Masterarbeit

The Use of Pursuits in Virtual Reality

Carl Oechsner
oechsner.carl@gmail.com

Zusammenfassung

Kurzzusammenfassung der Arbeit, maximal 250 Wörter.

Abstract

Short abstract of the work, maximum of 250 words.

Task

Kopie der Original-Aufgabenstellung

Motivation

Keine.

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig angefertigt, alle Zitate als solche kenntlich gemacht sowie alle benutzten Quellen und Hilfsmittel angegeben habe.

München, 7. Juni 2017

. . . .

Inhaltsverzeichnis

II

1   Introduction

2   Conceptual Framework

2.1   Eye Tracking

2.1.1   Pursuits

- Papers on Pursuits, Orbits, and where they got cited

2.2   Virtual Reality

- mention Fove - find papers in this direction - more generally: interaction in virtual reality and scenarios in which eye-tracking was used until now, its drawbacks and benefits and how vr applications could benefit from it

2.3   Eye-based interaction in VR

3   Design Space

3.1   User Bound Objects

3.2   World Bound Objects

4   Implementation

4.1   Communication with the Eye Tracker

4.2   Interpretation of Gaze Data using Pursuits

4.3   Measurement of the Field of View Wearing a Virtual Reality Headset

In order to report sizes and distances investigated in the main study, a method was needed to transform Unity units into visual degrees. In conventional eye-tracking studies that involve interaction with a screen, the person is positioned in a spot at a certain distance to the screen. Using this distance, the size and the resolution of the screen, visual angles can be calculated for the displayed objects. But due to different skull shapes the distance between the eyes and Vive lenses varies from person to person. Therefore an object with a certain size has varying values in visual degrees. At the beginning of each study session the participant was shown a Unity scene in which the individual values for the total available horizontal and vertical field of view were measured and matched to the corresponding pixel values.

First a horizontal scale is displayed that consists of 360 lines evenly distributed on an XZ-plane around the user's head in the virtual environment. Every five lines the corresponding value is displayed so that the participants can tell it to the experimenter. All scale elements are child elements of the virtual Vive camera in Unity so that the scale moves in the same manner as the user's head and its position in relation to the user's eyes remains constant. The scale is then rotated around the y-Axis so that eventually the line denoted with zero degrees meets the left edge of the screen area the user can perceive through the lenses. Now the user is asked to read the value off the scale at the very right of his field of view which constitutes the horizontal field of view the user has wearing the headset (see Fig. 4.3). In the same manner the vertical FOV is determined.

Subsequently it is determined to which pixel values the minima and maxima of the field of view correspond to. This is simply done by moving a vertical red line in front of the user until it meets the left and right edges of the available field of view and by recording the corresponding x-values in pixels. Accordingly the same procedure is performed with a horizontal line for acquiring the individual y-values. Given these six values, sizes, positions and distances can be calculated using the Law of Tangens (see Fig. 4.1) and Equation 4.2:
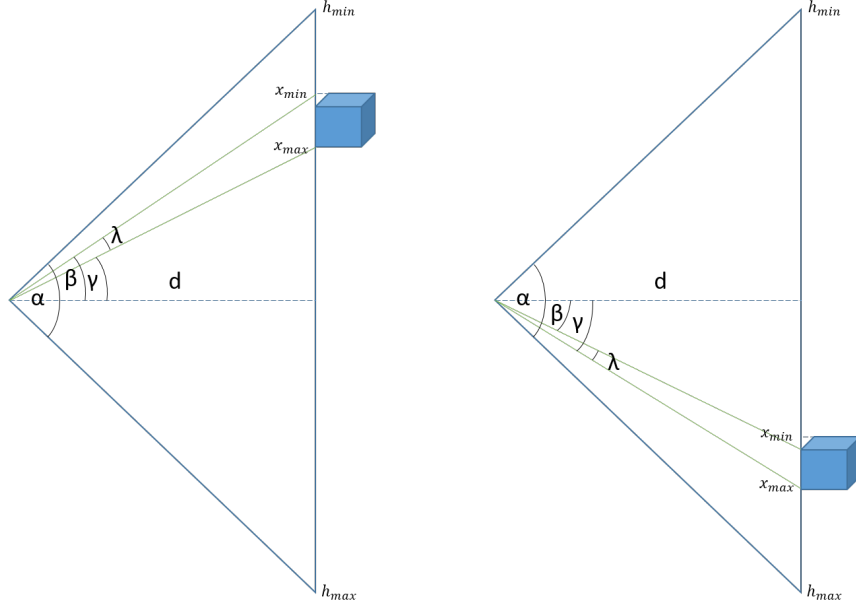
Figure 4.1: Model used for the calculation of the horizontal size of a displayed object in visual degrees. α: horizontal FOV, β: angle between the object's left edge and the horizontal center, γ: angle between the object's right edge and the horizontal center, λ: visual angle that the object occupies in the user's FOV, x_{min}: minimum x-value of the object in pixels, x_{max}: maximum x-value of the object in pixels, d: distance eye to screen, h_{min}: left edge of the FOV in pixels, h_{max}: right edge of the FOV in pixels. See 4.2

$$d = \frac{W}{\tan\left(\frac{\alpha}{2}\right)} \tag{1}$$

$$\beta = \arctan\left(\frac{\left|x_{min} - \frac{W}{2}\right|}{d}\right) \tag{2}$$

$$\gamma = \arctan\left(\frac{\left|x_{max} - \frac{W}{2}\right|}{d}\right) \tag{3}$$

$$\lambda = \begin{cases} \beta - \gamma & \text{if } x_{min} < \frac{W}{2} \wedge x_{max} < \frac{W}{2} \\ \beta & \text{if } x_{min} < \frac{W}{2} \wedge x_{max} \equiv \frac{W}{2} \\ \beta + \gamma & \text{if } x_{min} < \frac{W}{2} \wedge x_{max} > \frac{W}{2} \\ 0 & \text{if } x_{min} \equiv \frac{W}{2} \wedge x_{max} \equiv \frac{W}{2} \\ \gamma & \text{if } x_{min} \equiv \frac{W}{2} \wedge x_{max} > \frac{W}{2} \\ -\beta + \gamma & \text{if } x_{min} > \frac{W}{2} \wedge x_{max} > \frac{W}{2} \end{cases} \tag{4}$$

Figure 4.2: For better readability $W = \frac{h_{max} - h_{min}}{2}$ applies to equations 1 to 4. α: horizontal FOV, β: angle between the object's left edge and the horizontal center, γ: angle between the object's right edge and the horizontal center, λ: visual angle that the object occupies in the user's FOV, x_{min}: minimum x-value of the object in pixels, x_{max}: maximum x-value of the object in pixels, d: distance eye to screen, h_{min}: left edge of the FOV in pixels, h_{max}: right edge of the FOV in pixels.
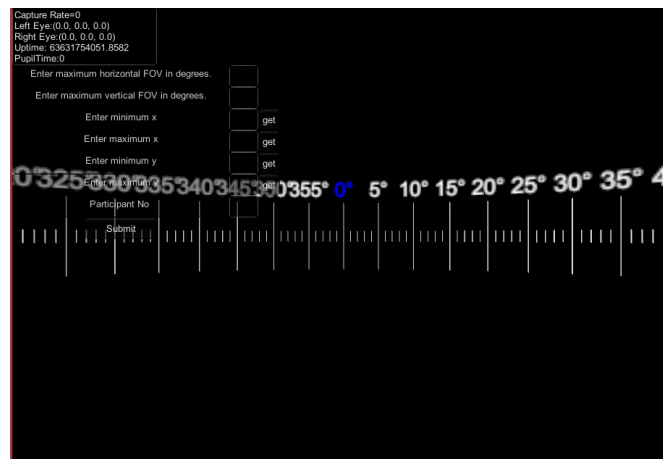
Figure 4.3: Horizontal scale that was presented to the users in order to determine their horizontal FOV

5 Evaluation

5.1 Trial Study

The system described in 4 is able to detect fixation on any user-bound object that is moving within a scene in Unity. Objects and their trajectories can have any shape, size or velocity. The trial study had several purposes: First, to identify a manageable set of independent variables that can be further investigated in the main study. Second, to exclude certain settings which don't work and there don't need to be investigated. Third, to find parameters for the correlation procedure that result in reliable results. The trial study was conducted with 14 participants. Measurements are indicated in Unity units which was changed in the main study using the procedure described in 4.3. One Unity unit corresponds to one meter in "real-world units". Objects moving on trajectories referred to as "horizontal" only alter the x component and objects on "vertical" trajectories only alter the y component of their local position respectively. In both cases the objects moved along the whole available visual range, i.e. along an trajectory with a length of 6 units.

5.1.1 Setting

The virtual objects were decided to be numbered cubes to easily tell the participants where to fixate on. All objects were located on a plane that was orthogonal to the camera's z-Axis and located at a distance of 6.78 units in front of the camera. The available visible space was determined empirically and reached from -3 to 3 units both horizontally and vertically on this very plane. When a fixation was detected by the software, an object was highlighted with a halo effect. There were several scenes created to investigate four different variables. In the following enumeration the scene is referenced with its number and the independent variable that was focussed on.

Scene 1a | *Number*: A cube with an edge length of 0.7 units was moving back and forth on a horizontal trajectory that ranged from one edge of the visible space to the other thus having a length of 6 units. To determine up to which number the system can correctly differentiate which object is selected, $\{3, 5, 7, 9\}$ cubes were shown at the same time. At each increase of the number of displayed objects, two trajectories were added at the top and the bottom of the existing ones. Each cube started moving from a random position on its trajectory. The trajectories had a vertical distance of 0.7 units to each other so that the cubes don't overlap. As the Pursuits approach would not work if all objects had the same velocity, they were all moving with unique speeds. The minimum difference between two speed values was 0.2 units per second.

Scene 1b | *Number*: Additionally, circular trajectories were investigated. $\{2, 3, 5\}$ numbered cubes were moving clockwise along a radial path on the aforementioned orthogonal plane with a radius of 2 units and a velocity of $45\frac{°}{s}$.

Scene 2a | *Position*: In this scene three objects were shown at a time at the edges of the visible area to determine if the object position has an impact on tracking performance. The objects moved along horizontal trajectories and were positioned either at the top or at the bottom of the field of view. The vertical difference between them was 1 unit with the outermost trajectory positioned at +/- FIND OUT VALUE!!. The objects started at random positions along the trajectory and had unique velocities with FIND OUT VALUES!!! differences.

Scene 2b | *Position*: Three objects on vertical trajectories were positioned at the left or at the right edges of the FOV. The trajectories always had a distance of 1 unit to each other with the outermost trajectory positioned at FIND OUT VALUE!!!. The velocity delta was 0.4 units per second to make differentiation possible.

Scene 3 | *Size*: : Cubes of three different sizes were compared to each other in terms of fixation detection. Three cubes were displayed at a time on horizontal trajectories with an edge length of {0.3,1.5} units to compare them to the standard 0.7 unit cubes. Their trajectories had a distance of one unit in the small and 2.5 units in the large case to prevent overlapping. Their velocities differed by one unit per second.

Scene 4 | *Trajectory Shape* : Three different trajectory shapes {horizontal, vertical, circular} are compared to each other within one scene each being equipped by a 0.7 unit cube. The linear trajectories had a length of 6 units and a speed of 2 units per second, the circular trajectory had a radius of 2 units and a speed of 45 degrees per second.

5.1.2   Procedure

At the beginning, each participant was shown how to put on the headset. Then, on the experimentee's head, it was adjusted in a way that it was attached safely but not too tightly to their head using the three Velcro fasteners of the Vive's harness. Being showed each scene the participants were asked to fixate on a certain object with their eyes by naming its digit which was documented in a Google Spreadsheet. The experimenter then controlled visually if the mentioned cube was being highlighted. If highlighting of an object was not constant or not given at all, detection quality was tried to improve by manipulating the values for the correlation process. HOW WAS THIS DONE? After each setting the users were asked which characteristics they found most convenient and represented best what they were looking at. The answers and other remarks were documented in a Google Spreadsheet.

5.1.3   Results and Discussion

During the course of the trial study, most reliable results were achieved with a Pearson threshold of 0.4, a correlation frequency of 3.33, a correlation time window of 300ms and a correlation averaging window of 900 milliseconds meaning that the average of the three last coefficients for an object was compared to the threshold (also see chapter 4).

All participants but one preferred circular over linear trajectories, because they were more convenient to follow. One participant mentioned that this was due to the larger distance between the objects. Furthermore, all but one person preferred smaller cubes over large cubes. Most of them agreed on them being easier to follow and to focus on. When it came to the objects' positions many agreed that differences between top and bottom (respectively left and right) were hard to tell. When they had to decide which position was more convenient, there was no discernible tendency except least users voted for the upper position. Regarding the number of objects, most experimentees preferred no more than five objects. As for the linear trajectories this was to the perceived drop in detection performance. Because there were hardly any detection errors with five objects on a circular trajectory three more cubes were shown additionally moving counterclockwise to sound out the limits of this setting. Although detection was still fairly reliable most agreed on a limit of five objects due to visual clutter.

As there was visual feedback on the selection provided the participants knew if a selection was right or wrong. If the software detected fixations in one setting better than in another this might have had an influence on their opinion.

5.2   Main Study

The main study had two primary goals: First, to find design guidelines for traceable objects in a virtual reality setting employing Pursuits. Second, by adding a task in which the user is moving within the VR area, to figure out to which amount pursuing objects is possible while performing

another task. The Correlator settings found credible in 5.1.3 were used throughout the main study. It consisted of two parts:

1. **Abstract Scenarios:** For evaluating the quality of the employed detection algorithm participants had to fixate certain objects in a low-stimulus virtual environment, once while being seated, once while moving (see 5.2.1).

2. **Use Case Scenarios:** Two scenarios were created to demonstrate and investigate the developed interaction technique in different use cases (see 5.2.2).

Before each session, the field of view of each participant was measured using the method described in section 4.3.

5.2.1   Abstract Scenario

In the abstract scenario we aimed for investigating certain variables in an isolated manner to acquire quantitative data for further analysis. There should be no biasing clues or embellishments distracting participants from their task and to make sure the observed effects are solely due to manipulations of the independent variables. For this reason, Chaperone bounds visualization was disabled and the Unity Camera Clear Flags were set to solid black, resulting all screen parts but the cubes to be coloured black. As the relative size of an object is a monocular depth clue, a script was used to ensure objects always have the same absolute size to the viewer, i.e. the object will be down-scaled when it moves towards the viewer (decreasing z-Value) and up-scaled when it moves away (increasing z-Value) [1]. Another key difference to the Use Case Scenario is that no feedback is provided if or which object is detected, which might have an impact on subjective ratings.

After each of the two scenes, participants had to answer a questionnaire Additionally every participant filled in a NASA-TLX questionnaire without pair-wise comparisons either in English or German to measure the perceived workload while performing the respective tasks [2] [3] [4]. This was done once after each of the two scenes. As neither *object velocity* and *position* had a mentionable effect and circular *trajectories* were rated both most reliable and comfortable to use in the trial study, these variables were decided to be fixed throughout the main study. Furthermore, as detection was rated most agreeable with up to five objects, this was the final choice for the number of objects per trajectory.

- **Independent variables**: Trajectory radius {0.5, 1.5, 2.5}, Object size {0.3, 0.65, 0.9}, Object Depth {2, 6, 16}, User is moving {yes, no}

- **Dependent variables**: Selection time [s], error rate {correct, not correct}

- **Fixed variables**: Object velocity {45 °/s}, position {center}, trajectory shape {circular}, Nr of objects per trajectory {5}

Static User

The seated user is shown every combination of size, radius and depth in a black surrounding in a randomized order. There are five coloured cubes of the same size moving along one circular trajectory. One cube is randomly chosen to be coloured red, the other four ones are coloured blue. Before each trial the user is told to fixate on the red cube. A cube counts as selected as soon as its coefficient is greater than the others' and exceeds the threshold. The procedure for calculating the threshold is described in section 4. For identification purposes the cubes are numbered and have their respective number as a black digit texture appearing on each of their faces. It counts as an error if a blue cube or nothing is selected. The trajectory centre is in the centre of the participant's

field of view. The objects appear and start moving when the experimenter presses the space bar. The objects disappear as soon as a cube is selected.

Moving User

In this scenario the experimentee is exposed to the same 27 combinations of radius, size and depth. To evaluate selection performance while the user is moving, the participant is told to perform a simple walking task. A white square on the floor indicates where the user is supposed to go. The square alternately appear in the middle of two opposing borders of the virtual play area. For safety purposes Chaperone bounds were enabled in this scenario. Additionally the participants were told to walk up to the bounds and to familiarize with them by feeling for the walls in the real environment. Time and error rate measurements are performed in the same way as in the static task.

In the original study design, the white square appeared in random locations within the play area. After some trials it appeared that the process of searching end re-orientating was much too demanding and resulted in feeling overwhelmed and in

5.2.2   Use Case Scenario

Besides the Abstract Scenario we wanted to examine the applicability of Pursuits in concrete use cases. There were two scenes created: one in which participants are required to be moving and one in which they are not. In both scenarios users had to fixate on objects moving along a circular trajectory to trigger certain actions in order to achieve a goal. The participant is provided input feedback on the correctness of the selected object. At the end of each setting qualitative feedback is gathered with both a NASA-TLX questionnaire and a User Experience Questionnaire (UEQ). The latter was employed to ascertain the overall impression the users received of the presented interaction technique . The UEQ is broadly used e.g. in evaluations on user satisfaction or quality assurance during software development. It consists of 26 scales in form of a seven grade semantic differential [5] [6].

Static User

The experimentee is asked to walk up to a virtual ATM machine and enter a PIN that is given to them by the experimenter. A number can be entered by looking at a cube tagged with the respective digit. The numbers one to five are moving along one circular trajectory in a clockwise manner, the numbers five to nine and zero are moving on a second trajectory in a counter-clockwise manner. There is a counter for every cube that is increased whenever the software detects a fixation (just like in the abstract setting). After a time limit $t_s \in \{0.3, 0.5, 0.7, 0.9\}$ TO BE DETERMINED!! the cube with the highest counter is considered as input. Every time an input is accepted, the user gets visual feedback in form of a temporal highlighting of all cubes so it is clear that they can move on to the next number.

Moving User

For the moving user task we decided on a game setting in which the user is asked to "shoot" at meteors that move into the field of view. Like in the abstract case the target object is coloured red while the other objects are coloured blue. There is only one target object visible at a time to clearly detect wrong selections. The positions and shapes of the object trajectories are randomly chosen by the software.

14

5.2.3   Objective Results

Statistics from the measured Data of the Abstract Scenario

5.2.4   Subjective Results

Results from NASA TLX, UEQ and additional questions of the participant form.

Abstract Scenario		Use Case Scenario	
Stationary	Walking	Stationary	Walking
21.08	36.38	47.13	47.27

Table 5.1: Averages of the Overall Rating Score measured for all participants after the respective Scenario

5.2.5   Discussion

5.3   Follow-Up Study

To differentiate between the effects of radius and velocity a Follow-Up study was conducted. Here its results and the assumptions drawn from it. Of course, due to the small sample size, further investigations have to be carried out on this matter to prove the assumptions.

5.3.1   Results and Discussion

6 Conclusion

7   Limitations and Future Work

Limitations:

Quality of Tracking Data relies on the eyes being in focus. As the cameras' focus can only be adjusted manually the headset has to be put on and off the head every time an adjustment has to be made. Due to the short distance it happened that even if pupils were in focus when looking straight, they might get blurry if the eyes move to fixate an object at the edge of the field of view. This might have had an influence on the tracking performance measured for large radii.

Measuring the field of view relied on the participants interpreting the presented scale correctly

It lies in the nature of gaze data that it is calibrated and fits only to one user perfectly. The presented system also works with other users using the same calibration but performance might vary between users.

DVD Contents

Literature

[1]   E. Bruce Goldstein. *Sensation and Perception*. Ed. by Mariane Tafliner et al. 6th ed. Pacific Grove, CA, USA: Wadsworth, 2002, pp. 227–233. ISBN: 0534539645.

[2]   S. G. Hart. "Nasa-Task Load Index (NASA-TLX); 20 Years Later". In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50.9 (2006), pp. 904–908. ISSN: 1071-1813. DOI: 10.1177/154193120605000909. arXiv: 9605103 [cs]. URL: `http://pro.sagepub.com/lookup/doi/10.1177/154193120605000909`.

[5]   B Laugwitz et al. "Subjektive Benutzerzufriedenheit quantitativ erfassen: Erfahrungen mit dem User Experience Questionnaire UEQ". In: *Usability Professionals* (2009), pp. 220–225. URL: `https://www.researchgate.net/profile/Martin%7B%5C_%7DSchrepp/publication/236462743%7B%5C_%7DSubjektive%7B%5C_%7DBenutzerzufriedenheit%7B%5C_%7Dquantitativ%7B%5C_%7Derfassen%7B%5C_%7DErfahrungen%7B%5C_%7Dmit%7B%5C_%7Ddem%7B%5C_%7DUser%7B%5C_%7DExperience%7B%5C_%7DQuestionnaire%7B%5C_%7DUEQ/links/00b7d52457033316ab000000.pdf`.

Web-References

[3]   Jens Grubert. *NASA Task Load Index German Short*. URL: http://jensgrubert.bplaced.net/nasa-tlx-short/TLX-German-short.html (visited on May 29, 2017).

[4]   Jens Grubert. *NASA Task Load Index English Short*. URL: http://jensgrubert.bplaced.net/nasa-tlx-short/TLX-English-short.html (visited on May 29, 2017).

[6]   Andreas Hinderks, Martin Schrepp, and Jörg Thomaschewski. *UEQ - User Experience Questionnaire*. URL: http://www.ueq-online.org (visited on May 29, 2017).