tilting of the device. It is simply a factor that alters the normal amount of perspective correction. So at 1.0 the effect appear as normal and at 0.0 the effect is disabled and the scene does not react to the tilting of the device.

field of view (FOV)
This parameter determines the field of view of the projection. Field of view is defined as the extent of the observable world that is seen at any given moment. This value also implicitly determines the distance of the camera to the scene, since the camera should have the box nicely fitted in its view.

Other factors that influence the 3d effect perceived by the user include the nature and layout of the scene, the real world environment, the eyes of the user and the experience of the user.

3.3.1 Realistic and Natural values

In this thesis the distinction is made between *realistic* and *natural* values for the different parameters.

Realistic value is the value of a parameter which results in the objectively most realistic 3d projection on the screen. This is the 3d projection that the participant would actually see if there was a real space behind the device containing objects.

Natural value is the value that a user *perceives* as the most realistic. This does not have to equal the actual realistic value.

The realistic values for each parameter can be easily deduced and calculated as can be seen in the corresponding sections(section 3.3.2 and section 3.3.3). The natural values for each parameter are unknown and further examined in the preliminary experiment in chapter 5.

3.3.2 Shearing Factor

The shearing factor is a factor that determines the amount of influence of the tilting of the device on the perspective correction applied to the image on the screen At a value of 1.0 the effect appear as normal, without any modification. At a value of 0.0 the effect is disabled and the scene does not react to the tilting of the device. At any value slightly above or below 1.0 the effect can be decreased or enhanced. In figure 3.2 the effect of altering the shearing factor is illustrated.

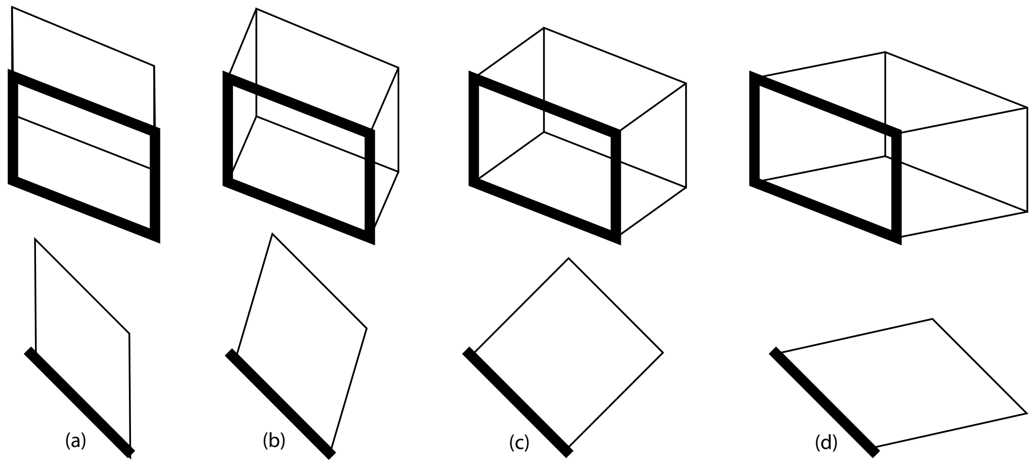


Figure 3.2: A simple illustration showing the effects of the shearing factor of the x-axis on the image. In it, (a) depicts a shearing factor that is 0, so the image is not corrected for the angle of tilt of the device, (c) depicts a shearing factor of 1, so the correction of the image corresponds to the angle of tilt of the device and (b) and (d) depict shearing factors that are respectively lower and higher then 1.

The realistic value for the shearing factor is the easiest to deduce because this is just a matter of definition. Since the shearing factor is simply a factor to enhance of decrease the effect tilting has on the projection, the most realistic value is 1.0. It will give the effect that is most realistic, because the angle at which the device is tilted correspond one on one with the angle used in the perspective correction.

3.3.3 Field of View

The FOV is a standard parameter in computer graphics that determines the extent of the observable world that is seen at any given moment. A low FOV shows observers a small portion of the world, while a high FOV shows observers a large portion of the world. The effect of different FOVs on the image is depicted in figure 3.3.

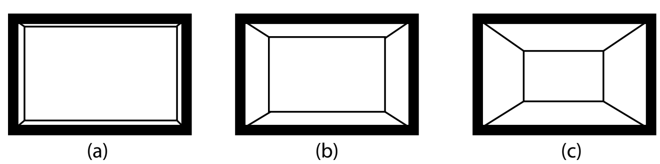


Figure 3.3: A simple illustration showing the effects of the FOV on the image. (a) depicts a low FOV, (b) a medium FOV and (c) a high FOV.

The shoebox model aims to create the illusion that there is a space behind the screen of the device. The FOV must therefore correspond to the viewing angle of the observer. Figure 3.4 shows the FOV in comparison to the viewing angle.

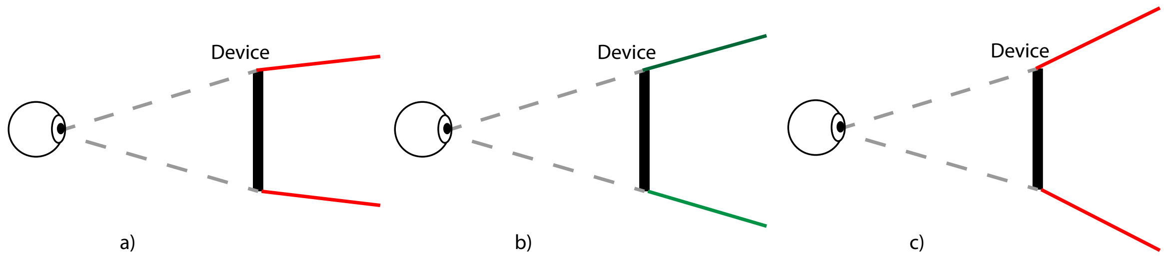


Figure 3.4: An illustration depicting the a FOV that is smaller then the viewing angle (a), equal to the viewing angle (b) and larger then the viewing angle (c).

The most realistic value for field of view equals the viewing angle of the observer. This can be calculated using simple trigonometry. A schematic illustration of this situation is shown in figure 3.5.

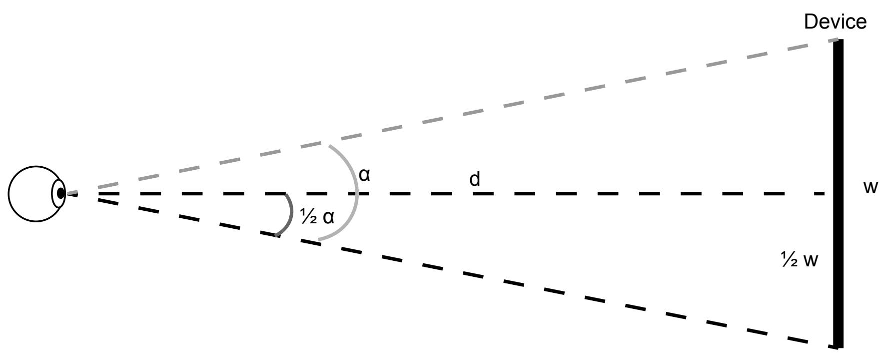


Figure 3.5: A illustration depicting the calculation of the most realistic value for the FOV parameter.

This results in the following equation:

$$tan(0.5\alpha) = \frac{0.5w}{d} \tag{3.1}$$

This equation can be rewritten as:

$$\alpha = 2tan^{-1}\left(\frac{0.5w}{d}\right) \tag{3.2}$$

To calculate the correct FOV, the screen width w and the distance of the observer d need to be known. Unfortunately the values for these variables are not fixed. To calculate a concrete value for the most realistic FOV, several assumptions will have to be made.

The screen size of the device depends on the type of the device and the concrete model. This report focusses on smartphones and not on tablets. The main smartphone used in this report is the HTC Desire, which has a screen width of 8 cm. Because of this, the screen width w is assumed to be 8 cm.

It is more difficult to make assumptions on the distance *d*, since this is highly dependent on the person using the device, the task, external conditions and other factors. When reading from a screen of a mobile phone users typically keep a distance of 17.5 cm to 58.0 cm [1]. From own measurements (with a small number of participants) a average distance of 40 cm was measured. So 40 cm is used as the distance to the device.

Using these assumptions and equation 3.2, the most realistic value for the FOV is approximatly 11°. Because of all the assumptions, the actual realistic FOV value will be in a range around 11°.

Chapter 4

3d Perception

To see what advantages the shoebox model has to a regular monitor view it is needed to look into the area of human 3d perception. This chapter explores depth perception because it is the most clear and measurable part of 3d perception. The experiments conducted in the rest of this thesis will also focus on depth perception for the same reasons.

4.1 Depth Perception

The shoebox model is likely to increase a user's depth perception without the use of stereoscopy. To understand how the shoebox model increases this depth perception it is wise to understand what information a human uses to determine depth in a presented scene.

To determine depth in a presented scene the human visual system uses *Depth Cues*. Different kind of depth cues exist.

4.1.1 Pictorial Depth Cues

Pictorial depth cues (sometimes refered to as *monocular* depth cues) are depth cues that exist in the two-dimensional image that aid the human visual system to interpret this two-dimensional image as three-dimensional.

Because this kind of depth cue depends on an interpretation of 2d information as being three-dimensional, it can cause some ambiguity. This ambiguity is often used in optical illusions, for example the etches of impossible buildings by M.C. Escher.

This is the kind of depth cue traditionally used in computer graphics, since most display systems are only capable of presenting 2d images. It can be computationally expensive to produce images with a wide variety of pictorial depth cues. Real-time applications therefore contain a lower amount these cues or they are less detailed. Table 4.1 explains the different pictorial depth cues.

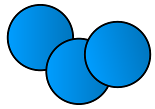
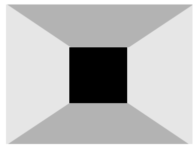
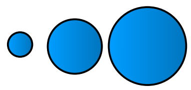
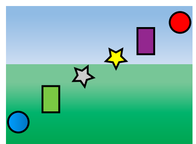
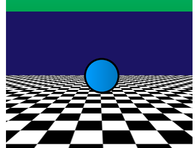
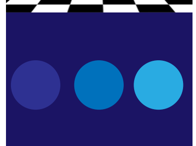
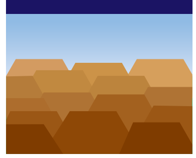
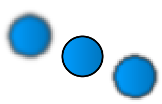
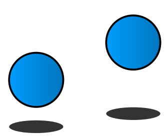
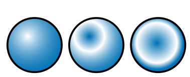
Depth Cue	Picture	Description
Occlusion		When one object is partially occluded by another object, the partially occluded object is perceived as further away.
Linear Perspective		Parallel lines seem to converge in their vanishing point.
Size Gradient		The perceived size of an object is inversely proportional to the distance of the object to the viewer.
Relative Height		Objects closer to the horizon (if present) appear farther from the viewer.
Texture Gradient		Textures get smoother and finer as they stretch further from the viewer.
Relative Brightness		The more intense the brightness of an object is the closer it appears to the viewer.
Aerial Perspective		The scattering of light in the atmosphere causes object further away to appear more blue and hazy.
Depth-of-Focus		The human eye only allows us to focus on one distance at a time. On this distance objects will appear sharp and on other distances they will appear more blurry.
Shadow		Shadow can gives our visual system a lot of depth information. If an object is in the shadow of another object this object is further away from the light source. Shadows on the ground plane can give information on the relative positions of the object that cast the shadows.
Shading		Shading gives information about the shape of object, most importantly, whether the surface of an object is concave or convex.

Table 4.1: The different pictorial depth cues explained [5].

4.1.2 Occulomotor Depth Cues

Occulomotor depth cues are depth cues based on the way the human eye adjusts to change in depth.

When a human looks at an object in 3d space, the eyes both aim at the object. This is called *convergence*. Because of the known angle between the eyes a human can approximate the distance to the object.

The second occulomotor depth cue is *accommodation*. If an objects appears closer to the perceiver, the perceiver need to focus the lenses in its eyes to continue to see object clearly. The curvature of the lenses inside the eyes change as objects move closer. The amount the lens has to accommodate is an indication of how close an object is located.

4.1.3 Binocular Depth Cues

Humans also use the two separate images from each eye to judge depth. This is called *stereoscopy* and uses the *binocular disparity Depth Cues*. The angular disparity between the images is used to estimate the depth of objects in the images. This can only happen if the two images overlap.

Stereoscopic 3d is used in 3d films and other 3d applications to give viewers the implied sense of depth. As a result of the large use, binocular disparity is the most studied depth cue in existing literature [19].

4.1.4 Motion-based Depth Cues

The last category of depth cues is the *motion-based* depth cues and consists of cues that are caused by motion.

Motion allows a viewer to see a scene from different angles and positions. It therefore offers information on the relative depth of different objects in the scene. This cue is called *motion parallax* and was used in old 2d side scrolling games, like Super Mario, to give levels some sense of depth.

Motion can also allow a viewer to see a single object from different angles if the object rotates around its axis. The overall shape of the object is better perceived by the viewer. This cue is called *kinetic depth*.

The information motion-based depth cues offer is comparable to the two different angles at which a viewer will see the world in stereoscopy. Except a viewer will now see the different views not at the same time but at different moments in time.

4.2 Depth Perception and the Shoebox Model

The shoebox model adds and enhances certain depth cues and therefore enhances the depth perception of the viewer.

The most important thing the shoebox model does, is that it introduces motion parallax if it is not present in the scene. Tilting the device, will cause motion in the image of the scene and therefore introduces this cue. To a lesser extend it will also introduce kinetic depth, because the tilting will allow users to look at different sides of an object. It will not however allow them to fully rotate the object and look at the backside of it.

Certain pictorial depth cues will be enhanced by the shoebox model. This is easily explained because the multiple positions from which the scene is viewed allow for multiple sets of pictorial depth cues. These multiple viewing positions can also remove a lot of ambiguity from the pictorial depth cues. Table 4.2 explains the influence of the shoebox model on each of the pictorial depth cues.

Depth Cue	Effect of shoebox	Explanation
Occlusion	Enhanced	The multiple viewing angles allow for multiple chances of objects to occlude each other.
Linear Perspective	Ambiguity removed	Because of the way the lines forming the linear perspective move while tilting the device, one can deduce which of the different possible scenarios is the case.
Size Gradient	Ambiguity removed	By tilting the device, one can deduce if a certain size difference is caused by size gradient or if the objects are just of different sizes.
Relative Height	Ambiguity removed	By tilting the device, one can deduce if a certain the height of an object is caused by its proximity to the horizon or its actual height in the 3d scene.
Texture Gradient	Ambiguity removed	By tilting the device, one can deduce if a visible texture gradient is caused by the perspective or it it simply a texture that changes detail.
Relative Brightness	Ambiguity removed	By tilting the device, one can deduce if the brightness of an object is caused by its depth or if it is simply of a lighter colour.
Aerial Perspective	Unaffected	-
Depth-of-Focus	Unaffected	-
Shadow	Enhanced	The change in perspective allows viewers to better judge which shadow belongs to which object.
Shading	Unaffected	-

Table 4.2: The effect of the shoebox model on the different pictorial depth cues.

The shoebox model has no effect on the occulomotor depth cues. It does not affect how the eyes have to focus and converge to see objects at different distances.

The shoebox model can enhance effect of seeing depth in stereoscopic systems by adding motion parallax to the depth cues, but it does not enhance the binocular depth cue itself. It is mostly an alternative to including the binocular depth, since it also works by allowing a user multiple views on the scene.

Chapter 5

Preliminary Experiment: Comparing realism and naturalness

5.1 General Setup

5.1.1 Aim and Goals

Before any experiments testing the perception of participants under different parameters can be executed, it is necessary to determine what values are useful to examine in the experiments. To do this, more knowledge on the natural values of the parameters is needed. This preliminary experiment is designed to gain more information on the natural values of these parameters and determine some approximate natural values. This experiment is also conducted to gain more practical experience with user evaluations, to be used in future experiments.

The goals of the experiment are as follows:

1. Get a concrete idea of the natural value for the parameters and how much it approaches the realistic value.
2. Test how the number and look of the reference planes influences the natural values.
3. Determine what parameters must be used in the following experiment, in which the users depth perception at different parameters is tested.

5.1.2 Main Task

The experiment consists of a number of scenes presented to the participant. The participants are asked to tweak certain parameters of the shoebox model until it appears most natural. The participants have control over these parameters using on screen buttons, and are asked to hit the 'ok' button if they are satisfied with the current settings. At this point the next scene will be presented to them. There is no time-limit imposed on this challenge. The participants are told beforehand how in many scenes they must tweak these parameters so they can adjust their speed to their liking.

5.1.3 Dependent Variables

In the experiment the participant is asked to tweak certain parameters until the result appears most natural to him. These parameters are called the *dependent variables*.

They are picked, because they have a clear influence on the appearance of the shoebox visualization and therefore the naturalness of the 3d effect. The following variables are picked:

Parameter	Realistic Value	Natural Value
Field of View	11	Unknown
Shearing Factor	1.0	Unknown

Table 5.1: The most realistic values for the different parameters. The most natural values are unknown.

For more details on these parameters see section 3.3.

It is expected that users will pick a slightly bigger natural value compared to the realistic one for the shearing factor to compensate for the lack of stereoscopic vision. It is expected that for the FOV users will pick a larger number as the most natural value, since they are used to bigger values for the field of view (console games traditionally use a field of view of 60° [20]).

Other factors that influence the 3d effect include the nature and layout of the scene, the real world environment, the eyes of the user and the experience of the user. These are left out as dependent variables because of they are difficult to quantify.

5.1.4 Independent Variables

If a scene has any planes that help to clearly see changes in perspective, we call these planes reference planes. How many reference planes a scene has, will likely influence the participant and therefore different number of reference planes is tested. So, as the first *independent variable* the number of reference planes in the scene is used. It is expected that the number of reference planes is influential in the ability of participant to judge how realistic a certain setting appears.

The look of these reference planes is also expected to have an influence. As a second independent variable the texture of the reference planes is used. Both a fully textured as well as a wire frame variants of the reference plane are used as values for the independent variable.

As a third independent variable the number of parameters a participant must tweak at the same time is used. If the users only has to tweak one parameter the other is fixed at its realistic value.

5.1.5 Participants and Environment

The ten participants, volunteers known by the author, were in the age range 20 to 25. Half of them had a background in computer science or related studies. Two of them were female, the rest male. Since this is just a preliminary experiment a low number of participants is justified. Because of this low number of participants no statistical analysis could be performed on the results.

The experiments took place in living rooms under quiet conditions. The participants were asked a couple of questions and a verbal explanation of the experiment was given.

5.2 Details

This section describes the experiment in detail.

5.2.1 Experiment Overview

1. Introduction to shoebox model (`Ref5`)
2. Tweaking values independently (order counterbalanced)
   (a) Tweaking shearing factor value (all scenes)
   (b) Tweaking FOV value (all scenes)
3. Tweaking shearing factor and FOV value at the same time (all scenes)

At the beginning and end of the experiment the participant is asked to answer a couple of questions (see questionnaire A.1).

In *part 2* the order of the two subtests *a* and *b* is counterbalanced to eliminate learning effects. *Part 1*, *part 2* and *part 3* always appear in order because the learning effect is useful in this case.

In *part 1* the participants are introduced to the shoebox model. This is to make sure they have a common understanding of the shoebox model. So this part must take place before the others.

In *part 3* the participants have to tweak both parameters at the same time. Because this can get intimidating and hard, a thorough understanding of the parameters is useful. That is why this test appears at the end of the experiment.

The last point is important to ensure that participants get some idea of what they are doing in scenes with a high number of reference planes before moving to the scenes with a low number of reference planes. These four options are presented in counterbalanced order so each option will appear approximately the same number of times.

As values for the fixed parameters the realistic values for the parameters are used.

As initial values for the tweak-able parameters the following values are used:

Shear 0.0 - This value makes sure there is no perspective shift when the device is tilted. It is chosen so the participant clearly sees the value is wrong and will adjust it to a higher value to approach the natural value.

FOV 60 - This is the value used most in console games and clearly too big. It is chosen so the participant clearly sees the value is wrong and will adjust it to a lower value to approach the natural value.

5.2.2 Scenes

Six scenes are used during the experiment.

All the scenes include basic shapes to avoid influence of absolute size of known objects. The shapes are distributed across the whole space of the box. The shapes are given primary colours to avoid influence of certain colour on the scene. Also the combination of red and green is avoided, as not to cause problems for people who have red-green colour blindness.

If a scene has any planes that help to clearly see changes in perspective, we call these planes reference planes. How many reference planes a scene has, will likely influence the participant and therefore different numbers of reference planes are tested. It is also tested whether the look of a reference plane will influence the participant, thus both fully textured as well wire frame variants are present.

This results in the following scenes (also see figure 5.1):

`Ref5`: Basic configuration in an 8x4.8x8 box.

`Ref5w`: Same as `Ref5` but the reference planes are represented as wireframes.

`Ref1`: Basic configuration on a ground plane.

`Ref1w`: Same as `Ref1` but. the reference planes are represented as wireframes.

`Ref0`: Basic configuration in an empty scene.

`Ref0a`: A bunch of randomly floating shapes in a empty scene.

The scenes `Ref5`, `Ref1`, and `Ref0` each have a different number of reference planes. The goal is to test the influence of reference planes on the ability to judge the natural values of the parameters of the shoebox model.

The box that is used in`Ref5` and `Ref5w` simulates a box of 8cm of depth behind the screen of the device. For the screen size, the size of the HTC desire is used, which is 8cm x 4.8cm. This results in a box of 8x4.8x8.

In `Ref0` the lack of reference planes makes it harder to see the 3d effect of the shoebox approach. To compensate, a higher density of objects is used in a second scene called `Ref0a`. In this scene 50 red cubes and 50 blue and yellow spheres are randomly rotated, scaled and placed at random positions within the scene.

In *part 1* only `Ref5` is used. In *part 2a*, *part 2b* and *part 3* of the experiment all of the different scenes are used. Their order is determined by four options:

- `Ref5 - Ref1 - Ref0 - Ref5w - Ref1w - Ref0a`
- `Ref5 - Ref1 - Ref0a - Ref5w - Ref1w - Ref0`
- `Ref5w - Ref1w - Ref0 - Ref5 - Ref1 - Ref0a`
- `Ref5w - Ref1w - Ref0a - Ref5 - Ref1 - Ref0`

These four possible orders accomplish three things:

- It counterbalances whether the wire-framed variants or the normal scenes are presented first.
- It counterbalances which one of the two `Ref0` variants is used at which time.
- It ensures the scenes are always presented in an order form a high number to a low number of reference planes.

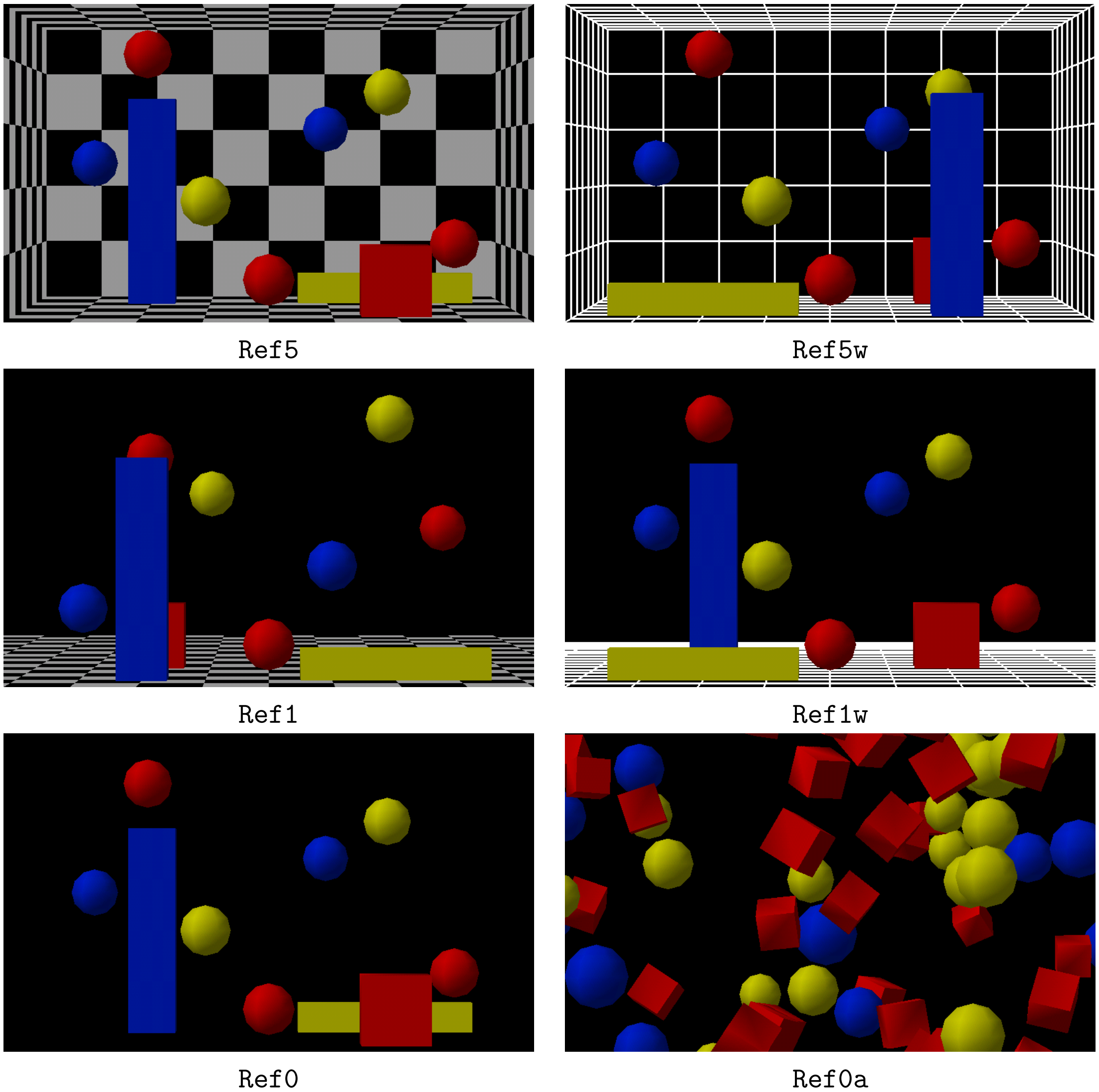


Figure 5.1: The scenes used in the experiment

5.2.3 Position of the Objects

To rule out any influence by the placement of the objects, different positions for the objects are used. To decide the position of the box-like objects in every scene (except `Ref0a`) the floor plane is divided in 4 equal quadrants and the three box-like objects are placed in the middle of three of the four quadrants. All the possible orderings on these quadrants appear in a counterbalanced order to the participants.

For the floating spheres in the every scene, one of two possible sets of positions is chosen at random. These sets each have different coloured spheres distributed so each cm on the horizontal axis of the screen has one sphere. The spheres are also distributed at different heights and depths so the entire space is used.

5.2.4 User Interface

During the test the participant is asked to tweak two different parameters. Depending on which parameter the participant is asked to tweak, different interface elements can be shown:

On the left side the element to tweak the shearing factor parameter is shown, which consists of:

- Plus button that raises the shearing factor value by 0.1 to a maximum of 5.0.
- Minus button that lowers the shearing factor value by 0.1 to a minimum of 0.0.
- An icon that indicates the function of the plus and minus button. The function is also explained in text at the beginning to make sure the user understands what the parameter means.

On the right side the element to tweak the FOV parameter is shown, which consists of:

- Plus button that raises the FOV value by 1 to a maximum of 120.
- Minus button that lowers the FOV value by 1 to a minimum of 4.
- An icon that indicates the function of the plus and minus button. The function is also explained in text at the beginning to make sure the user understands what the parameter means.

In the middle bottom, just one button is located:

- OK button which confirms the chosen parameters and continues to the next scene.

The minimum and maximum values for the FOV and shearing factor value are both chosen because they are extreme values that would yield results which are too exaggerated. This way the participants clearly see values are not realistic.

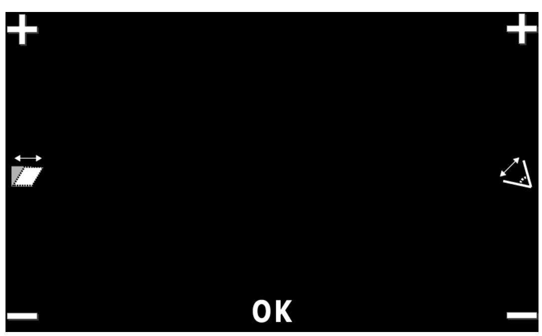


Figure 5.2: The gui used in the experiment

5.3 Hypotheses

The following hypotheses are formulated about the outcome of the experiment:

1. The natural FOV value is higher then the realistic FOV value.
    - Since the participants are used to high field of views in existing applications, such as console games.
2. The natural shearing factor value is higher then the realistic shearing factor value.
    - Since the participants have to compensate for lack of stereoscopic vision to still have the same experience.
3. A higher number of reference planes will result in natural values that are closer to the realistic values.
    - This is because reference planes help them to better judge the currently set parameters toward values that they are used to in every day life.
4. When full textures are used in the reference plane(s) instead of wire frames, the natural values will be closer to the realistic values.
    - Because when a reference plane is fully textured the entire surface can be used as a reference instead of just the edges connecting the vertices.
5. More objects in the scene will result in natural values that are closer to the realistic values.
    - Because the participant can use more objects as reference.

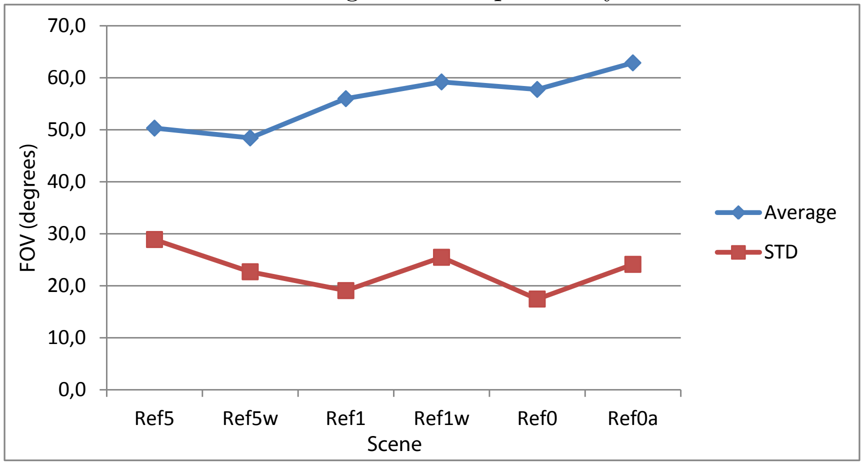
5.4 Results

5.4.1 Natural FOV

Hypothesis 1: "The natural FOV value is higher then the realistic FOV value."

The results in figure 5.3 show that participants consistently pick natural values for the FOV that are a lot higher than the realistic 11°degrees. The actual value that is picked is only slightly below the 60 degrees that console games use (between 50°and 60°). So hypothesis 1 is **confirmed**.

Tweaking FOV independently:



Tweaking FOV when also tweaking the shearing factor:

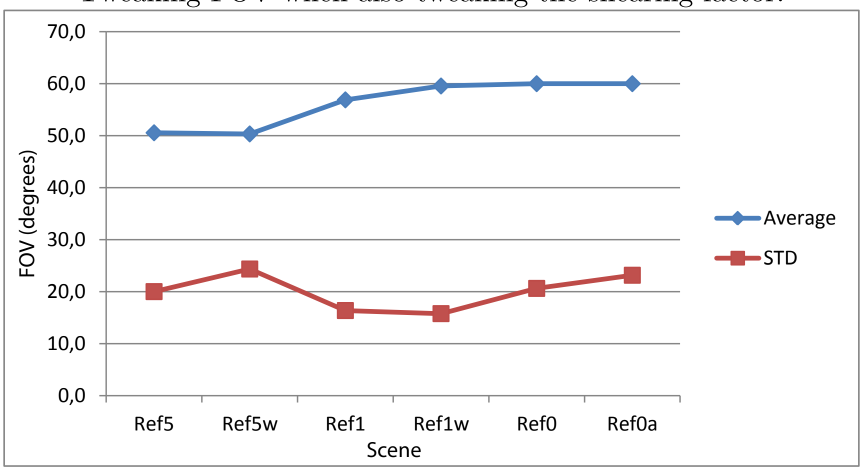


Figure 5.3: Results from the experiment part 2 and 3 on the FOV. The first chart depicts the average selected FOV when tweaking the FOV independently. The second chart depicts the average selected FOV when tweaking the FOV together with the shearing factor.

The FOV is initialized at 60° when participants have to pick a value for the field of view, so this could influence the participants. The participants however, were first shown a scene that had a realistic FOV of 11°(although they were not told that this was the most realistic). Furthermore, when just tweaking the shearing factor, the participants were also presented a FOV of 11°. So it is hard to say how much influence the initial 60° had on the participants judgement of the naturalness. However, this does not devalue the fact that participants are more used to high FOVs in applications. One participant even made a remark that it was difficult for him to tell whether certain FOV is looked natural, since they all looked natural to him. He explained this by telling that he used to play competitive computer games and changing the FOV all the time to have an advantage over other players (seeing more versus accurate aiming). Another participant asked how deep the box was and told that that would influence the FOV he would see as most natural. In the final questionnaire most participants also indicated the FOV was difficult to set to a natural value. This indicates that people are used to varying FOVs and it is hard for them to pick one field of view over another. They see them as more or less equally natural. They tend to pick a FOV close to the 60° they are used to in console games and other applications.

5.4.2 Natural Shearing Factor

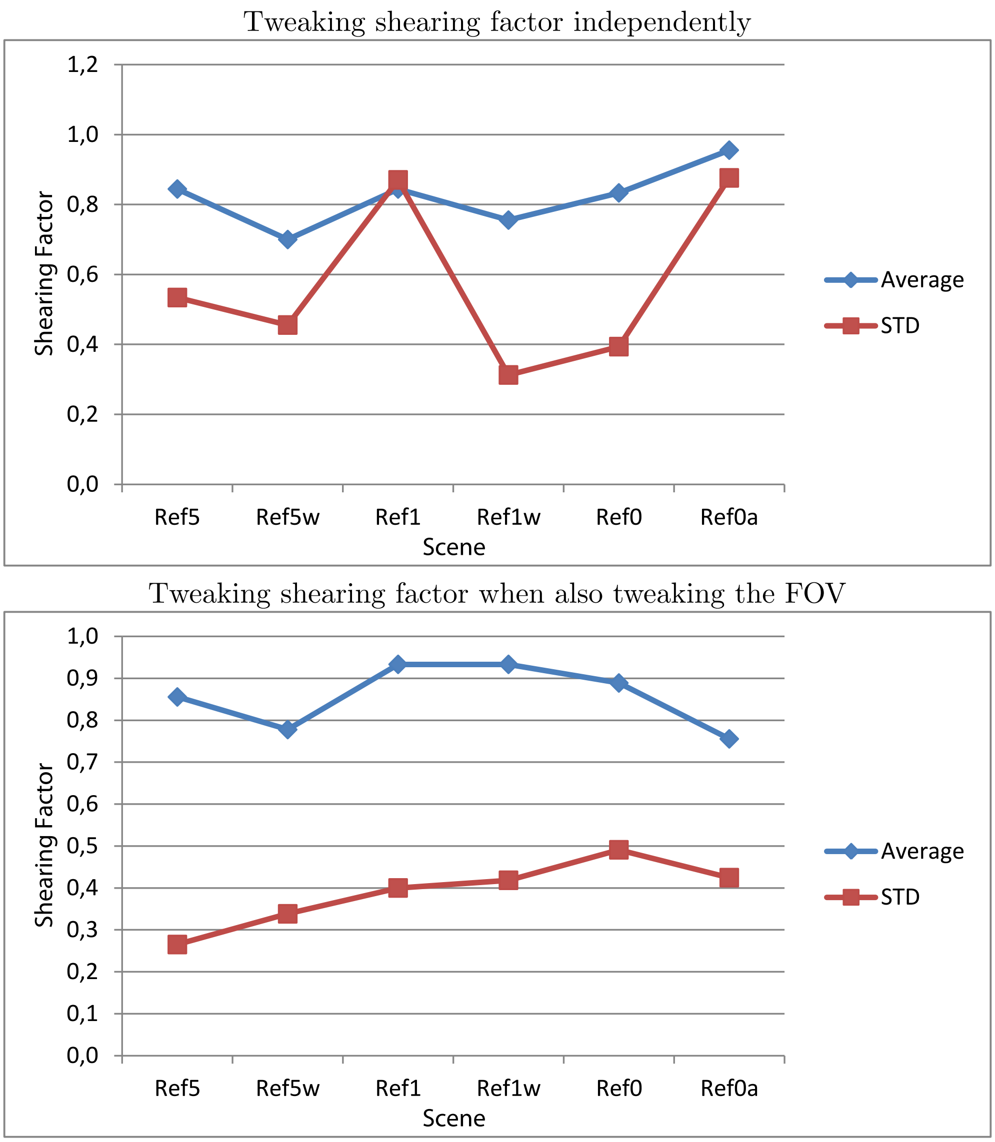


Figure 5.4: Results from the experiment part 2 and 3 on the shearing factor. The first chart shows the average selected shearing factor when tweaking the shearing factor independently. The second chart shows the average selected shearing factor when tweaking the shearing factor together with the FOV.

> Hypothesis 2: "The natural shearing factor value is higher then the realistic shearing factor value."

The results in figure 5.4 show that participants consistently pick values for the shearing factor that are below the most natural value of 1.0. The value picked is mostly between 0.7 and 1.0. This is completely against the expectation. So hypothesis 2 can be **rejected**.

Participants noted visual distortions of the objects, particularly the spheres. This can be explained by the following images:

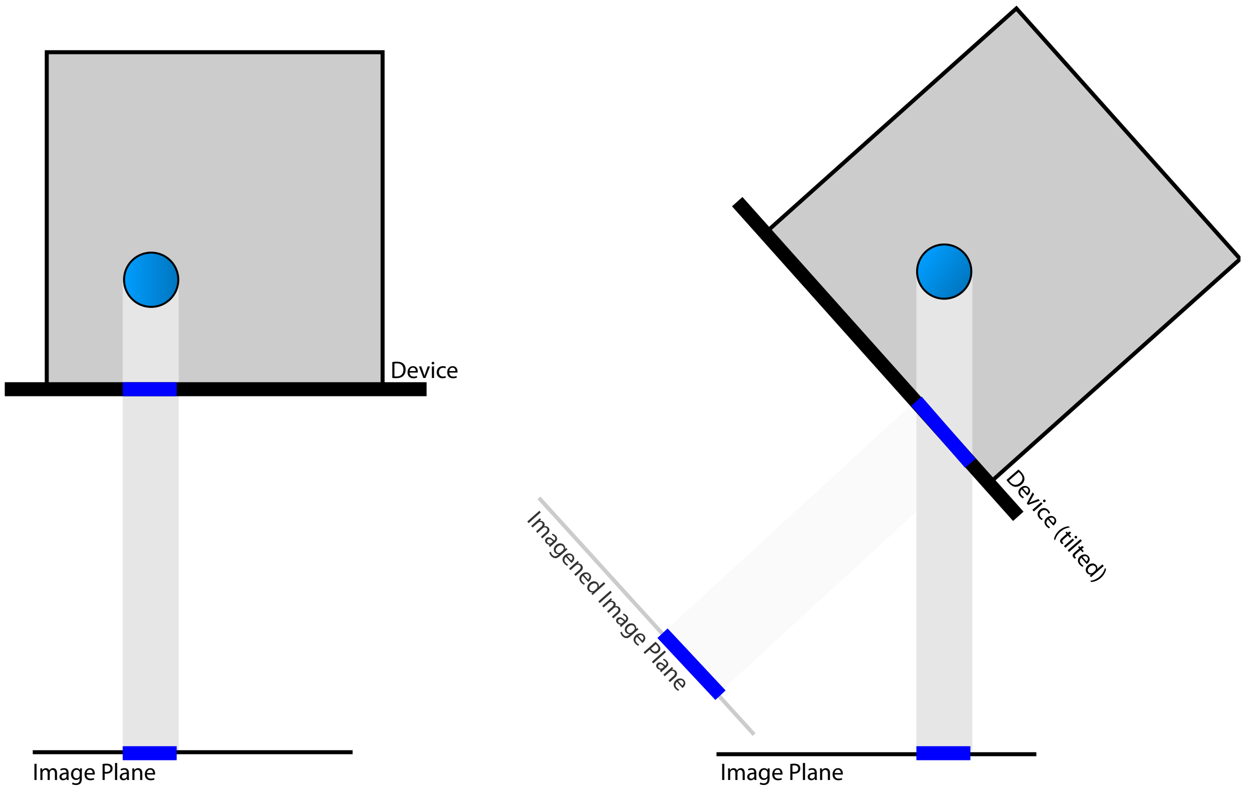


Figure 5.5: The projection to the image plane if the device is held straight (left) or tilted (right).

In the left image of figure 5.5 the normal situation, when the device is not tilted, is visible. In this image, the device and the imaginary scene behind the device are shown. The image plane represents the human visual system. It can be clearly seen that the projection of the blue sphere on the screen of the device and on the image plane result in no visual distortions of the sphere.

In the right image of figure 5.5 the situation a tilted device is shown. It can be seen that the projection on the screen of the device is visually distorted. However since the image plane is located at an angle to the device, the projection on the image plane is distorted back and the user should perceive an undistorted image. Several participants however noted the visual distortion on the screen and even used them to judge the correctness of the current shearing factor. An explanation for this is that people are used to watch screens at an angle (while watching television, reading information panels of road signs, etc.). So mentally, it seems as though they look at the screen from the front. This is illustrated in by the imaged image plane. It can be seen that the projected image is still as distorted as the image on the screen of the device.

This offers an explanation as to why the shearing factor was selected so low. The participants want to keep visual distortions to a minimum and therefore select a lower shearing factor. This effect may be enhanced because of the stereoscopic view of the participants. This enables them to more clearly see image on the screen of the device as a flat image and interpret it as something that should be viewed from the front.

5.4.3 Influence of Scene

> Hypothesis 3: "A higher number of reference planes will result in natural values that are closer to the realistic values."

When participants had to select a value for the FOV, the participants selected a higher value (and thus further from the 11°) for the FOV as the number of reference planes decreased. This is visible both in the 2nd part and the 3rd part of the experiment. However 50°(in the case of the 5 reference planes) is still very far from 11°. Nonetheless, in the case of the FOV, hypothesis 3 can be **confirmed**.

However, when the participants had to select a value for the shearing factor the effect was not visible. In the 2nd part of the experiment the effect even seems to reverse, although the difference of not significant enough to draw any conclusion. The lack of the effect on the selection of the right shearing value is probably caused by person using the deformation the spheres as a reference and not the reference planes. So in the case of the shearing factor, hypothesis 3 can be **rejected**.

> Hypothesis 4: "When full textures are used in the reference plane(s) instead of wire frames, the natural values will be closer to the realistic values."

The experiment showed no real difference between the fully textured reference planes and the wire-frame reference planes. This suggests that the users use the lines on the reference plane (in this case the checkerboard pattern) as the guide to judge depth and not the surface of the plane. Since both the checker-board pattern and the wire-frame have the same reference lines on them it makes no difference which one is used. Hypothesis 4 is therefore **rejected**.

> Hypothesis 5: "More objects in the scene will result in natural values that are closer to the realistic values."

Sometimes `Ref0a` performs better then `Ref0`, sometimes the other way around. This is most likely due to the fact that it is difficult to see any real changes (as some participants told) if one changes the parameters in these scenes. This inability is most likely caused by the lack of reference planes. This indicates that clear reference planes are important of the judgement of realism. Hypothesis 5 can be **rejected**.

5.5 Conclusions

The first goal of the experiment was to get a more concrete idea of the natural values of the parameters. The results can be seen in table 5.2.

Parameter	Realistic Value	Natural Value
Field of View	11°	50°–60°
Shearing Factor	1.0	0.8–0.9

Table 5.2: The most realistic and natural values for the different parameters.

The second goal was to find out how the number and look of the reference planes influences the natural values. From the experiment it can be concluded that a lower number of reference planes results in a natural values that are farther away from the realistic values. Whether the reference plane is textured does not seem to influence the natural values. This is probably since users use the lines of the reference plane tot judge the scene and as long as the wire-framed version consists of the same visual lines it has the same effect on the naturalness of the shoebox model.

In the follow up experiment the depth perception under different parameter values will be tested. The third goal was to determine educated values for the parameters to be used in this experiment. In the next experiment it is wise to pick FOVs at values that are relevant because the are used in other applications, since the experiment indicated a large influence of existing application on the natural value. The values that are wise to be examined are 11°, 30°, 45°, 60° and 90°.

In the follow up experiment is wise to pick values for the shearing factor at 0.6, 0.8, 1.0, 1.2 and 1.4. This way a fair distribution of values is picked in the range that the experiment suggested while still looking at value above this range because this probably increases depth perception.

Chapter 6

Experiment: Relative Order Depth Perception

6.1 General Setup

6.1.1 Aim and Goals

The preliminary experiment in chapter 5 has given us some more information on the natural values of the parameters. The following experiment is designed to test the influence of the shoebox model with varying parameters on the depth perception of users. It tests the ability of users to judge the relative order of two objects. This will be referred to as relative order depth perception. Also the influence of some aspects of the scene will be tested. To accomplish these aims the following goals are formulated:

1. Test how the shoebox model in general, with realistic values for its parameters, influences the relative order depth perception of the participants, with objects at different distances.
2. Test the how the FOV parameter influences the relative order depth perception of the participants.
3. Test how the shearing factor parameter influences the relative order depth perception of the participants.
4. Test how the existence and look of a reference box influences the relative order depth perception of the participants.
5. Test how the position of the objects influences relative order depth perception of the participants.

6.1.2 Main Task

In this experiment the participants are asked to perform a certain task under different conditions. During this task the participants are shown two spheres at the same time. They are asked to select the sphere that is closest to the viewer. A single instance of this test will be referred to as a *depth test* in the remainder of this section.

6.1.3 Dependent Variables

In this experiment the dependent variable will be the percentage of correctly judged depth tests at each test. This is used as a measurement of how good the depth perception of users is under certain settings.

6.1.4 Independent Variables

In this experiment the independent variables consist of the FOV, the shearing factor and the scene.

For the FOV, the values that are examined are 11°, 30°, 45°, 60°and 90°. These are chosen because they are important often used values in existing applications. A influence of the values in existing applications was suggested in the results from the preliminary experiment (chapter 5).

For the shearing factor, the values that are examined are 0.6, 0.8, 1.0, 1.2 and 1.4. These are chosen because the preliminary experiment (chapter 5) suggested that natural values consist around the most realistic 1.0.

As a third independent variable the scene is used. Three scenes are used in the experiment: box with checker-board pattern, box with evenly coloured grey surfaces and no box at all.

6.1.5 Participants and Environment

Most participants for this test were students participating in the "Ontwerpen Inactivive Systemen" course at Utrecht University in 2011. Their presence was mandatory for the course. This resulted in 29 participants and this number was increased to 32 by three volunteers. All participants had a background in computer science. The participants consisted of 2 women and 30 men.

The experiments were conducted in a computer room in which several other experiments were being conducted at the same time. The experiments were done with two participants at the same time if possible (each on a different device). Two devices were used in the experiment: a HTC Desire smartphone and a Samsung Galaxy S smartphone. Each device was used in 16 of the experiments.

6.2 Details

6.2.1 Experiment Overview

At the beginning of the experiment a general explanation of the shoebox model is given to the participant. The participant is shown the shoebox model and can play around with it a bit. The participant is then asked to answer a couple of questions (see questionnaire A.2). Next, the participant is asked to sit down, hold the device a certain distance from the head (+/- 40 cm) and begin the experiment.

1. Introduction (3 depth tests)
2. Monitor vs Shoebox Test (2 x 27 depth tests)
3. Testing different values (a and b in counterbalanced order)

(a) 5 different Shearing factor values $(5 \times 9 depthtests)$

(b) 5 different FOV values $(5 \times 9 depthtests)$

Test 2 and 3 are then repeated at different scenes until all scenes are being tested.

Test 2 and 3 appear in counterbalanced order to ensure the influence of order is dealt excluded from the results. Test 3a and 3b also appear in counterbalanced order for the same reason.

The positions of the spheres appear in counterbalanced order. The sub positions of the spheres appear also appear in in counterbalanced order. The definition of these positions and sub positions is given in section 6.2.3.

At last, the participant is asked to answer a couple of follow up questions (see questionnaire A.2).

6.2.2 Scenes

Three scenes are used in this experiment, `sCB`, `sP` and `sNB` (see figure 6.1). The first two scenes, `sCB` and `sP`, consist of the inside of a box. The box that is used simulates a box of 8 cm of depth behind the screen of the device. For the screen size, the size of the HTC desire is used, which is 8 cm x 4.8 cm. This results in a box of 8x4.8x8. The walls of the box `sCB` have a simple checkerboard pattern. The walls in `sP` have a plain grey texture. Scene `sNB` consists of an empty scene.

sCB sP sNB

Figure 6.1: The three scenes used in the experiment.

6.2.3 Position of the Objects

The target spheres can appear in four different orientations with respect to each other. They are placed on opposite corners of a theoretical subbox of 2x2x3 cm as depicted in figure 6.2.

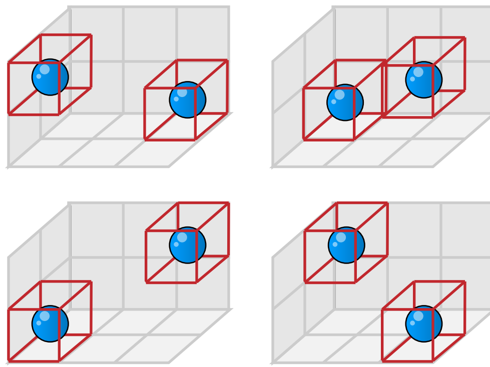


Figure 6.2: The four possible orientations of the target spheres inside a sub box.

This way the spheres are always 1 cm apart in depth and there is a nice horizontal gap between the spheres so they will not overlap when the device is tilted. These subpositions of spheres appear in counterbalanced order. These subboxes are positioned in the large box at different coordinates. They are positioned in a 3x3x3 grid inside the big box so the sub boxes touch the corners of the big box. These positions appear in counterbalanced order.

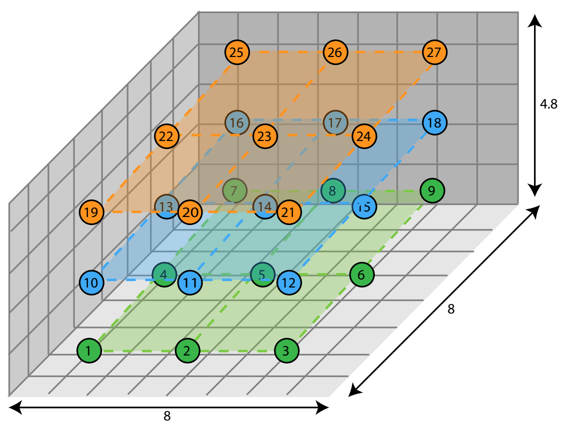


Figure 6.3: The positions the subboxes in the scene.

Some tests use a subset of all possible positions. In this case positions are picked so they appear in the whole space. These subsets appear in counterbalanced order.

6.2.4 Colour of the Spheres

Both spheres always have the same colour so the difference in colour cannot be of influence when determining which is closest. To rule out the influence of this colour on the perception, three different primary colours are used: red, blue and yellow. One colour is randomly picked each time the two spheres need to be relocated. Because of the large number of depth tests a randomized approach is justified.

6.2.5 User Interface

During the test the participant is asked select the sphere that is closest to the screen. This done using two buttons:

- The left button which is located on the left side of the screen and consists of the word *Left*.
- The right button which is located on the right side of the screen and consists of the word *Right*.

6.3 Hypotheses

To support the goals of the experiment, several concrete hypotheses are formulated:

1. The shoebox model will increase the users' perception of relative order of objects.

   Reason Because of the addition of the motion parallax depth cue.
   Tested by Test 2

2. The checker-board scene will yield better results then the plain scene.

   Reason Because the checker-board texture allows for a lot of reference lines the users can follow with their eyes.
   Tested by Test 2

3. The empty scene will yield worse results then the other scenes.

   Reason Because the lack of reference plains will negatively influences the ability of users to correctly judge the distance of the spheres.
   Tested by Test 2

4. A bigger shearing factor value will increase the users' perception of relative order of objects.

   Reason Because a bigger shift in perspective will emphasize the different depths of two objects by increasing the relative movement between them while shifting.
   Tested by Test 3a

5. A bigger FOV value will increase the users' perception of relative order of objects.

   Reason This is because a larger FOV will increase the visible distance in depth of two objects.
   Tested by Test 3b

6. It will be easier to judge order on objects close to the screen.

   Reason Because the distance between the objects will appear to be more when the relevant objects are closer to the screen.
   Tested by All tests

6.4 Results

6.4.1 Influence of the Shoebox Model

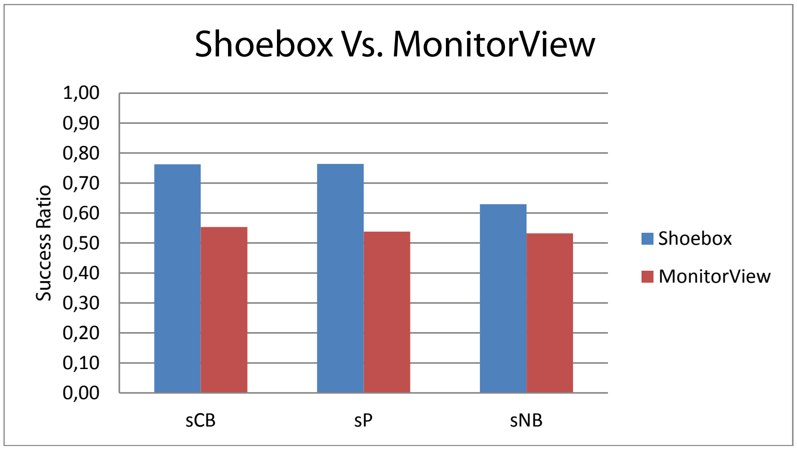


Figure 6.4: Results of part 2 of the experiment. Comparing the monitor view against the shoebox approach and checking the influence of the scene.

> Hypothesis 1: "The Shoebox model will increase the users' perception of relative order of objects."

Part 2 of the experiment mainly focusses on the influence of the shoebox approach compared to the traditional monitor view on the depth perception. The results of this part of the experiment are presented in the chart in figure 6.4. The chart indicates a higher success rate when the shoebox approach is used compared to when the monitor view is used. On the `sNB` scene this difference is less then on the other scenes, but still very clear.

To see if this visible difference is significant a statistical analysis was performed. The statistical analysis consists of a simple students t-test on each scene between the two concepts. The results are shown in table 6.1. This table shows that in all three cases the t-value is below 0.05, meaning the observed differences are statistically different. This shows that hypothesis 1 can be **confirmed**.

T-Test	Scene	T-Value
Monitor-Shoebox	`sCB`	0.000
	`sP`	0.000
	`sNB`	0.004

Table 6.1: Results of paired two tailed t-tests between the success rate on the Monitor View and the success rate on the Shoebox Model in part 2 of the experiment.

Another remarkable observation is that the success ratio on the monitor view is not the expected 0.5 but averages around 0.54. The random chance of 0.5 was expected because important depth queues such as relative size, occlusion and shadows are not visible. However small variations of lighting on the spheres offer an explanation for a slight increase in chance on picking the correct sphere.

6.4.2 Influence of Scene

In the result of part 2 of the experiment, shown in figure 6.4, the influence of the different scenes can also be observed.

The first thing that can be noticed is that there seems to be no clear difference between scenes in the monitor view. This can be explained by the fact that inclusion reference plains offers no clear advantage in the monitor view. It will not allow users to better judge changes in perspective, since there are none.

> Hypothesis 2: "The checker-board scene will yield better results then the plain scene."

However with the shoebox model applied, different observations can be made. It was expected that the checker-board scene (`sCB`) would yield better results then the plain scene (`sP`). But this pattern can not be observed in the results. The two scenes seem to have a roughly equal success rate. This is further supported by the statistical analysis in table 6.2 which shows no statistical difference between the success rates in the the two scenes. This shows that this hypothesis 2 must be **rejected**.

A explanation for this is that two parallel lines in the shoebox model are enough to define a reference plane. It must be noted that the lines do not appear parallel due to perspective projection. The additional lines that the checker-board offers do not add any real information to the user on the shape of the plane.

> Hypothesis 3: "The empty scene will yield worse results then the other scenes."

However there is a clear difference between the success rate on the two boxed scenes (`sCB` and `sP`) and the empty scene `sE`. The empty scene shows a lot lower success rate then the boxed scenes. This is supported by the statistical analysis in table 6.2 which states that the differences with both pairs of scenes compared to the empty scene are statistically significant. So it can be concluded that reference planes are important to the ability to judge depth in the shoebox model. This shows that this hypothesis 3 can be **confirmed**.

Concept	T-Test	T-Value
Shoebox	`sCB - sP`	0.968
	`sP - sNB`	**0.002**
	`sNB - sCB`	**0.001**
Monitor	`sCB - sP`	0.558
	`sP - sNB`	0.764
	`sNB - sCB`	0.348

Table 6.2: Results of paired two tailed t-tests between the success rates on different scenes in part 2 of the experiment. Significant values are emphasized by a bold font.

6.4.3 Influence of Shearing Factor

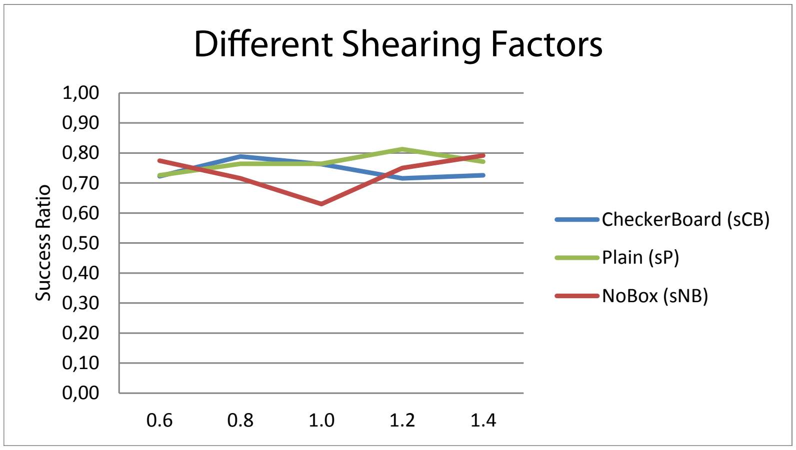


Figure 6.5: Result of part 3a of the experiment. Comparing the success rate on different values of the shearing factor.

> Hypothesis 4: "A bigger shearing factor value will increase the users' perception of relative order of objects."

In part 3a of the experiment different values for the shearing factor are tested to see how the success rate is influenced by these values. The results of this part of the experiment are presented in the chart in figure 6.5. The chart shows no clear relationship between the shearing factor and the success rate.

Also ANOVA statistical analysis (shown in table 6.3) confirms the null hypothesis that all results are drawn from the same distribution. This can be concluded because the F values are lower then the F-Crit values on all scenes.

This indicates no relationship between shearing factor and success rate on all of the three scenes. It was expected that a higher shearing factor would result in a better success rate because a bigger shift in perspective would emphasize the different depths of the two objects by increasing the relative movement between them while shifting. The relative movement between object is of course increased by a high shearing factor but this does not result in the expected better success rate. So this increase of relative movement is not of influence on judgement of order. This can be explained by the fact that a low relative movement already emphasizes the fact that one object is closer then the other. A larger relative movement will not influence this.

In the experiment it was only tested if a user could judge the relative order of two objects. The ability to see if one object is closer then another is a qualitative judgement of order. A user either judges it correctly, or does not.

However, if one would introduce another test, a influence might become apparent. The possibility exists that a higher shearing factor does increases the ability of users to make quantitative judgements on the distance between objects. Another possible test is to see if a user can judge size of objects. So a user will not increase the ability to judge which object is closer, but might increase the ability to perform other tests that test 3d perception.

This means that hypothesis 4 has to be **rejected**, under a relative order depth test.

ANOVA

Scene	Source of Variation	SS	df	MS	F	P-value	F crit
sCB	Between Groups	0,1065	4	0,0266	0,984	0,4182	2,4320
	Within Groups	4,0556	150	0,0270			
	Total	4,162	154				
sP	Between Groups	0,1454	4	0,0363	1,4123	0,2326	2,4320
	Within Groups	3,8616	150	0,0257			
	Total	4,0071	154				
sNB	Between Groups	0,4713	4	0,1178	2,6041	0,0382	2,4320
	Within Groups	6,7868	150	0,04524			
	Total	7,2581	154				

Table 6.3: ANOVA statical analysis of the influence of the shearing factor on success rate in part 3a of the experiment.

6.4.4 Influence of Field of View

> Hypothesis 5: "A bigger FOV value will increase the users' perception of relative order of objects."

In part 3b of the experiment different values for the FOV are tested to see how the success rate is influenced by these values. The results of this part of the experiment are presented in the chart in figure 6.6. The chart shows no clear relationship between the shearing factor and the success rate.

Also ANOVA statistical analysis (shown in table 6.4) confirms the null hypothesis that all results are drawn from the same distribution.

This indicates no relationship between FOV and success rate on all of the three scenes. It was expected that a higher FOV would result in a better success rate because a larger FOV will increase the visible distance in depth of two objects. This bigger visible depth will result in a larger relative movement between the objects. However, as shown in section 6.4.3, a low relative movement already emphasizes the fact that one object is closer then the other.

This means that hypothesis 5 has to be **rejected**, under a relative order depth test. Other 3d perception tests are necessary to draw any further conclusions on the influence of FOV on the 3d perception of users.

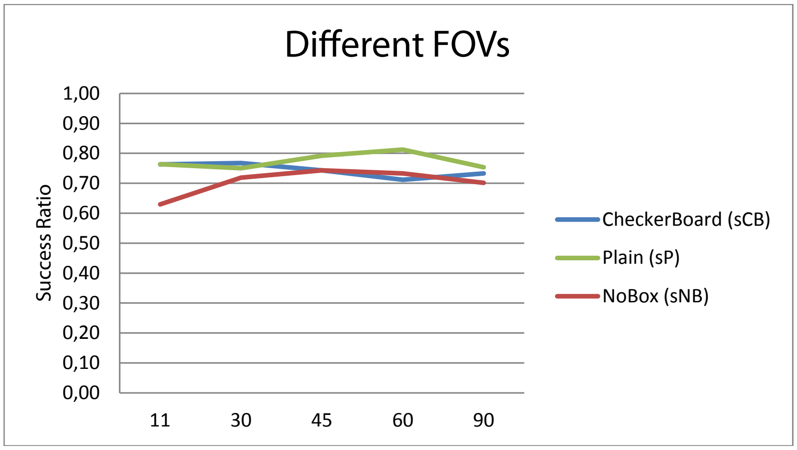


Figure 6.6: Result of part 3b of the experiment. Comparing the success rate on different values of the shearing factor.

ANOVA

Scene	Source of Variation	SS	df	MS	F	P-value	F crit
sCB	Between Groups	0,0626	4	0,0157	0,4690	0,7584	2,4320
	Within Groups	5,0060	150	0,0334			
	Total	5,0686	154				
sP	Between Groups	0,0897	4	0,0224	0,8424	0,5003	2,4320
	Within Groups	3,9923	150	0,0266			
	Total	4,0820	154				
sNB	Between Groups	0,2491	4	0,0623	1,2560	0,2899	2,4320
	Within Groups	7,4368	150	0,0496			
	Total	7,6859	154				

Table 6.4: ANOVA statical analysis of the influence of the FOV on success rate in part 3b of the experiment.

6.4.5 Influence of Position

Hypothesis 6: “It will be easier to judge order on objects close to the screen.”

The experiment also tests the influence of position on the depth perception. In figure 6.7 the success rates at objects at different positions are plotted. The chart shows the success rate on the positions of the spheres on each axis. This chart assumes a origin in the middle of the scene (that is the 8x8x4.8 box).

The chart shows no clear relationship between the position of the object and the success rate. Also ANOVA statistical analysis (shown in table 6.5) confirms the null hypothesis that all results are drawn from the same distribution. This means that hypothesis 6 has to be **rejected**.

This is surprising, especially on the z-axis. It was expected that objects on the z-axis would see better results if objects are placed farthest from the viewer. It could be that this positive effect does occur but is negated by another effect. This is the effect that if two object with the same distance between them are placed further away the distance between them seems smaller. It could also mean that the problem identified in section 6.4.3 occurs and the discreet nature of the depth test causes the behaviour observed.

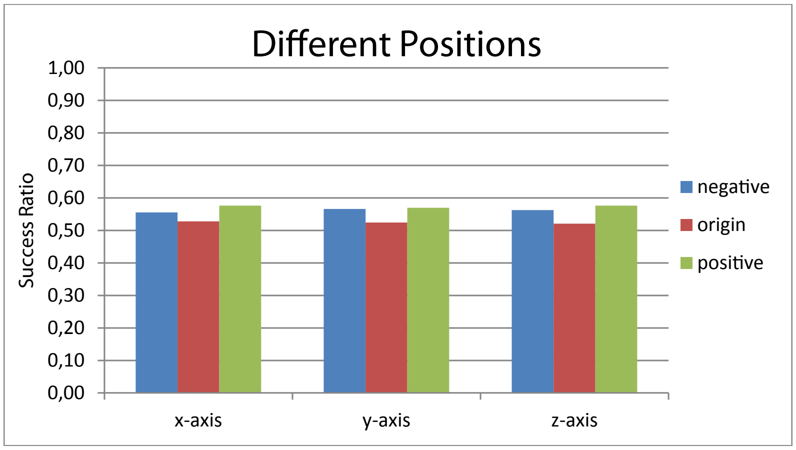


Figure 6.7: The success rate on different the positions in part 2 of the experiment.

ANOVA

Axis	Source of Variation	SS	df	MS	F	P-value	F crit
x	Between Groups	0,0381	2	0,0190	0,8106	0,4477	3,0943
	Within Groups	2,1836	93	0,0235			
	Total	2,2217	95				
y	Between Groups	0,0404	2	0,0202	0,7151	0,4918	3,0943
	Within Groups	2,6258	93	0,0282			
	Total	2,6662	95				
z	Between Groups	0,0535	2	0,0267	1,2166	0,3009	3,0943
	Within Groups	2,0448	93	0,0220			
	Total	2,0983	95				

Table 6.5: ANOVA statical analysis of the influence of the position on success rate in part 2 of the experiment.

6.5 Conclusions

The first goal of this experiment was to test how inclusion of the shoebox model in general, with realistic values for its parameters, influences the relative order depth perception of users. The experiment showed that the shoebox model significantly increased the performance of users on tasks depending on judgement of relative order in depth compared to the simple monitor view. This was expected because of the addition of motion-based depth cues by the inclusion of the shoebox model as well as the lack of other depth cues in the scene.

The second goal was to test the influence of the FOV parameter on the relative order depth perception of users. No relationship between the FOV parameter and the relative order depth perception of user was observed.

The third goal was to test the influence of the shearing factor parameter on the relative order depth perception of users. In the experiment there was no clear relationship between the shearing factor parameter and the relative order depth perception of users.

The fourth goal was to test the influence of the reference box on the relative order depth perception of users. The experiment indicated no clear relationship between the look of the reference box and the relative order depth perception of users. However the presence of a reference box does have a significant influence on the performance of users doing a task involving relative order depth perception.

The fifth goal was to test in influence of the position of objects on the relative order depth perception of users. The experiment indicated no clear relationship between the position of the object and the relative order depth perception of users.

Some expected effects did not occur. This is most likely due to of nature of the depth test. It is wise to do further experiments using a depth test of different nature that include judging the size of the distance in depth.

Chapter 7

Experiment: Line Discrimination

7.1 General Setup

7.1.1 Aim and Goals

As suggested in chapter 6, the conclusions of the experiment on relative order depth perception might be influenced by the nature of the depth test. It was suggested the influence of parameters on the performance of users on other tasks had to be investigated. In this follow up experiment a line discrimination task is chosen to fit this goal. It is chosen to test the ability of users to judge sizes of objects in located in 3d space using the shoebox model.

This experiment aims to complete the following goal:

1. Test how the FOV parameter influences the performance of users on a line discrimination task.

The influence of the shearing factor was dropped from the experiment. This was mainly done to cut the length of the experiment. The choice was made to drop the shearing factor because in both the preliminary experiment and the relative order depth perception experiment, the shearing factor provided no interestingly enough results. The FOV was kept because additional information was found important because of the large gap between the natural and realistic values for the FOV.

The comparison of the performance on the shoebox model and the monitor view was also dropped. The experiment on relative order depth perception proved that the inclusion of an additional depth cue, while there are almost no other depth cues present, significantly increase the performance of user on tasks involving depth perception. Since both the reason and presence of the increased depth perception are clear in this case, further investigation on this issue is not necessary.

7.1.2 Main Task

In this experiment users are asked to perform a task under different conditions. Participants will be presented with two lines of random orientation and length at fixed positions in the room. One line is always 1.2 times longer then the other. The participants will have to select the line that is longest. A single instance of this task will be referred to as a *line discrimination test* in the remainder of this section.

7.1.3 Dependent Variables

In this experiment the dependent variable will be the percentage of correctly judged line discrimination tests at each setting. This is used as a measurement of how well users can differentiate length at different settings.

7.1.4 Independent Variables

In this experiment the independent variable that is examined is the FOV. For the FOV values of 10°, 15°, 30°, 45° and 60°are tested. These are picked to include commonly used values as well as including realistic values and to cover the whole range of relevant FOVs.

7.1.5 Participants and Environment

The participants for this test were students participating in the "Multimodal Interaction" course at Utrecht University in 2012. Their presence was mandatory for the course. This resulted in 28 participants. All participants had a background in computer science. The participants consisted of 2 women and 26 men.

The experiments were conducted in a computer room in which several other students were working quietly, but no other distractions were present. The experiments were done with four participants at the same time if possible (each using a different device). Four devices were used in the experiment: a HTC Desire smartphone, a Samsung Galaxy S smartphone and two Motorola Milestone smartphones. This experiment was conducted at the same time as the experiment in chapter 8.

7.2 Details

7.2.1 Experiment Overview

At the beginning of the experiment a general explanation of the shoebox model is given to the participant. The participant is shown the shoebox model and can play around with it for a small amount of time. After that, the participant is given an explanation of the task he has to perform. Then, the participant is asked to answer a couple of questions (see questionnaire A.3). Next, the participant is asked to sit down, hold the device at certain distance from the head (+/- 40 cm) and begin the experiment.

The experiment exists of the following:

1. 3 test line pairs (3 line discrimination tests)
2. Tests with different FOV values (5×15 line discrimination tests)

At last, the participant is asked to answer a couple of follow up questions (see questionnaire A.3).

7.2.2 Position, Orientation and Length of Lines

The lines appear at two fixed positions in the middle of the scene. The length of the shortest line is randomly picked between 1 cm and 2 cm. The longest line is always 1.2 times longer then the other. The orientation of each line is randomly generated.

7.2.3 User Interface

During the test the participant is asked select the line segment that is longest. This done is using two buttons:

- The left button which is located on the left side of the screen and consists of the word *Left*.
- The right button which is located on the right side of the screen and consists of the word *Right*.

7.3 Hypotheses

To support the goals of the experiment, several concrete hypotheses are formulated:

1. A higher FOV will increase the users ability to compare the length of 3d line segments.

   Reason Because a higher FOV will result in bigger perspective depth in the image.

2. A higher FOV will have a positive influence on the decision time.

   Reason Because the bigger perspective depth that results from the high FOV clarifies the difference in length between the line segments.

7.4 Results

7.4.1 Influence on Performance

> Hypothesis 1: “ A higher FOV will increase the users ability to compare the length of 3d line segments.”

In the experiment the influence of FOV on the success rate is tested. It was expected that a higher FOV would lead to a better success rate. The results, which are shown in figure 7.1, however seem to reject this hypothesis. In the chart it appears that the success rate is unaffected by the chosen FOV. Statistical analysis, shown in table 7.1, confirms the null-hypothesis that the samples are drawn from the same distribution. This means that hypothesis 1 can be **rejected**.

It was expected that a higher FOV would result in better performance because a higher FOV gives more perspective depth in the picture. This was not the case. A possible explanation can be that a similar behaviour as in chapter 6 appears. The motion parallax effect caused by the shoebox gives enough information to perform the task. Since it just asked which line is longer, and not how much longer, the amount visible difference in length of the line segments does not matter, as long as the amount is visible. So this behaviour can be explained by the simple binary nature of the task, and the results might be different in a task where users are asked to quantify the difference.

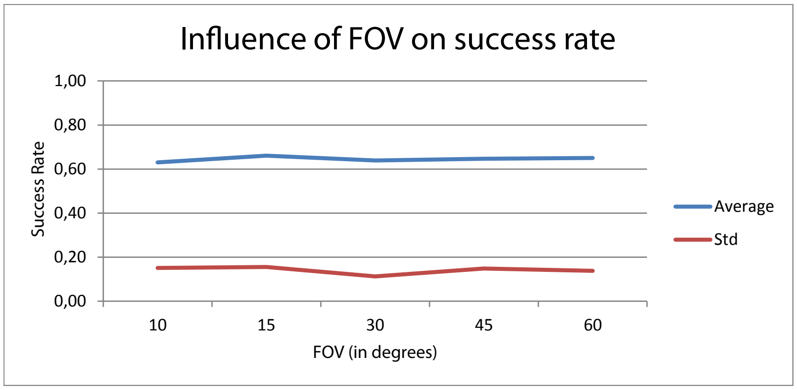


Figure 7.1: The average success rate on each FOV.

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0,0128	4	0,0032	0,1584	0,9588	2,4506
Within Groups	2,325	115	0,0202			
Total	2,338	119				

Table 7.1: ANOVA statical analysis of the influence of the FOV on the success rate.

Hypothesis 2: "A higher FOV will have a positive influence on the decision time."

It was also expected that a higher FOV would have a positive influence on decision time. Figure 7.2 show a very small increase in the decision time. So that is the opposite from what was expected. It can be noted that a higher FOV decreases the noticeability of the shoebox effect. This most likely counters the positive effect of the FOV. To see if the observed effect was significant an ANOVA statical analysis was performed (visible in table 7.2). The resulting data confirms the null-hypothesis that the samples for the different FOVs are drawn from the same distribution. So it can be concluded that FOV has no influence on the decision time of users performing a line discrimination task. Therefore hypothesis 2 can be **rejected**.

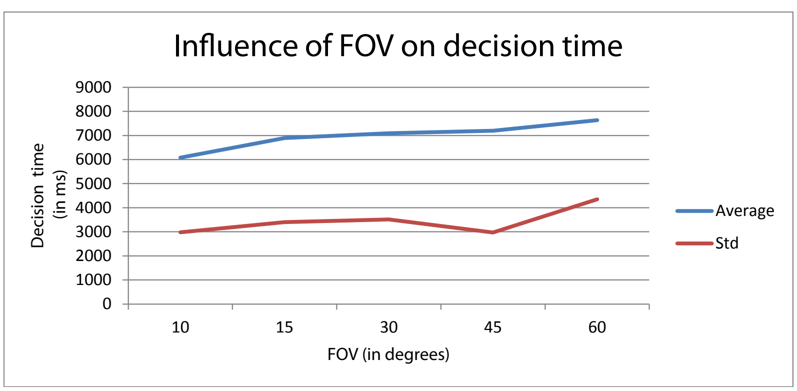


Figure 7.2: The average descision time on each FOV.

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	30074837,1	4	7518709,275	0,5986	0,66442	2,4542
Within Groups	1381671099	110	12560646,35			
Total	1411745936	114				

Table 7.2: ANOVA statical analysis of the influence of the FOV on the decision time.

7.5 Conclusion

The goal of this experiment was to test the how the FOV parameter influences the performance of users on a line discrimination task. No influence of the FOV on the success rate of users was observed. The FOV also has no influence on the time it takes users to make decisions on the line discrimination task.

The nature of the line discrimination test might be causing the behaviour. Just as the relative order depth perception task in chapter 6, this test consists of a simple binary choice. Further experiments should involve a task where users are asked to quantify a distance.

Chapter 8

Experiment: Depth Estimation

8.1 General Setup

8.1.1 Aim and Goals

In previous experiments the nature of the perception task might have been responsible for the observed behaviour. The previous tests participants were faced with a simple binary choice during a task. This experiment is designed so participants have to quantify a distance.

The experiment tries to accomplish the following goals

1. Determine the influence of the FOV on the users' ability to estimate depth.
2. Determine the influence of the FOV on the users' thoughts of realism of the effect.
3. Determine the influence of the FOV on the users' perceived performance on estimating depth.

8.1.2 Main Task

In this experiment participants are asked to perform a task under different conditions. Participants are presented with a scene containing three spheres. The sphere in the middle is always positioned in the exact centre of the scene. Participants have to judge the depth of the left and right sphere. They are given a paper containing a top view map of the scene (see figure 8.1). On it, the position of the middle sphere is marked with an 'X'. On the left and right of the middle sphere are two dotted lines marking the possible positions of the left and right sphere. The task of the participants is to mark the positions of these spheres on this paper by placing an 'X' at the correct depth on these dotted lines.

A single instance of this test will be referred to as a *estimation test* in the remainder of this section.

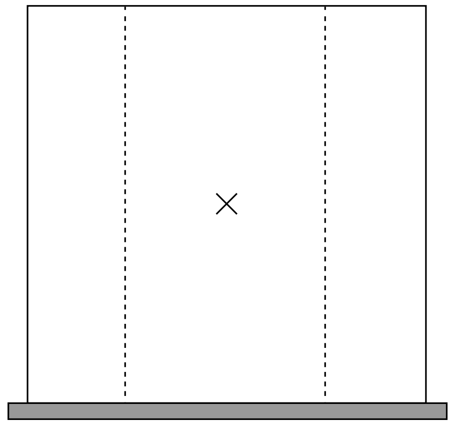


Figure 8.1: The map participants receive to mark the positions of the left and right sphere. It depicts a top view of the scene behind the device. The thick line on the bottom represents the device itself and the X marks the sphere in the middle. The dotted lines indicate the lines on which the left and the right sphere are located.

8.1.3 Dependent Variables

The most important dependent variable of the test is the average error (in mm) between the actual position of spheres and the position marked on the answer paper.

8.1.4 Independent Variables

In this experiment the independent variable is the FOV. FOVs of 10°, 15°, 30°, 45°and 60°are examined. These are picked to include commonly used values for the FOV as well as including realistic values and thus covering the whole range of relevant FOVs.

8.1.5 Participants and Environment

The participants for this test were students participating in the "Multimodal Interaction" course at Utrecht University in 2012. Their presence was mandatory for the course. This resulted in 28 participants. All participants had a background in computer science. The participants consisted of 2 women and 26 men.

The experiments were conducted in a computer room in which several other students were working quietly, but no other distractions were present. The experiments were done with four participants at the same time if possible (each on a different device). Four device were used in the experiment: a HTC Desire smartphone, a Samsung Galaxy S smartphone and two Motorola Milestone smartphones. This experiment was conducted at the same time as the experiment in chapter 7.

8.2 Details

8.2.1 Experiment Overview

At the beginning of the experiment a general explanation of the shoebox model is given to the participant. The participant is shown the shoebox model and can play around with it a bit. After that, the participant is given a explanation of the the task they have perform. Then, the participant is asked to answer a couple of questions (see questionnaire A.3). Next, the participant is asked to sit down, hold the device a certain distance from the head (+/- 40 cm) and begin the experiment.

The experiment exists of the following:

1. Testing 5 different FOV values (5 × 6 estimation tests)
   (a) 6 Estimation tests
   (b) Question: "How realistic do you think this setting looks? (-5 to 5)"
   (c) Question: "How well do you think you performed on this setting? (-5 to 5)"

The possible order of the FOV values is counterbalanced between all participants. After the experiment the participant is asked to fill the rest of the questionnaire(see questionnaire A.3).

8.2.2 Position of the Objects

The left and right spheres are positioned on imaginary lines 2 cm left and right of the centre of the scene. Each sphere can be positioned on 6 places on its corresponding line (see figure 8.2). One sphere is positioned on places 1, 2 or 3 and the other on 4, 5 and 6. The results in 18 possible couples of positions. Of this a fixed subset of 6 couples is picked to be presented at each FOV value. This subset is picked manually to represent the whole spectrum of possible couples of positions. This subset is fixed for pragmatic reasons, namely to aid the manual measurement of error. The couples of positions in the subsets are presented in 10 fixed orders. These are fixed for the same pragmatic reasons.

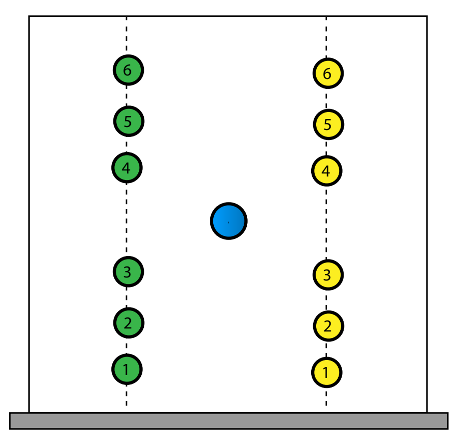


Figure 8.2: The possible positions of the left and right sphere.

8.2.3 Colour of the Spheres

To rule out any influence of colour, the spheres are randomly given one of the primary colours red, blue or yellow. The left and right sphere are given the same colour and the middle one is given another colour. This is so users can easily identify the middle one and know on which spheres they have to perform the test.

8.2.4 User Interface

Participants are presented with one button on screen. This button reads the word *Ok* and is used if a participant wants to move to the next scene.

8.3 Hypotheses

To support the goals of the experiment, several concrete hypotheses are formulated:

1. A higher FOV will increase the user's ability to estimate depth.

   Reason Because it increases the visible size of the z-axis, and therefore users can better judge distances on this axis.

2. Users will judge FOVs somewhere between the actual realistic FOV and the conventional FOV of 60°to be most realistic.

   Reason Because of the influence of conventional FOV on the users' judgement of realism.

3. Users will think they performed best at values they feel are realistic instead of their actual performance on the settings.

   Reason Because the naturalness gives the users the perception of performance.

8.4 Results

8.4.1 Influence on Performance

Hypothesis 1:"A higher FOV will increase the user's ability to estimate depth."

The most important thing tested in this experiment is the relation between FOV and the ability to estimate depth. In figure 8.3 the results of the average error (in mm) in relation to the FOV is plotted in a chart. The chart indicate a slight relation between the FOV and the average error. Larger FOV values give a smaller error value.

To see if this observed relation is significant an ANOVA statistical analysis (shown in table 8.1 was performed. The results confirm the null hypothesis that the samples for the different FOVs are drawn from the same distribution.

Therefore hypothesis 1 has to be **rejected**.

So the conclusion can be drawn that even if there is a relationship, it is very small and not significant.

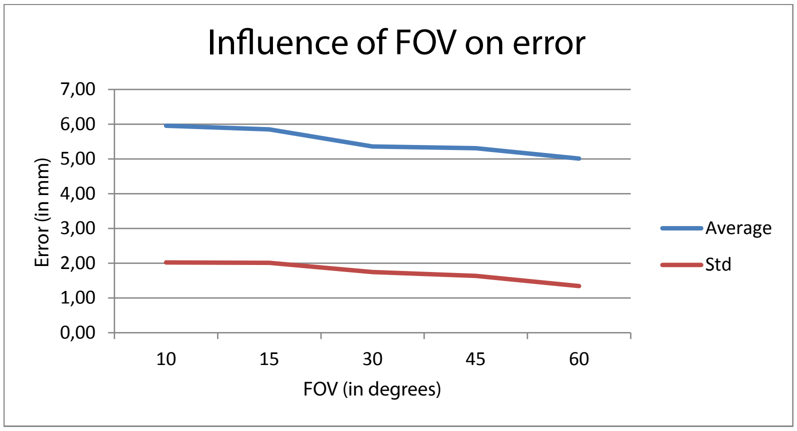


Figure 8.3: The average error on each FOV.

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	16,8969	4	4,2242	1,3503	0,2549	2,441
Within Groups	406,6879	130	3,1284			
Total	423,5848	134				

Table 8.1: ANOVA statical analysis of the influence of the FOV on the error.

8.4.2 Influence on Naturalness

> Hypothesis 2: "Users will judge FOVs somewhere between the actual realistic FOV and the conventional FOV of 60°to be most realistic."

In this test, it was asked how realistic they thought the effect looked at each FOV. Participants were asked to judge the realism of the setting with a grade between -5 and 5. It was suspected that values somewhere between the actual realistic field of view and the conventional field of view would be most natural. Figure 8.4 seems to confirm this suspicion. Values of around 30° get the best scores on realism from the participants. ANOVA statistical analysis in table 8.2 however confirms the null hypothesis that the samples at the different FOVs are drawn from the same distribution.

This mean that hypothesis 2 has to be **rejected**.

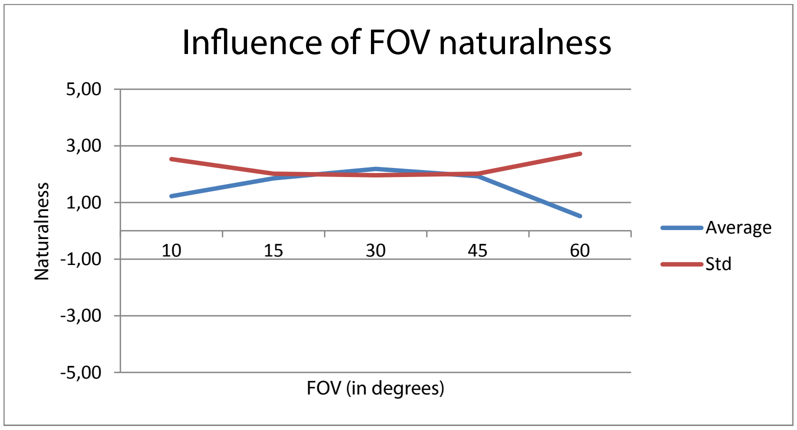


Figure 8.4: The average judged realism (naturalness) (from -5 to 5) on each FOV.

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	8,0296	4	2,0074	0,6154	0,6523	2,4414
Within Groups	424,0741	130	3,2621			
Total	432,1037	134				

Table 8.2: ANOVA statical analysis of the influence of the FOV on the naturalness.

8.4.3 Influence on Subjective Performance

> Hypothesis 3: "Users will think they performed best at values they feel are realistic instead of their actual performance on the settings."

The test also asked participant how well they think they performed at each FOV. They were asked to judge their performance with a grade between -5 and 5. It was suspected that the most natural values would yield the best results. Figure 8.5 seems to confirm this. Values of around 30°get the best scores on perceived performance from the participants, which was also the value which was judged most realistic. If users would correctly judge their performance they would give the higher FOV values the best grades.

Unfortunately ANOVA statistical analysis in table 8.3 confirms the null hypothesis that the samples at the different FOVs are drawn from the same distribution. It has to be noted that the difference between the F-value and the critical F-Value us very small, and just a small difference in values would lead the null hypothesis to be unconfirmed.

However, hypothesis 3 has to be **rejected**.

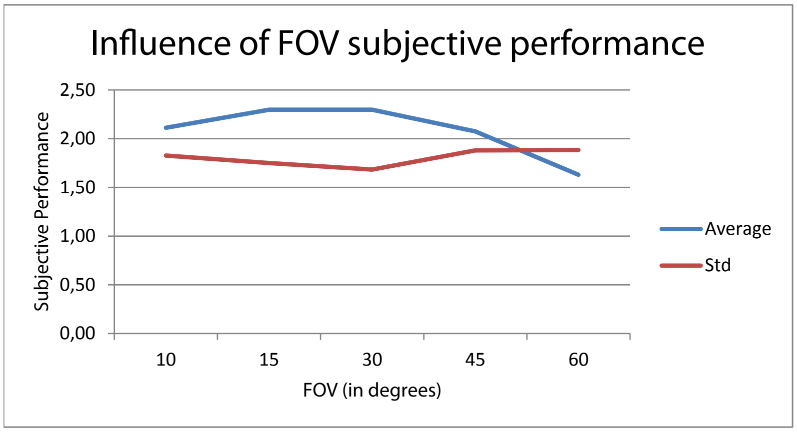


Figure 8.5: The average error on each FOV.

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	48,7851	4	12,1963	2,3638	0,0564	2,4414
Within Groups	670,7407	130	5,1595			
Total	719,5259	134				

Table 8.3: ANOVA statical analysis of the influence of the FOV on the subjective performance.

8.5 Conclusion

The first goal of this experiment was to determine the influence of the FOV on the users' ability to estimate depth. It was observed that higher FOVs gave slightly better performance. However this proved not to be significant. So even when the nature of the task was changed to a task where users had to quantify the distance, there was no significant influence of FOV on performance.

The second goal of this experiment was to determine the influence of the FOV on the users' thoughts of realism of the effect. It was observed that FOVs closer to 30° were judged as most realistic. This also proved not significant.

The last goal of the experiment was determine the influence of FOV on the users' perceived performance on estimating depth. Participants judged FOVs closer to 30° slightly better then other FOVs. However, the difference was not significant.

Chapter 9

User Feedback

During each of the experiments the users were asked to fill in a questionnaire. The chapter is a summary of some of the results of these questionnaires as well as general remarks of users.

9.1 General User Reception of the Shoebox Model

The user opinion on the shoebox model was generally positive. Some participant in the experiment even used words as 'nice' and 'cool' to describe the effect. Other participants described the effect as 'decent'. There were however some complains about a couple of areas.

Some participants noted that the implementation of the shoebox model was a bit laggy and the image contained some jitter. The jitter is caused by the noise on the signals from the accelerometers of the device. This is further enhanced by the choice of a more realistic low field of view which makes the perspective shift more noticeable. To counteract some jitter, the current reading of the accelerometer is averaged with the previous reading. While this helps to reduce the amount of jitter, it does cause some lag. One way to prevent this is to have a device equipped with inertial sensors that are less susceptible to noise. An other is solution is to have a more intelligent noise cancellation algorithm that causes less lag and better suppresses the jitter.

Participants also noted visual distortions of the objects in the scenes. The effect is most likely caused by the lack of stereoscopy, as explained in section 5.4.2. This is an inherent problem with the shoebox model.

Several participants made comments about the unintuitive direction of the perspective shift when the device is tilted. They expected it to be shifted in the exact opposite direction. The explanation for this can be found in the participants' experience with conventional 3d applications. Most conventional 3d applications have an *ego-centric* frame of reference for its viewport controls [11]. This frame of reference is for example used in first person shooters. The world is shown through an active participant in the virtual world. So if the participant rotates the device clockwise, he expects the 'camera' to rotate to clockwise. To give this illusion the the virtual world is rotated counter clockwise. However, the shoebox model uses an *exo-centric* frame of reference for its viewport controls [11]. In this frame of reference the user looks from the outside into the scene. The controls directly apply to the box. If the participant rotates the device clockwise, the world will be rotated clockwise. The mapping of the rotation of the device to the rotation of the virtual world is depicted in figure 9.1

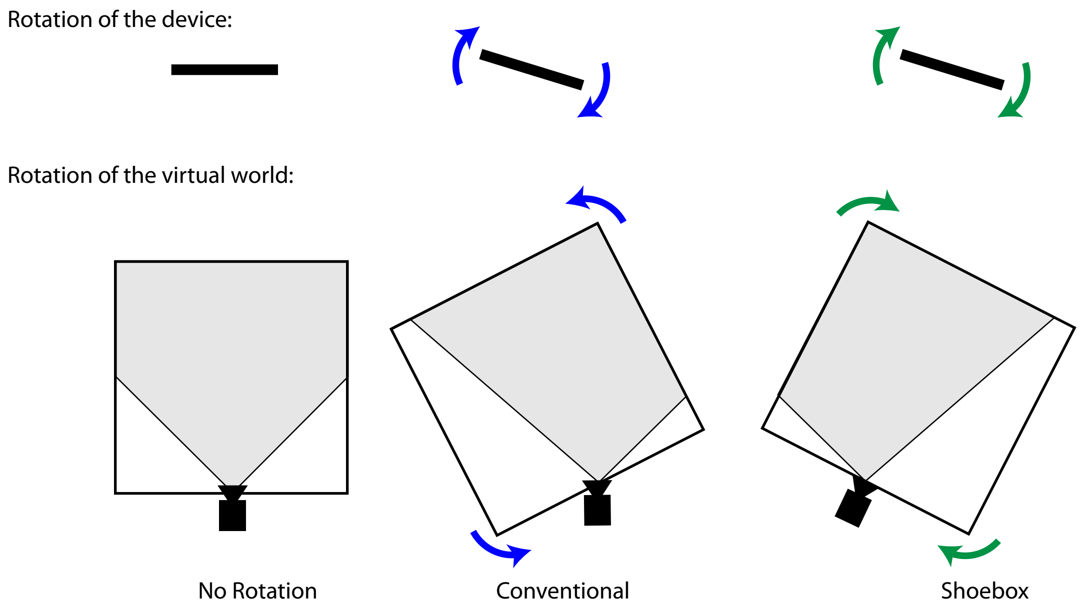


Figure 9.1: The mapping of the rotation of the device to the rotation of the virtual world in a conventional egocentric frame of reference and in the shoebox model. Note the non-corresponding orientation of the camera and the field of view in example of the shoebox model depicting the off-axis projection.

9.2 User Opinion on Proposed Application Areas

In all questionnaires participants were asked their opinion on the use of the shoebox model in games and other applications. They were asked to rate possible use of the shoebox model in each application area with a grade between 0 and 5. The results of these two questions are shown in figure 9.2.

Participants rated the possible use in games with an average of 3.4. The opinion on use in other applications was rated with a comparable average of 3.5. The general response to the use of the shoebox model in applications can be interpreted as moderately positive. It must be noted that there were some out outliers in both directions. Several participants were very positive of the use of the shoebox model. Others didn't like the shoebox model and therefore rated it's possible use in both areas with a 0 or 1.

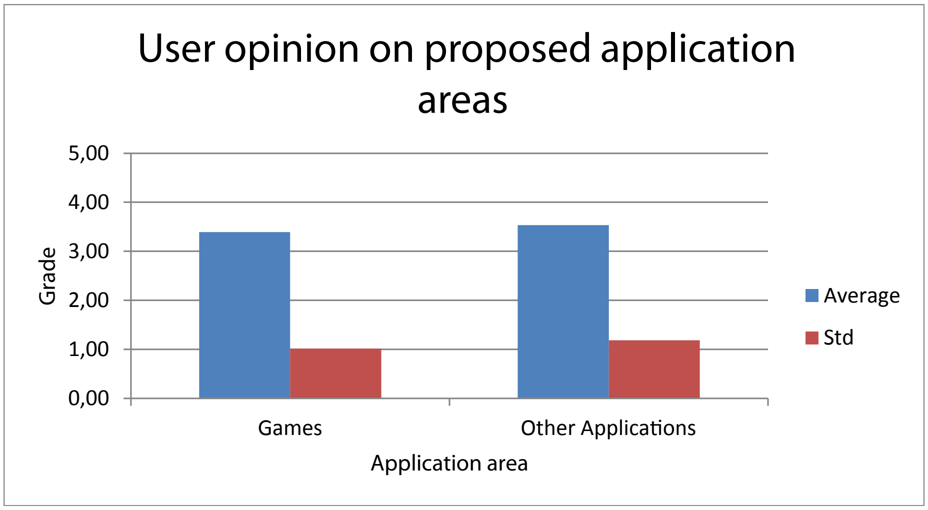


Figure 9.2: User opinion on proposed application areas rated with a grade between 0 and 5.

9.3 User Comparison to Auto Stereoscopic Screens

In questionnaire A.2 and A.3 participants who had experience with auto stereoscopic screens were asked to compare the 3d effect of auto stereoscopic screens to that of the shoebox effect. It was asked as an open question. The results are shown in figure 9.3.

Participants generally thought auto stereoscopic screens provided a better 3d effect then the shoebox model. They did noted eye sore, headaches, blurry vision, general fatigue on the brain and a limited range of allowed head positions as major drawbacks of auto stereoscopic screens. Therefore most participants preferred the shoebox visualisation. Participants typically noted that the shoebox visualisation was easier on the eyes and performed at a larger range of head positions.

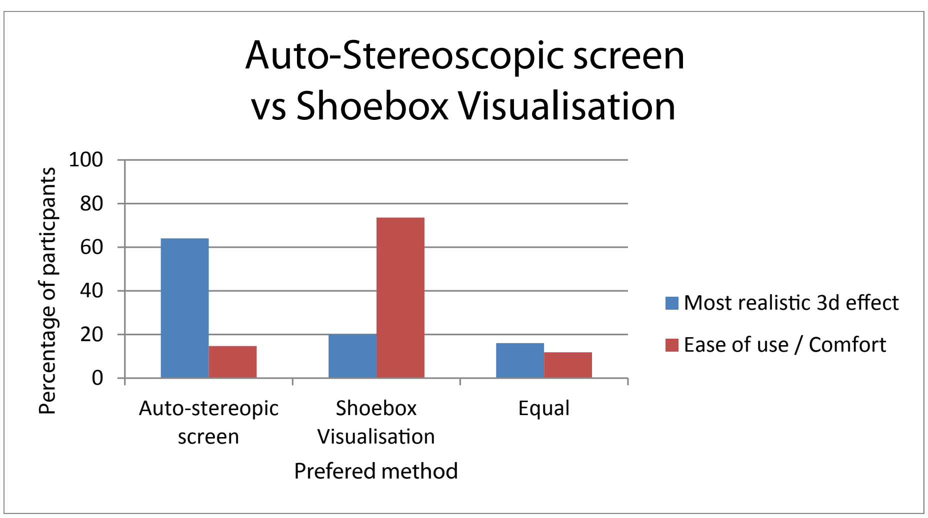


Figure 9.3: User preference of auto stereoscopic screens versus the shoebox model.

Chapter 10

Conclusions and Future Work

10.1 Conclusions and Summary

Recent years saw a big rise in the adoption of mobile devices such as smartphones and tablets. These devices allow for new 3d visualization concepts. One of these new concepts is the shoebox model, which uses perspective correction on the image on screen while tilting the device to enhance the image's 3d effect. This report researched the effect of the shoebox model, on users' 3d perception and experience.

It did so by the use of three separate user studies. All of the user studies used simplified graphics to exclude influences of the scene and rendering on the results. All user studies (except the preliminary experiment) featured a series of small tasks for the user to perform that required depth perception. Data gathered in these user studies was subjected to statistical analysis before conclusions were drawn. During the user studies each user was asked to fill out a questionnaire to gain subjective feedback on the study and the shoebox model.

> Does the shoebox model achieve a better 3d perception of users compared to the conventional 3d visualisation?

To study this, depth perception was used as an indication of 3d perception. The depth perception of users was significantly better in the shoebox model then in the conventional 3d visualisation on simple scenes that contained a low number of depth cues.

> How does 3d perception in the shoebox model depend on the involved parameters?

The two most important parameters that influence the look of the shoebox model are the shearing factor and the field of view. They are shown however to have no significant influence on the depth perception of users. The presence of a reference planes significantly increases users' performance on tasks involving depth perception. Without reference planes the effect of the shoebox model is not clear enough to users. The look and texture of the reference planes have no significant influence on the users' performance while performing tasks involving depth perception.

> Is the shoebox model beneficial to a users sense of realism?

The user reaction to the realism of the 3d effect was moderately enthusiastic. Most users described the realism of the 3d effect as 'decent'.

> How does the sense of realism of users in the shoebox model depend on the involved parameters?

The influence of the shearing factor was not researched for pragmatic reasons. The influence of the FOV on users sense of realism was not significant. However, a small preference for FOVs around 30° was observed. The lack of influence on users sense of realism is most likely caused by the experience of users with varying FOVs in existing applications.

> Is the shoebox model beneficial to an application's visual appeal?

The general response in the questionnaires to the use of the shoebox model in different applications is moderately positive. Users generally think auto stereoscopic screens provided a better 3d effect then the shoebox model. However most users prefer the shoebox model because of the drawbacks of auto stereoscopic screens, such as eye fatigue, headaches and limited range of head positions.

10.2 Discussion and Future Work

One of the reasons to do this research was to give application developers more information on how to best apply the shoebox model. From this research can be concluded that the shoebox model is a good method to give users more depth information of the 3d scene presented on the screen. However both the FOV and shearing factor parameter are shown to have no significant influence on the depth perception of users. The FOV is also shown to have no influence on users sense of realism. Application developers are therefore recommended to choose a value for these parameters that results in the most visually appealing effect or is most practical for its intended application.

However, the user studies performed in this report all made use of simplified graphics without most pictorial depth cues. Real life applications will have more depth cues present in the image on screen. So its possible that these real life applications don't benefit as much from the shoebox model in terms of depth perception as the simple scenes used in the user studies.

Another consequence of the simplified graphics is that the objects on screen don't represent real objects, just simple geometrical shapes such as spheres and cubes. In real life applications, the scenes would depict real objects. The objects can be represented in a 1:1 scale, so the virtual box behind the screen represents a box with a size in the magnitude of cm's. Another possibility is that objects on screen can represent real objects that are much bigger, so the virtual box behind the screen represents a room or even a landscape. The effect of the scale of the scene on the users' sense of realism is unknown, and could be subject to further research.

It should also be worthwhile investigating the performance on users of real life tasks involving depth perception. These real life tasks could involve finding certain patterns in medical data or exploring a three-dimensional graph.

Chapter 11

Acknowledgements

First of all the author would like to thank dr. W.O. Hürst for his help and advice. Without him this research could not have been realised. In addition the author is thankful to all participants in the user studies, especially the ones who volunteered.

Special thanks goes out to all the friends and family who supported the author in the process of the thesis and to all the students in room 684 of the BBL for providing the much needed distractions as well as the occasional advice during the work on this thesis. The author would also like to thank David Weterings and Piet de Jong for proof reading this report.

At last, the author would like to acknowledge the Department of Computer Science of the Utrecht University, for providing the necessary equipment and conditions to perform the user evaluations as well as providing a room to work in.

Glossary

counterbalanced order If the order of a number of different conditions in an experiment is counterbalanced it means that all possible orders are uniformly distributed across all participants.

device coupled perspective correction Perspective correction that is directly coupled to the orientation of the viewing device.

field of view The field of view (abbreviated FOV) is the extent of the observable world that is seen at any given moment.

fishtank model In the fishtank model concept the virtual world is attached to the device and the user looks into this world though the screen of the device. The perspective of the projection on screen is corrected to account for the user's head position using head-tracking.

fixed world concept The fixed world concept is a 3d concept in which the user views a virtual world that is aligned with the real world through the screen of their mobile device. The device acts as a viewport looking into the virtual world.

head coupled perspective correction Perspective correction that is directly coupled to the position of a user's head..

natural The value of a parameter is described as natural when users judge this value as giving the most realistic result.

off axis projection Off axis projection is a way of perspective projection in which the centre axis of the projection plane is not aligned with the camera position.

perspective correction The process correcting the perspective to account for a users viewing position.

realistic The value of a parameter is described as realistic when the resulting image is objectively closest to the image a person would see if their was an actual space behind the screen.

shearing factor The shearing factor is a factor that determines the amount of influence of the tilting of the device on the perspective correction applied to the image on the screen.

traditional concept The traditional concept is the 3d concept in which a user controls a view of the virtual world on a monitor that has a fixed position in the real world. It is also called the monitor view concept..

Appendix A

Questionnaires

During each experiment the participant were asked to fill in a questionnaire.

A.1 Questionnaire Preliminairy Experiment

- Participant nr:
- Name:
- Age:
- Sex:
- Questions asked before experiment:
  - Do you have any problem with your vision? no / glasses / color blind /problem with stereoscopy
  - How often do you play games? never/rarely/regularly/often
  - Do you play games on your mobile phone? never/rarely/regularly/often
  - Have you played any games which included this shoebox model? No/maybe/maze game/ yes
- Questions asked after the experiment:
  - Did you find the field of view hard to set to a natural value? (1-5)
  - Did you find the parameter that influenced the change in perspective hard to set to a natural value? (1-5)
  - Did you find it hard to set both parameters a once? (1-5)
  - Would you like to see more games apply the shoebox model? (1-5)
  - Would you like to see other applications apply the shoebox model (visualizing data, etc)? (1-5)
- General remarks:

A.2 Questionnaire Experiment: Relative Order Depth Perception

- Participant nr:
- Name:
- Age:
- Sex:
- Questions asked before experiment (after introduction to shoebox):
  - Do you have any problem with your vision? no / glasses / color blind / problem with stereoscopy / Other, ...
  - How often do you play 3d games? never/ <monthly / monthly / weekly / daily
  - How often do you play mobile 3d games? never/ <monthly / monthly / weekly/ daily
  - Have you played any games which included this shoebox model? No/maybe/maze game/ yes
  - Do you have any experience with stereoscopy? No/ Movies(glasses) / Screen (NDS, Mobile)
- Questions asked after the experiment
  - 2. What strategy did you use to determine which sphere was closer?
  - ....
  - ....
  - Would you like to see more games apply the shoebox model? (1 - 2 - 3 - 4 - 5)
  - Would you like to see other applications apply the shoebox model (visualizing data, etc)? (1 - 2 - 3 - 4 - 5)
  - How does this shoebox approach compare to stereoscopic displays? (Asked if experience with stereoscopy)
  - ....
  - ....
  - General remarks:
  - ....
  - ....

A.3 Questionnaire Experiment: Line Discrimination and Experiment: Position Estimation

- ParticipantNr:
- Name:
- Age:
- Sex: m / f
- Device: H / M1 / M2 / S
- 1.Do you have a problem with your vision?
  - no / glasses / color blind / problem with stereoscopy / Other, ...
- 2.How often do you play 3d games?
  - never / <monthly / monthly / weekly / daily
- 3. How often do you play 3d games on a mobile phone?
  - never / <monthly / monthly / weekly / daily
- 4. Have you played any games that included this shoebox model?
  - never / <monthly / monthly / weekly / daily
- 5. Do you have any experience with stereoscopy with glasses (3d movies)?
  - yes / no
- 6. Do you have any experience with auto stereoscopy without glasses (Nintendo 3ds, certain phones)?
  - yes / no
- 7. What strategy did you use in the estimating distances test?
  - ....
  - ....
- 8. What strategy did you use in the line discrimination test?
  - ....
  - ....
- 9. Would you like to see more games apply the shoebox model?
  - No,thank you 0 1 2 3 4 5 Yes, please
- 10. Would you like to see other applications apply the shoebox model (Data visualization, google map, etc) ?
  - No,thank you 0 1 2 3 4 5 Yes, please
- If you have experience with auto-stereoscopy:

- 11. How does this shoebox approach compare to auto-stereoscopic displays? (With regard to 3d effect, ease of use, headaches, feeling on eyes, etc)
  - ....
  - ....
- 12. General remarks:
  - ....
  - ....

Bibliography

[1] Yuliya Bababekova, Mark Rosenfield, Jennifer E. Hue, and Rae R. Huang. Font size and viewing distance of handheld smart phones. *Optometry and Vision Science, Vol. 88, No. 7*, 2011.

[2] Lynne Baillie, Harald Kunczier, and Hermann Anegg. Rolling, rotating and imagining in a virtual mobile world. In *Proceedings of the 7th international conference on Human computer interaction with mobile devices & services*, MobileHCI '05, pages 283–286, New York, NY, USA, 2005. ACM.

[3] Paul Bourke. Calculating stereo pairs. `http://paulbourke.net/miscellaneous/stereographics/stereorender/`, 1999.

[4] Tolga Capin, Kari Pulli, and Tomas Akenine-Mllere. The state of art in mobile graphics research. *Communications of the ACM, Vol.36, No.71*, 1993.

[5] Zeynep Cipiloglu, Abdullah Bulbul, and Tolga Capin. A framework for enhancing depth perception in computer graphics. *APGV '10 Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, 2010.

[6] George W. Fitzmaurice. Situated information spaces and spatialy aware palmtop computers. *Communications of the ACM, Vol.36, No.71*, 1993.

[7] George W. Fitzmaurice, Shumin Zhai, and Mark H. Chignell. Virtual reality for palmtop computers. *ACM Transactions on Information Systems, Vol.11, No 3*, 1993.

[8] Jeremie Francone and Laurence Nigay. Head tracking for desktop vr displays using the wiiremote. `http://www.youtube.com/watch?v=Jd3-eiid-Uw`, 2011.

[9] Rebecca Grootelaar. Perspective correction for 3d graphics on mobile devices. Master's thesis, Utrecht University, 2011.

[10] Matthias Helder. Vr environments on handheld devices with applications for gaming. Master's thesis, Utrecht University, 2011.

[11] Ian P. Howard. Pictorial communication in virtual and real environments (2nd ed.). chapter Spatial vision within egocentric and exocentric frames of reference, pages 338–358. Taylor & Francis, Inc., Bristol, PA, USA, 1993.

[12] Wolfgang Hürst and Tair Bilyalov. Dynamic versus static peephole navigation of vr panoramas on handheld devices. In *Proceedings of the 9th International Conference on*

Mobile and Ubiquitous Multimedia, MUM '10, pages 25:1–25:8, New York, NY, USA, 2010. ACM.

[13] Wolfgang Hürst and Matthias Helder. Mobile 3d graphics and virtual reality interaction. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*, ACE '11, pages 28:1–28:8, New York, NY, USA, 2011. ACM.

[14] Jane Hwang, Jaehoon Jung, and Gerard Jounghyun Kim. Hand-held virtual reality: A feasibility study. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 356–363. ACM Press, 2006.

[15] J. Adam Jones, J. Edward Swan, II, Gurjot Singh, Eric Kolstad, and Stephen R. Ellis. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, APGV '08, pages 9–14, New York, NY, USA, 2008. ACM.

[16] Billy Lam, Ian Stavness, Ryan Barr, and Sidney Fels. Interacting with a personal cubic 3d display. In *ACM Multimedia Technical Demonstration*, pages 959–960, October 2009.

[17] Johnny Chung Lee. Head tracking for desktop vr displays using the wiiremote. `http://www.youtube.com/watch?v=Jd3-eiid-Uw`, 2007. Carnegie Mellon University.

[18] Jack Loomis and J. Knapp. Visual perception of egocentric distance in real and virtual environments, 1999.

[19] Jonathan Davis Pfautz. Depth perception in computer graphics. Technical report, University of Cambridge, 2002.

[20] FZD School. Fov in games. `http://www.youtube.com/watch?v=blZUao2jTGAl` and `http://www.youtube.com/watch?v=S1XsPYPGcl0`, 2010.

[21] FridgeCat Software. atilt 3d labyrinth. `https://play.google.com/store/apps/details?id=com.fridgecat.android.atiltlite&hl=nll`, 2010.

[22] Ian Stavness, Billy Lam, and Sidney Fels. pcubee: a perspective-corrected handheld cubic display. In *Proceedings of the 28th international conference on Human factors in computing systems*, CHI '10, pages 1381–1390, New York, NY, USA, 2010. ACM.

[23] Başar Uğur, Ali Vahit Şahiner, and Işik Bariş Fidaner. Off-axis stereo projection and head tracking for horizontal display. *Presentations of ACE2009, International Conference on Advances in Computer Entertainment Technology*, 2009.

[24] Colin Ware, Kevin Arthur, and Kellogg S. Booth. Fish tank virtual reality. In *Proceedings of the INTERACT '93 and CHI '93 conference on Human factors in computing systems*, CHI '93, pages 37–42, New York, NY, USA, 1993. ACM.