

# 3D HMD Image Browsers: Optimal User Placement and Exploration Mechanisms

K. van Nieuwenhuizen\*  
Utrecht University

G.J. Schaap†  
Utrecht University

W.Hürst‡  
Utrecht University

## ABSTRACT

A lot of different image browsers have been proposed over the past few years. However, most of them are displayed on non-stereoscopic 2D displays. Image browsers that are displayed in three dimensions in a virtual environment can be created due to the current state of head mounted displays. Because of these developments, questions arise on how to use these HMDs in the best possible way to improve image browsing. The focus in this paper is on different options for user placement and exploration mechanics in such browsers.

Three different cylindrical image browser interfaces that differ in terms of placement (inside vs. outside) and exploration (scrolling and head tracking vs. head tracking) are introduced. How these different variables influence performance and experience is measured using known-item search tasks and a questionnaire respectively.

Results show that it is better to place the user in the center of the cylinder to maximize user experience. It is also better to use both the head tracking ability as well as scrolling for exploring to maximize the user experience while minimizing physical strain on the human body. In terms of performance it does not matter if the user is placed inside or outside the cylinder, nor does it matter if scrolling is included or not.

**Index Terms:** H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, Augmented and Virtual Realities; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Evaluation/Methodology;

## 1 INTRODUCTION

Nowadays it is extremely simple for people to capture a wide variety of images of their surroundings because of devices like mobile cell phones and tablets. Image browsers can be used to display a multiple of these images in such a way that it is efficient to look at them and find them, while providing an enjoyable experience for the user. A wide variety of image browsers have been proposed[17], but these are mostly displayed on a non-stereoscopic 2D display.

However, the idea of interfaces displayed in three dimensions in the real world have been around for quite some time. Movies like *Star Wars*[8] and *Total Recall*[9] use holographic interfaces to display information. Recent movies like *Pacific Rim*[10] and *Edge of Tomorrow*[11] have portrayed these holographic interfaces even more realistically. Thus the line between fiction and non-fiction is getting thinner and thinner.

\*e-mail: kevinvannieuwenhuizen@gmail.com

†e-mail:Johnnyschaap1@gmail.com

‡e-mail:huerst@uu.nl

These holographic interfaces can be used for image browsing and can be mimicked within a virtual environment by using a head mounted display (HMD), a device worn on the head that has a single or two small displays with lenses embedded in it[20]. This device tracks the head movement of the user and updates the screen according to the change in head orientation. These aspects gives the illusion that he is 'inside' a virtual environment because it directly connects to how we look around in the real world. Thus, it potentially gives a better user experience in terms of enjoyability and of having the sense of presence in the virtual world.

HMDs have been around for quite some time, however they were not readily available to the public[26]. With the developments in the field of HMDs[20], they will be accessible to any person in the near future, making it possible to use them for applications in everyday life.

An example of a HMD is the Oculus Rift[14], a relatively in-expensive[13], high quality HMD which currently can be purchased as a developer kit. Other alternatives exists on the market[6][24]. However, they need additional hardware which takes away from the simplicity the Oculus Rift has.

HMDs can be used as an interaction device in image browsers. Placing a three dimensional interface and the user in a virtual environment creates the feeling that the interface is physically surrounding the user (figure 1). This creates the same feeling as the holographic interfaces portrayed in movies and thus gives a potentially better experience in terms of enjoyment with respect to non-stereoscopic 2D image browsers. This concept of creating an image browser inside a virtual environment using a HMD will be referred to as *HMD 3D image browsers*.

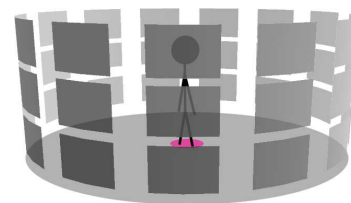


Figure 1: The user and the interface of the image browser are placed inside the virtual environment in such a way that it seems if the interface is physically surrounding of the user.

A common problem of image browsers is that they have to display a lot of images. It can happen that not every image will be visible on the display at the same time. An efficient solution needs to be found to make sure the user can view all the images. One way to solve this problem is to make use of scrolling, sliding the content of the image browser across the display horizontally or vertically. The user his view remains static, while the images move in front him.

If a HMD is used as an interaction device, another solution is to take advantage of the fact that the user can look around by using the ability to track head movement. This effectively creates a larger area which he can see within the virtual world and thus a larger area to position and display images.

The suggested solutions have to perform well in terms of the time it takes to find specific images (known-item search) and minimize the amount of errors the user makes while trying to find them while maintaining an enjoyable experience. The solutions also have to perform well when the user just wants to browse through the images. Other options exist to solve the problem of limited space. However, the suggested options are the most straight forward solutions.

It is not clear what interface shape is the best, nor is it clear where to place the user with respect to the interface in such a way that the performance and user experience are optimized. A 2D non-stereoscopic image browser is restricted in the way the interface is shaped and placed, it has to be visible and operative with respect to the non-stereoscopic 2D display. A HMD 3D image browser has much less restrictions in this respect. Using the ability to track the head movement of the user, a significant portion of the virtual environment can be utilized for an interface.

This paper introduces interfaces with the shape of a cylinder because it seems to be the most logical choice; Every position in one row on the shell of the cylinder has the same distance to the center. When placing the user in the center and the images on the shell, there will be no bias with regard to the positions of the images. Every image has a distance to the center based on the radius of the cylinder and the difference of the height of the user and the row. In other words, no image will be given preference based on position (figure 2), and in turn no image will be given bias when the user looks for a specific image in the HMD 3D image browser. Placing the user in the center of the cylinder can potentially be distracting, but can also be better for the user experience in terms of feeling present in the virtual environment due to the fact that he has a lot of information surrounding him.

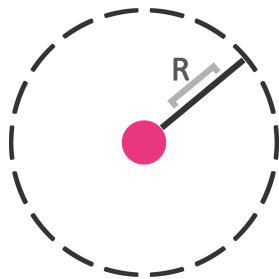


Figure 2: A top view of a cylinder. Every image placed on the shell of the cylinder has distance  $R$  to the center.

Another option is to place the user outside the cylinder at a distance of  $2 * CylinderRadius$  from the center, ensuring that the distance of the images on the cylinder directly in front of him is based on the radius of the cylinder and the difference in height between the user and the row. The area in front of the user is also directly in the perceptual span of the human eye[18] and is the most important part of the cylinder when looking at, or finding, specific images. This is the same as when he is placed in the center of the cylinder (figure 3). Also, when the user is placed outside the cylinder, he has less information surrounding him and thus it is

potentially better in terms of performance, while lowering the user experience.

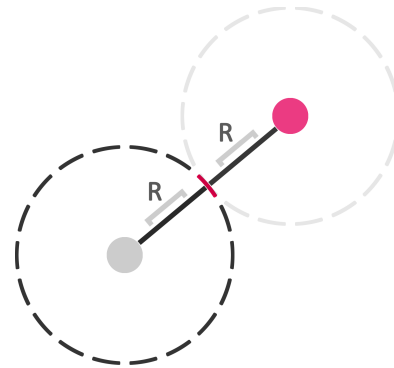


Figure 3: A top view of a cylinder. The distance from the center of the cylinder to the image on the shell (colored in red) is the same as from the point outside the cylinder.

A cylindrical interface is also suitable for implementing a scrolling mechanism. Rotating the cylinder will still respect the arguments mentioned in this section to use a cylindrical shape as an interface.

In this paper three cylindrical HMD 3D image browsers are presented and related aspects regarding the user experience as well as the performance are researched. To be more precise, this paper gives an answer to the following research questions;

*"What is the most efficient way to implement an exploration mechanism for HMD 3D image browsers using a cylindrical interface, in terms of performance and user experience."*

*"Should the user in the virtual environment be placed inside or outside a cylindrical interface to optimize user experience and performance."*

Section 2 contains work related to this research field. Section 3 will give an in-depth explanation about the three cylindrical HMD 3D image browsers as well as give a more in depth explanation of the used hardware. Section 4 and 5 introduces the experiment as well as the results and section 6 contains the conclusion and future work.

## 2 RELATED WORK

HMD 3D image browsers are a relatively scarce topic. However, a lot of research has been done on non-stereoscopic 2D image browsers. The research in this section are all implemented on a non-stereoscopic 2D display unless noted otherwise.

2D, 2.5D and 3D interfaces A. Cockburn and B. McKenzie[3][4] researched the difference in performance between virtual and physical interfaces as well as difference in performance between 2D, 2.5D and 3D interfaces. They created a physical and virtual interface for each of these dimensions and did an experiment to find the difference. They concluded that the higher the dimension, the more the user would move his body. They also concluded that physical interfaces were slightly (but not significantly) faster than their counterpart. The same could be said when comparing the 2D and 2.5D interfaces against 3D interfaces. It is important to note that even though the recorded results suggested that the 3D interfaces performed worse, it felt more natural and the users felt like

they were faster using the 3D interface. The same feeling can potentially be applied to HMD 3D image browsers without lowering performance.

**Shape Comparison** Klaus Schoeffmann and David Ahlström[12] researched which shape, with images displayed on it, would lead to the best performance. They first identified the problems of 2D interfaces, including the lack of possibility to show a lot of images at the same time on the display and the way users, when searching for an image, tend to scroll back to the top instead of continuing from the end after they reached the end of the dataset without finding the image. To solve these problems they developed 3D cylinder and globe interfaces. Their experiment showed that the 3D interfaces were indeed better in terms of performance. P. Cubaud et al.[16] also used cylinders to display book covers on. This is used as some sort of library that the user can use to find a certain book.

**Scrolling** S. Mehra[22] researched the difference between spatial and spatial-temporal memory by comparing dynamic versus static peephole navigation. Dynamic peephole navigation performed better. S. Boring et al.[21] compared scrolling (spatial-temporal) and tilting (spatial) on a mobile phone against each other and concluded that tilting had better performance in terms of speed. Even though A. Cockburn[2] suggests that there is no benefit in using spatial memory in 3D interfaces on a non-stereoscopic 2D display, it is not concluded that this is the case when using the head tracking of HMDs. Thus it is potentially better than using standard scrolling methods when implementing HMD 3D image browsers.

### 3 IMPLEMENTATION

The HMD used for this project is called the *Oculus Rift DK1*[15]. It is worth to mention that the Oculus Rift has a resolution of  $1280 \times 720$ [15]. Although a higher resolution will likely have a positive impact on the user experience for image browsing, we do not expect any negative influence on the results because our HMD 3D image browsers are designed in such a way that every image seen on the display of the Oculus Rift can be differentiated from each other. Thus, the results of the experiment when performing known-item search are not influenced by the low resolution.

The 3D HMD image browsers are implemented using Unity3D[28]. Unity3D provides an empty 3D environment where 1 unit equals 1 meter. This is the basis of our implementation.

Input is done using a *Xbox 360 Controller* (figure 4). We chose this device because it is the most simple and straight forward to use. Other options exist, but considering all of them is outside of the scope of this paper.



Figure 4: Illustration of the Xbox 360 controller.

#### 3.1 Core Implementation

The HMD 3D image browsers are using a cylinder as a medium to show the images. The images are placed on the shell of the cylinder (section 1). The placement of this cylinder in the virtual environment is arbitrarily chosen. For this experiment, the center of the cylinder is placed on the origin of the virtual environment.

The cylindrical interfaces use a radius of 3 units and a length of 3.25 units. The length and the radius are chosen in such a way that the maximum amount of images can be viewed through the HMD while the user can still differentiate them. These values are the same for all three HMD 3D image browsers for consistency and are chosen based on the resolution of the HMD.



Figure 5: Illustration of how an image is presented in the image browser. It is surrounded by a border in order to make all images as uniform as possible.

Each image is resized such that it fits within a  $5.5 \times 5.5$  unit frame and is surrounded by a square border (figure 5). The images are placed on the shell of the cylinder in such a way that there are 3 rows and 24 columns of images. There are a total of 72 images showed on the cylinder. The distance between each image is 2.5 units.

The placement of these images on the shell of the cylinder is determined by sorting them based on the HSV color model[25]. This color model consists of three components; *Hue*, *Saturation* and *Value*. The first component is used to sort the images horizontally, the third component is used to sort the images vertically. A speedup of about 20% can be achieved when compared to unsorted images when performing a known-item search in a grid type interface[19][23].

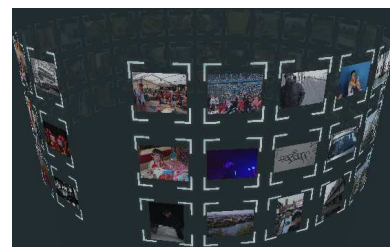


Figure 6: Illustration denoting the gap in the cylinder. This gap can be used as a reference point for the user when finding images.

The cylinder has a single column without images to introduce discontinuity in the shape (figure 6). This is done to make sure that the user has a reference point in the image browser. This reference point can be used as an anchor to remember the placement of images when finding certain images in the HMD 3D image browser.

When the user is looking at an image in the interface, it is highlighted by enlarging it as well as shifting it slightly towards him, making it clear to the user which image is currently being looked at (figure 7). Unlike other solutions (like the mouse cursor on computers), this hovering mechanism does not generate any graphical distractions.



Figure 7: (Left) - Image when not being looked at. (Right) - Image when being looked at. It is slightly enlarged as well as shifted slightly towards the user.

An option is given to the user to zoom in on a selected image. This is done by multiplying the original size with a factor of 3. When the user hovers over an image, it can be selected and enlarged by using the *A* button. Pressing this button again will deselect the image and shrinks it to its original size.

Our HMD 3D image browsers are designed in such a way that it does not matter what type of image database is used to show images. The implementation used in this experiment uses the *Yahoo! Webscope Yahoo Flickr Creative Commons 100M (YFCC-100M) dataset* [29]. This dataset consists of 100 million images and videos. Images are downloaded and sorted on the fly. Because they are downloaded in real time, loading times may differ depending on the internet connection speed.

### 3.2 Inside View without Scrolling

The first interface places the user in the center of the cylinder as seen in figure 2 and will be referred to as *IN*. The view of the user can be seen in figure 8. When performing known-item search with this version, spatial memory[1] plays an important role due to the fact that every image has a position that does not change.



Figure 8: The view of the user in the image browser when he is placed in the center of the cylinder.

### 3.3 Inside View with Scrolling

The second interface is identical to *IN* except for the fact that this version allows the user to rotate the cylinder around the axis through the center (figure 9) by using the *Left Trigger* button to rotate to the left and the *Right Trigger* button to rotate to the right. This interface will be referred to as *IS*.

### 3.4 Outside View with Scrolling

The third interface is identical to *IS* in every way except for the placement of the user. This version places the user outside of the cylinder with a distance of  $2 * CylinderRadius$  from the center of

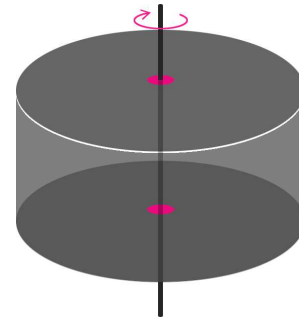


Figure 9: The axis of rotation of the cylinder.

the cylinder as can be seen in figure 3 and will be referred to as *OS*. The view of the user can be seen in figure 10. The user has to use the scrolling mechanism in order to view every image on the cylinder because of occlusion.



Figure 10: Image browser where the user is placed outside the cylinder at a distance of  $2 * CylinderRadius$  from the center.

## 4 EXPERIMENT

The HMD 3D image browsers in section 3.2 and 3.3 are used to determine the best way to display an amount of images that are not all viewable at once for the user. The only variable between these two image browsers is the ability to scroll. The comparison on the influence of this variable will be referred to as the *scrolling experiment*. The HMD 3D image browsers in section 3.3 and 3.4 are used to determine which placement of the user with respect to the interface maximizes the user experience and performance. This experiment will be referred to as the *placement experiment*. The only variable between these two image browsers is the placement of the user.

Both experiments will be validated using the time it takes to find the image as well the amount of errors the user made in the known-item search as a measure for performance. The user experience will be measured using a questionnaire.

**Set-Up** 19 males and 3 females participated in the experiment. This group had an age ranging from 18 to 57 and used contemporary image browsers and game controllers at least once a week. They had no experience with the Oculus Rift. The average duration of the experiment was 45 minutes. The participants had to sit in a stationary chair and were given the controller. An explanation of the experiment was then given to the participant. This included

telling the participants that they had to test the three different interfaces in sequence by performing a known-item search and that they had to fill in a questionnaire afterwards. Before testing each interface, the participants were told the name of it, how it looked and how the interaction with the controller worked. This included telling the user that the *Left Trigger* and *Right Trigger* were used to scroll to the left and right respectively (only applicable to *IS* and *OS*), the *A* button was used to select images and the *Y* button was used to confirm if the participant had found the image they were given at the beginning of a test round for the known-item search. The known-item search was repeated 10 times for each of the three image browsers to reduce the influence of outliers. The order in which these three image browsers were tested were chosen in such a way that every possible permutation happened at least 3 times and at most 4 times given the number of participants. Thus, each of the participants had another sequence of image browsers to test. Choosing a semi-random sequence of image browsers for each participant was done because participants could get used to the hardware and perform better on the third image browser than on the first. The randomness was used to avoid habituation of the hardware influencing the results.

**Testing** The testing of each of the three image browsers was done as follows; 72 random images were loaded in the image browser from the Flickr database. Due to certain hardware and software restrictions, only a tenth of the database was used for this experimentation called  $\alpha$ . This was chosen arbitrarily. Performing known-item searches was done on 72 images randomly chosen from  $\alpha$ , making sure that certain (dis)advantages for known-item searches in the chosen datasets were averaged out between all participants. These images stayed the same throughout the testing of the same image browser.

The participant had to put on the HMD and was then given the chance to explore the interface. He had a maximum of three minutes to explore the image browser and to get familiar with the shown images. This phase could be ended early if the participant did not need the full 3 minutes.

After the exploration phase, the participant was shown a randomly chosen image chosen from the 72 already viewed in the previous phase. He was told to memorize this image and to press the *A* button of the Xbox 360 controller when he was done memorizing. The participant had an unlimited amount of time to do this because each person differs in terms of how long it takes to memorize images. After pressing the *A* button, the image disappeared and the interface of the image browser the participant familiarized with in the exploration phase was shown. The task was to find the image as fast as possible, while minimizing the amount of wrongly chosen images. If the participant found the image, he had to select the image by pressing the *A* button to select and verify that this was his choice by pressing the *Y* button. As soon as the participant found the right image and verified his choice, feedback was given that the round was over. If the participant did not find the image within a time frame of three minutes, the application automatically ended the current round. The known-item search was repeated 10 times, each round showing a new random image chosen from the 72 images in the image browser.

After 10 rounds of each interface, information collected in the application was written to a file. For each round, this file included the time it took the participant to find and select the correct image (with the *A* button), the time it took the participant to verify that he had chosen the right image (with the *Y* button) and the amount of false verifications the user did (pressing the *Y* button on the wrong image). This information was used to calculate the performance of the known-item search task.

The amount of rotation to the left and the right by using the *Left Trigger* and *Right Trigger* buttons, as well as the amount of rotation to the left and right using the head tracking were recorded. This information was used to determine which type of exploration mechanism was preferred more when the participant had the option to use both (applicable to the image browsers specified in section 3.3 and 3.4).

After testing the three HMD 3D image browsers, the participant had to fill in a questionnaire. Answers had to be given based on a five-point Likert Scale.

## 5 RESULTS

This section will summarize the results gathered from the experiment as explained in section 4. The results gathered for the *scrolling experiment* will be summarized in subsection 5.2 and the results gathered for the *placement experiment* will be summarized in subsection 5.3. Two-Tailed T-Tests with an  $\alpha$  value of 0.05 were used to draw conclusions between the various variables gathered from the experiment.

### 5.1 Shared Data

On average, the participants were really enthusiastic about the application (rated 4 out of 5) and they would use HMD 3D image browsers more if they had the chance (rated 3 out of 5). However, they also said that these kind of applications, in the state they experienced them, were not better nor worse than contemporary image browsers (rated 3 out of 5 on average). The low resolution of the Oculus Rift (section 3) did not affect the participants (rated 3.2 out of 5). Even though the device has a low resolution, creating an application with this in mind makes sure that any negative effects caused by the low resolution are minimized.

There is no clear preference between the three interfaces in general. However, there is a difference in terms of the experience. Participants rated *IS* and *OS* higher than *IN* ( $p = 0.032$  and  $p = 0.036$  respectively). This could be due to the fact that *IN* requires to turn the whole body and thus produces physical stress. They preferred *IS* the most. This is possibly due to the fact that being placed in the center of the cylinder benefits the user experience, while being able to scroll using the controller reduces the physical stress and increases how easy it is to view all images.

Participants did not think they would make more errors in one set-up than the other. In terms of speed, participants felt that *OS* led to the least errors, while *IN* led to the most errors. There was no significant difference between *IN* and *IS* ( $p = 0.076$ ) and between *IS* and *OS* ( $p = 0.107$ ). However, there is no actual difference in terms of speed between the three interfaces.

### 5.2 Scrolling Experiment

**Performance** The average time it took the participants to complete the known-item search each round for *IN* and *IS* can be seen in figure 11. Even though *IS* is the fastest, there is no significant difference between that interface and *IN* ( $p = 0.122$ ). In terms of speed, it is enough to use the head tracking ability as an exploring mechanism when using a cylindrical interface where the user is placed in the center of the cylinder.

A minimal amount of errors were made in the known-item search by the participants as can be seen in figure 12. This can be due to the fact that the participants had to select and enlarge the image first with the *A* button before confirming their choice with the *Y* button. Having an extra (enlarged) look at the image, and thus extra time to think about their choice, may have helped to reduce



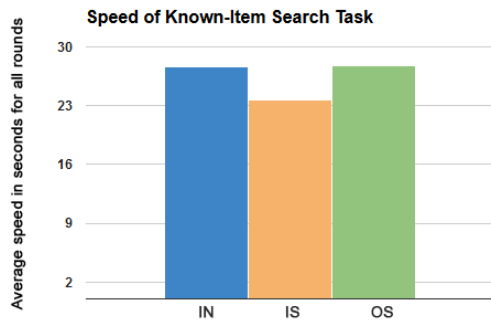


Figure 11: The average speed in seconds for all rounds to perform the known-item search task (pressing the Y button on the controller when the correct image was selected).

the amount of errors the participants made. There is no significant difference between the two different interfaces ( $p = 0.511$ ). This means that the participants did not make more errors in a particular interface and thus the amount of errors is not influenced by the different exploring methods used in *IN* and *IS*.

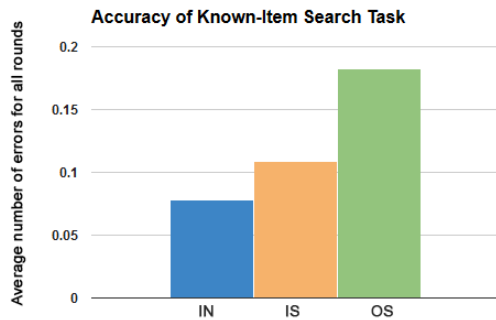


Figure 12: The average number of errors the participants made for all rounds (pressing the Y button when a wrong image was selected).

**User Experience** Participants were asked if they felt like they were really surrounded by the pictures. There is no significant difference between *IN* and *IS* ( $p = 0.135$ ). They were also asked if it was easy to view all the images in the HMD 3D image browser when comparing the two interfaces. No significant difference was found between the two ( $p = 0.144$ ). It is worth to mention that the fact that the participants could only make use of the head tracking ability in *IN* did not mind having to use the head tracking as only ability to explore when placed in the center of the cylinder.

Participants felt that the controls were intuitive and easy to learn for both *IN* and *IS* (rated 4.5 out of 5 on average). This includes the usage of the head tracking as an exploring mechanism, as well as using the multiple buttons on the controller. Using the head tracking ability to explore was as easy to use as the controller (rated 3 out of 5).

**Physical and Mental State** There is no significant difference between the two interfaces in how easy it was to remember where the images were located ( $p = 0.266$ ). This is also supported by the fact that the two interfaces performed the same for the known-item search tasks in terms of speed and the amount of errors.

*IN* did not have the ability to use the *Left Trigger* and *Right*

*Trigger* buttons of the controller to scroll. Participants could only make use of the head tracking ability. They did endure physical stress on their neck and head. Even though the values are not that high (rated 2.2 out of 5 on average), it is still worth to mention. There is a significant difference when comparing *IN* with *IS* ( $p = 0.009$ ). Even though the head tracking did not bother any of the participants (rated 3 out of 5), it is advised to not rely on the head tracking ability to explore since this can cause physical stress.

The participants were also asked if the application affected their mental state in a bad way, for example nausea or stress. There is no significant difference between *IN* and *IS* ( $p = 0.605$ ). They said that *IS* caused the most nausea and/or stress (rated 2.2 out of 5). This could be due to the fact that the user could use the head tracking ability as well as the controller to explore. The combination of these two could lead to more nausea and/or stress. *IN* does not have this problem, since it cannot make use of the controller.

### 5.3 Placement Experiment

**Performance** No significant difference was found in the amount of time it took the participants to complete the known-item search task when comparing *IS* with *OS* ( $p = 0.095$ ). Placing the user in the center of the cylinder or with a distance of  $2 * \text{CylinderRadius}$  from the center of the cylinder is the same in terms of speed (figure 11). The amount of errors (figure 12) are also not influenced by the different placement of the user in the two interfaces ( $p = 0.435$ ).

There is also no significant difference between the usage of the controller and the usage of the head tracking ability to explore in *IS* ( $p = 0.469$ ). The participants made slightly more use of the *Left Trigger* and *Right Trigger* buttons to explore as can be seen in figure 13.

As expected, there is a significant difference between the usage of the controller and the usage of the head tracking ability in *OS* ( $p = 0.001$ ). Participants made about twice as much use of the scroll buttons. This is most likely due to the fact that they were forced to use the controller when they wanted to view all images because of the placement outside of the cylinder. Also, the images in front of the participants are directly in their perceptual span (section 1) and thus the need to use the head tracking ability is reduced. However, the fact that it is still being used is most likely due to the fact that it feels very natural to look around inside the virtual environment, just like humans do in real life.

There is also a significant difference between the usage of the head tracking ability when comparing *IS* and *OS* ( $p = 0.033$ ). This can also be related due to the fact that the participants felt more immersed in terms of having a sense of presence in the virtual world in *IS* than in *OS* ( $p = 0.001$ ). This can also be seen in figure 13, participants made about 1.5 times as much use of the head tracking ability in *IS* when compared to *OS*. There is potentially a correlation between the usage of head tracking and the user experience.

**User Experience** There is no significant difference in how easy it is to view all the images in the HMD 3D image browser when comparing *IS* and *OS* ( $p = 0.576$ ). Both interfaces had the head tracking ability as well as the option to use the controller. Participants did find that *IS* had the highest user experience in terms of that they felt like they were really surrounded by the pictures (rated 4 out of 5) while they gave *OS* a lower score (3 out of 5), with a significant difference ( $p = 0.001$ ). However, they also mentioned that they were more focused on the known-item search task in *OS* than in *IS* ( $p = 0.008$ ). This could be due to the fact that they were

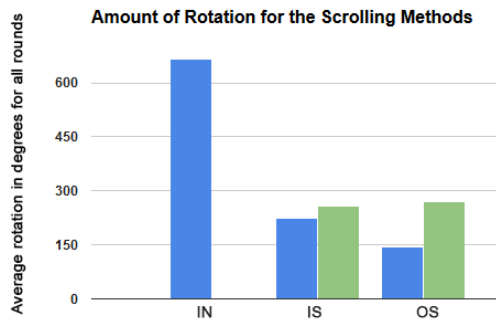


Figure 13: The average rotation in degrees for all rounds using the HMD(Blue) and the controller (Green).

less distracted by the images surrounding them and thus had more attention towards the given task. It is important to note that this significant difference is not reflected in the performance in terms of speed and the amount of errors. Thus, placing the user outside the cylinder with a distance of  $2 * CylinderRadius$  from the center of the cylinder severely lowers the user experience, while giving the feeling to the user that he is more focused on the given task.

**Physical and Mental State** There is again no significant difference between the two interfaces in how easy it was to remember where the images were located ( $p = 0.544$ ). Based on this score, the placement of the user potentially does not matter in terms of memory and performance. There is also no significant difference in the endurance of physical stress between *IS* and *OS* ( $p = 0.076$ ).

A significant difference was found between how the interfaces affected the mental state of the participants ( $p = 0.034$ ). As seen in section 5.2, *IS* caused the most nausea and/or stress (rated 2.2 out of 5). A score of 1.6 out of 5 was given to *OS*. This could potentially mean that using a controller while being placed outside the cylinder as well as lowering the amount of information the user can see causes less nausea.

## 6 CONCLUSION & FUTURE WORK

This paper discussed the usage of head mounted displays for image browsers. The Oculus Rift DK1 was used to create an image browser by creating an interface 'inside' a virtual environment. Three different cylindrical interfaces were introduced for this experiment, in which 22 users participated. They were all very enthusiastic about the interfaces and they would definitely use the application more if they had the chance. Out of the three interfaces, *IS* was liked the most. This interface gave the highest feeling of being present inside the virtual environment, while being easy to control in terms of the interaction with the interface. The following questions were answered based on the experiment;

*"What is the most efficient way to implement an exploration mechanism for HMD 3D image browsers, using a cylindrical interface, in terms of performance and user experience."*

Results show that, in terms of performance, it does not matter if only the head tracking ability (*IN*) or if both the head tracking ability and the ability to scroll (*IS*) are included when the user is placed in the center of the cylinder. However, not including the scrolling mechanism will generate significantly more physical strain on the participants. Thus, utilizing both the head tracking ability as well as the scrolling is advised. An important thing to note is that the head tracking ability created the sense of presence

and being surrounded by the images in the virtual world.

*"Should the user in the virtual environment be placed inside or outside a cylindrical interface to optimize the user experience and performance."*

In terms of performance, it does not matter if the user is placed at the center of the cylinder (*IS*) or outside the cylinder with a distance of  $2 * CylinderRadius$  from the center of the cylinder (*OS*). However, there is a significant difference in terms of user experience. Participants did not feel like they were being surrounded by the images of the interface in *OS* and mentioned that they were more focused on the known-item search tasks and experienced less physical stress and nausea. Thus, placing the user in the center of the cylinder increases the experience at the cost of some physical stress and less focus on the given task.

The implementation of our HMD 3D image browsers used cylindrical shapes. These shapes are the most obvious to use when one of the goals is to increase the feeling the user is surrounded by the interface. Future research could explore a lot of other shapes, each with their own unique properties, which may or may not perform better than cylindrical interfaces in certain situations. The same can be said for different radii for the cylindrical interfaces and different placements of the user inside the virtual environment. There are an infinite amount of (non-trivial) possibilities which could improve the experience as well as the performance of the interfaces.

Another interesting topic is how interaction and feedback is handled. This paper used a controller to interact with the interface and used a simple hovering method to inform the participants that a certain image was highlighted. However, interaction can be done by many other devices and the feedback can be done in many other ways like, using cursors or color coding.

One of the problems of the Oculus Rift DK1 was the low resolution of the device ( $1280 \times 720$ ). Even though participants mentioned that they were not bothered by the low resolution, the interfaces can be extended with the use of the Oculus Rift DK2. This version of the Oculus Rift has a resolution of  $1920 \times 1080$ . This could improve the user experience as well as the performance of the user.

An interesting idea to expand the HMD 3D image browser is by using folders. Most images on the internet have tags associated with them. This is also the case with the Flickr database used in this paper. These tags can be used to browse through the database more quickly by creating folders based on these tags. When a folder is selected, newly loaded images are shown based on the tags of the folder. This may speed up the process of searching specific images since the user can navigate with the help of these tags instead of having to browse through the whole database. It may also increase the user experience due to the fact that the user can browse through the database more user-friendly.

## REFERENCES

- [1] R. P. K. A. Postma and M. van Asselen. How the brain remembers and forgets where things are: The neurocognition of objectlocation memory. In *Neuroscience & Biobehavioral Reviews*, volume 32.8, pages 1339–1345, 2008.
- [2] A. Cockburn. Revisiting 2d vs 3d implications on spatial memory. In *Proceedings of the fifth conference on Australasian user interface - Volume 28*, pages 25–31, 2004.
- [3] A. Cockburn and B. McKenzie. 3d or not 3d?: evaluating the effect of the third dimension in a document management system. In *Pro-*

- ceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 434–441, 2001.
- [4] A. Cockburn and B. McKenzie. Evaluating the effectiveness of spatial memory in 2d and 3d physical and virtual environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 203–210, 2002.
- [5] A. Cockburn and B. McKenzie. Evaluating the effectiveness of spatial memory in 2d and 3d physical and virtual environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 203–210, 2002.
- [6] Google Inc. Google Cardboard - Virtual reality on your smartphone. <https://cardboard.withgoogle.com/>, 2014. accessed Oct. 2014.
- [7] J. Gregory. *Game engine architecture*. CRC Press, 2009.
- [8] IMDb.com Inc. Star Wars: Episode IV - A New Hope. [http://www.imdb.com/title/tt0076759/?ref\\_=fn\\_al\\_tt\\_1](http://www.imdb.com/title/tt0076759/?ref_=fn_al_tt_1), 1977. accessed Oct. 2014.
- [9] IMDb.com Inc. Total Recall. [http://www.imdb.com/title/tt0100802/?ref\\_=nv\\_sr\\_2](http://www.imdb.com/title/tt0100802/?ref_=nv_sr_2), 1990. accessed Oct. 2014.
- [10] IMDb.com Inc. Pacific Rim. [http://www.imdb.com/title/tt1663662/?ref\\_=nv\\_sr\\_1](http://www.imdb.com/title/tt1663662/?ref_=nv_sr_1), 2013. accessed Oct. 2014.
- [11] IMDb.com Inc. Edge of Tomorrow. [http://www.imdb.com/title/tt1631867/?ref\\_=nm\\_filmg\\_act\\_4](http://www.imdb.com/title/tt1631867/?ref_=nm_filmg_act_4), 2014. accessed Oct. 2014.
- [12] K. Schoeffmann and D. Ahlström. 3D Interfaces to Improve the Performance of Visual Known-Item Search. [http://vidosearch.com/publications/3DInterfaces\\_preprint.pdf](http://vidosearch.com/publications/3DInterfaces_preprint.pdf), 2014. accessed Oct. 2014.
- [13] S. M. A. M. K. Young, G. B. Gaylor and B. Bodenheimer. A comparison of two cost-differentiated virtual reality systems for perception and action tasks. In *Proceedings of the ACM Symposium on Applied Perception*, pages 83–90, 2014.
- [14] Oculus, V.R. Oculus rift-virtual reality headset for 3d gaming. <http://www.oculusvr.com>, 2012. accessed Oct. 2014.
- [15] Oculus, V.R. *Oculus Development Kit 1 SDK Overview*, May 2014.
- [16] P. S. P. Cubaud and A. Topol. Binding browsing and reading activities in a 3d digital library. In *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, pages 281–282, 2002.
- [17] C. Plaisant, D. A. Carr, and B. Shneiderman. Image browsers: Taxonomy, guidelines, and informal specifications. 1994.
- [18] K. Rayner. The perceptual span and peripheral cues in reading. In *Rayner, K. (1975). The perceptual span and peripheral cues in reading. Cognitive Psychology*, volume 7.1, pages 65–81, 1975.
- [19] K. Rodden, W. Basalaj, , D. Sinclair, and K. Wood. Does organisation by similarity assist image browsing? In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 190–197, 2001.
- [20] J. P. Rolland and H. Hua. Head-mounted display systems. In *Encyclopedia of optical engineering*, pages 1–13, 2005.
- [21] M. J. S. Boring and A. Butz. Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. In *OZCHI '09 Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, pages 161–168, 2009.
- [22] P. W. S. Mehra and M. Worring. Navigating on handheld displays: Dynamic versus static peephole navigation. In *ACM Transactions on Computer-Human Interaction*, volume 13.4, pages 448–457, 2006.
- [23] K. Schoeffmann and D. Ahlstrm. Using a 3d cylindrical interface for image browsing to improve visual search performance. In *Image Analysis for Multimedia Interactive Services*, pages 1–4, 2012.
- [24] Shoojee GmbH & Co. KG. Durovis Dive - the world's first hands-free VR headset for smartphones. <http://www.durovis.com/index.html>, 2013. accessed Oct. 2014.
- [25] A. R. Smith. Color gamut transform pairs. In *ACM Siggraph Computer Graphics*, volume 12. No. 3, pages 12–19, 1978.
- [26] I. E. Sutherland. A head-mounted three dimensional display. In *fall joint computer conference*, volume I, pages 757–764, Dec. 1966.
- [27] M. Tavanti and M. Lind. 2d vs 3d, implications on spatial memory. In *Information Visualization, 2001. INFOVIS 2001. IEEE Symposium on*, pages 139–145, 2001.
- [28] Unity Technologies. Unity3D Game Engine. <http://unity3d.com/>, 2014. accessed Oct. 2014.
- [29] Yahoo! Labs. The Yahoo! Webscope Yahoo Flickr Creative Commons 100M (YFCC-100M) dataset. <http://webscope.sandbox.yahoo.com/catalog.php?datatype=i&did=67>, 2014. accessed Oct. 2014.



## 7 ANNOTATED APPENDIX

### 7.1 Introduction

The annotated appendix contains detailed information which cannot be found in the paper. This section will describe the workload of the authors as well as give an overview of the content of the deliverables for this research. Section 7.2 explains the technical details of the application. This includes how to start the application, how it works and how to change certain parameters in the program. The database used for this application as well as a guide on how to download the database is explained in section 7.3. An overview of the results of all participants measured by the application is shown in section 7.4 and an overview of the results of thall participants collected by the questionnaire is shown in section 7.5.

#### 7.1.1 Workload

This research can be divided in a couple of important parts. These parts, as well the hours spend on each part, can be found in table 1. Wolfgang was supervising and his time spent is not included in the table. *Research* contains the time it took for preliminary research as well as literature research. *Discussion* contains the time it took for all meetings. *Application* contains the time it took to implement the HMD 3D image browser. *Experiment* contains the time it took to conduct the experiment, *document* contains the time worked on the document and *total* it the total amount of time we worked on the project.

The work is divided in such a way that we, Kevin and Johnny, worked evenly on every part. The general workflow was that one of us worked on a part and when finished for the day, the other took over. This is because each of us sees things in a different way and thus we supplement each other very well. When working both on a single part, for example the application, each of us might find errors as well as come up with solutions to problems while the other does not and vice versa.

Part of Research	Time in Hours
Research	109
Discussion	29
Application	285.75
Experiment	89
Document	302
Total	814.75

Table 1: Table containing the work in hours divided in several important parts of the research.

#### 7.1.2 Deliverables

Everything we created for this project as well as all the data gathered will be included into the deliverables. The root folder has the following folder structure;

**Application** Contains a Unity3D project, which is a standard Unity3D directory including source code, assets and libraries.

##### Database

**Guide** Contains a .TXT file that explains how to download the Flickr database.

**Database Formatter** Contains a C# application and the source code which processes parts of the Flickr database to remove redundant data.

##### Documents

**Latex files** Contains the latex file as well as all images and other necessary files to compile the latex file.

**Paper** Contains a .PDF file of the paper.

**Info for the participant** Contains the information presented to the participant at the start of the experiment.

**Questionnaire** Contains the questionnaire used in the experiment.

**Log** Contains a detailed spreadsheet of all the information on who worked on what part of the project and when.

##### Media

**Images** Contains extra images not used in the paper itself.

**Videos** Contains a demo video and video footage used to create the video.

##### Results

**Application Measurements** Contains the raw .TXT files with information generated by the application.

**Application results** Contains a detailed spreadsheet of the information recorded by the application.

**Questionnaire results** Contains a spreadsheet with the answers of the questionnaire.

## 7.2 Implementation

**How to Start** The application is made in Unity3D. To open the project, you can either load it from within Unity3D or open one of the *.unity* files in the *Assets/Scenes* directory. When the project is open, you can click the play button in Unity3D to start the application. Three different scenes are created for the three different interfaces. To switch between the different interfaces another scene needs to be opened.

**How the Application work** How the application works has been explained in section 3. Parameters can be changed in the inspector of the objects in the scene. Make sure to do this before starting the application. Relevant parameters can be found in the *Main*, *ImageHandler* and *PhaseHandler* GameObjects.

**Additional Information** At the beginning of this project, we implemented a version of the HMD 3D image browser which included 'folders'. Each of these folders had a single tag associated with it and when a folder was selected, new images from the Flickr database associated with that tag were loaded (e.g. clicking on a folder with the tag 'trees' would result in new images with trees being loaded from the database). This made browsing through the database easier. However, this idea is not used in the final version of the application because the focus of the research changed in another direction. It can still be turned on if needed in the *Main* GameObject in the Unity3D editor.

There are also simple *Scroll Buttons* which can be turned on in the *Main* GameObject. These buttons were used to 'scroll' through different pages of images. This type of scrolling was later changed by using the controller buttons and the HMD head tracking functionality because we changed the interface to a cylinder surrounding the user instead of having one third of a cylinder as interface.

## 7.3 Database

The Flickr database was used for this experiment. There are several steps to acquire and get permission to use the database. These steps will be listed here.

First you need to go to the following link and request access to the database: <http://webscope.sandbox.yahoo.com/catalog.php?datatype=i&did=67>. Once your request has been granted, you get an email stating step-by-step instructions. However, because of issues downloading the database on a Windows machine, we downloaded the *S3 Browser* application and added an external bucket with the link provided in the email. We could then download the database by opening this bucket, right-click and press download. Make sure you have the right features activated within *Amazon*, you need to give your credit card information as well as some other personal information.

The database consists of large text files with the url to the image as well as a lot of tags for each image. The database formatter created by us throws away the redundant information, filters the images based on given parameters and reformats it to make sure the files are much smaller in size (and thus less loading times are required). The usage of the database formatter is self explanatory, everything is explained in the application itself.

## 7.4 Quantitative data analysis

When a participant finished the experiment in an interface, data was written to a file in the *Data/Setup* directory. This file contains the following information;

- Name of the interface.
- Seed value used for the random generator.

- For each KIS-task round:
  - Time it took for the introduction phase.
  - The image shown in the introduction phase.
  - If the participant found the image.
  - Time it took for the KIS-task.
  - Time it took for the participant to select the right image.
  - Number of errors a participant made.
  - The amount of rotation of the Oculus Rift as well as from input.
  - Heatmap of where the participants looked at.

The relevant information was then processed and put into spreadsheets. This information can be found in figure 14, 15, 16 and 17.

## 7.5 Qualitative data analysis

The questionnaire was used to measure the user experience of the user in terms of having the feeling of being inside the virtual environment surrounded by images. All the results can be found in the paper itself, section 5. The questionnaire itself was divided into several parts to make it more understandable for the participants as well as make it easier to analyse the results by grouping questions together. This was done as follows;

**Personal Data** Information regarding gender, age as well as usage of the hardware.

**General Data** How much the participants enjoyed the interfaces.

**Immersion** How the participants felt in terms of the feeling of being inside the virtual environment.

**Interaction** How easy and natural it was to interact with the interfaces.

**Physical & Mental state** If the participants experienced stress on the body or mind.

**Other** Additional remarks of the participants.

The raw data can be found in the deliverables.

Average number of errors for all rounds			
Participant ID	IN	IS	OS
0	0	0.1	0
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0.1	0
5	0	0.3	1.3
6	0	0.1	0
7	0	0	0
8	0	0	0
9	0	0.1	0.1
10	0.2	0.8	0.1
11	0.1	0.1	0.1
12	0	0	0
13	0.6	0	0
14	0.4	0.4	0
15	0	0	0
16	0	0	0
17	0	0.1	0.1
18	0	0.2	0
19	0	0	1.6
20	0.1	0.2	0.3
21	0	0	0.4
22	0.4	0	0.2
<b>Average</b>	<b>0.07826086957</b>	<b>0.1086956522</b>	<b>0.1826086957</b>

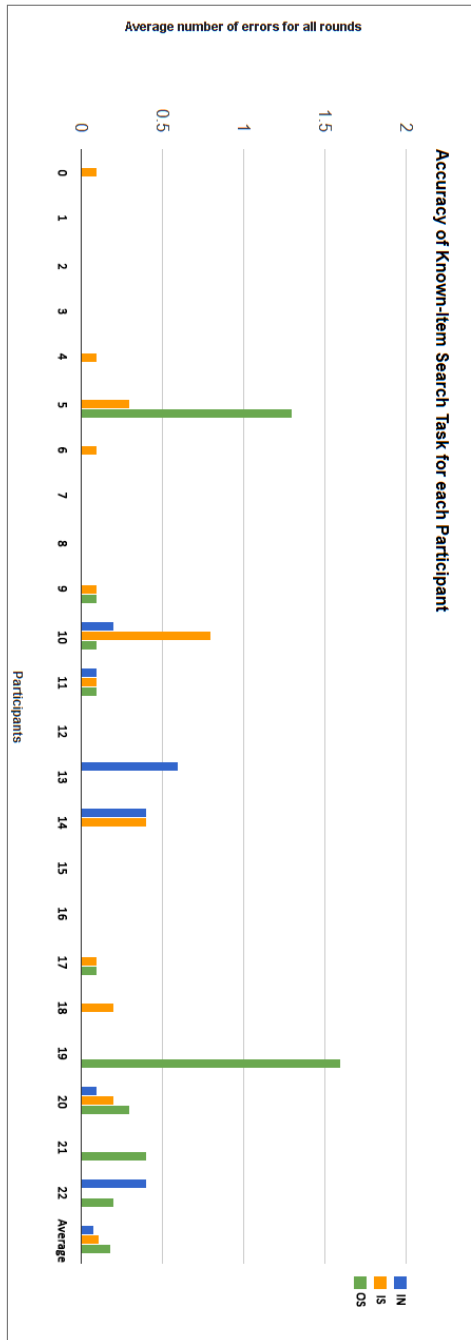


Figure 15: The average number of errors each participant made for all rounds of the three interfaces. It also includes the average number of errors of all participants combined. (pressing the Y button when a wrong image was selected).

Average speed in seconds for all rounds			
Participant ID	IN	IS	OS
0	21.419	15.853	9.957
1	23.223	14.276	17.875
2	26.49	22.041	13.054
3	12.669	23.493	23.528
4	61.625	39.953	32.166
5	36.554	29.761	45.973
6	27.912	21.716	29.956
7	28.493	22.21	42.229
8	15.855	19.234	16.065
9	29.109	27.019	22.215
10	28.709	54.946	49.25
11	46.501	20.238	43.011
12	22.382	14.851	15.793
13	24.132	12.691	11.86
14	51.548	47.013	49.477
15	21.309	12.406	13.778
16	8.429	10.524	6.659
17	18.05	13.36	17.292
18	23.439	29.074	12.348
19	17.468	17.593	29.958
20	40.371	12.7	30.271
21	12.213	12.713	34.788
22	38.047	49.568	70.221
<b>Average</b>	<b>27.60204348</b>	<b>23.61882609</b>	<b>27.72156522</b>

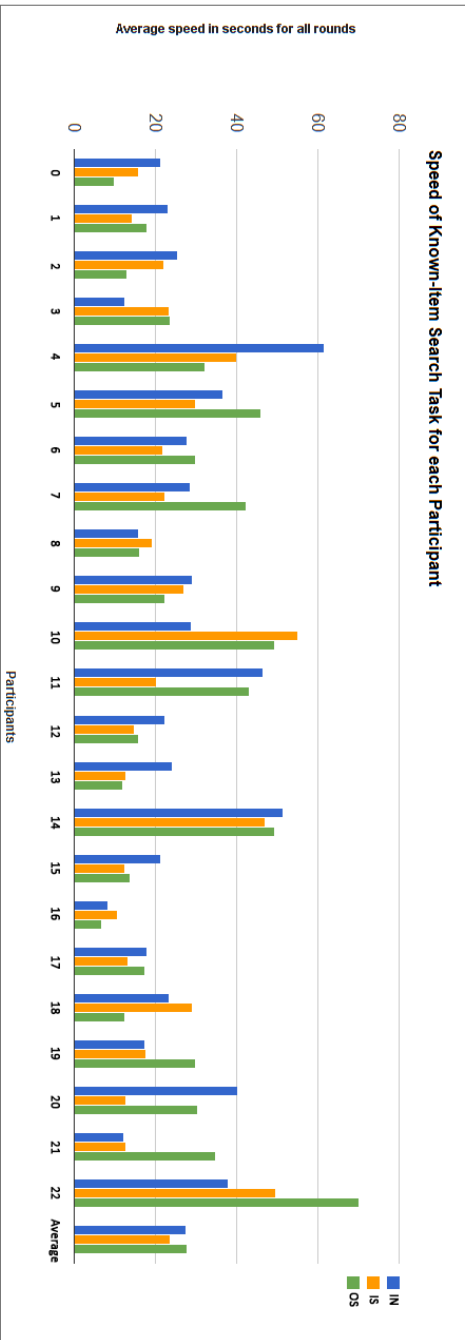


Figure 14: The average speed in seconds of each participant for all rounds of the three interfaces to perform the known-item search task. It also includes the average speed of all participants combined. (pressing the Y button on the controller when the correct image was selected).

Average rotation in degrees of the HMD for all rounds			
Participant ID	IN	IS	OS
0	737.6579	606.01422	89.344031
1	636.58406	188.710669	114.508543
2	612.224545	144.948032	59.696731
3	443.703227	100.569378	80.917611
4	816.55815	149.662127	160.332059
5	764.25319	145.452501	179.76827
6	711.73068	431.09173	448.47869
7	501.276725	93.621913	123.65539
8	377.903554	204.238425	87.54423
9	513.442	151.614057	88.283492
10	702.188009	242.601008	266.474295
11	714.52066	150.463612	200.674123
12	703.40028	87.806742	40.4914379
13	687.79791	333.742781	41.704774
14	1153.167938	418.325478	342.699412
15	573.91666	41.8224526	34.2952799
16	498.951272	557.010325	38.432718
17	680.489359	155.920666	41.6788147
18	840.650655	329.675799	97.034288
19	650.097785	209.115959	175.411546
20	803.56545	122.136537	56.6563811
21	551.30305	68.954577	288.003655
22	681.9435	207.684294	228.724366
<b>Average</b>	<b>667.7098943</b>	<b>223.5297075</b>	<b>142.817676</b>

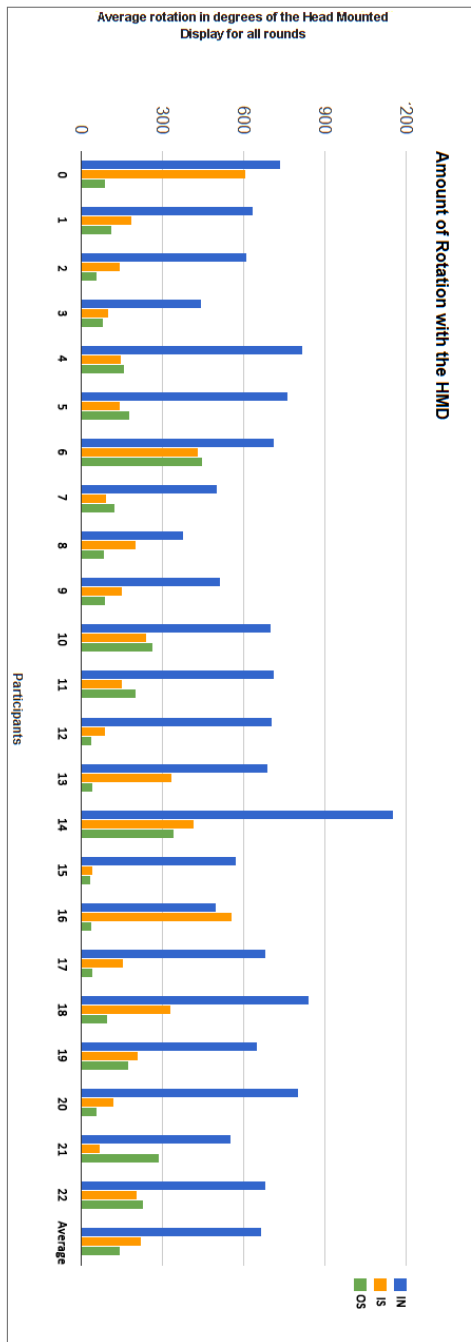


Figure 17: The average rotation in degrees each participant made using the HMD for all rounds of the three interfaces. It also includes the average rotation in degrees of all participants combined.

Average rotation in degrees of the controller for all rounds			
Participant ID	IN	IS	OS
0	0	18.9471	209.5703008
1	0	154.946384	188.4171
2	0	255.163659	137.008541
3	0	285.29296	237.570312
4	0	355.78778	194.92927
5	0	374.991155	377.601423
6	0	398.124573	332.276199
7	0	274.065329	269.4608548
8	0	244.390507	160.985738
9	0	162.549622	132.3469431
10	0	643.976884	490.026587
11	0	227.011241	224.030419
12	0	219.76882	216.566536
13	0	192.761604	175.519468
14	0	370.276619	421.922554
15	0	239.201826	177.917573
16	0	0.2191867	140.8216199
17	0	305.33727	322.01075
18	0	330.519844	216.947489
19	0	240.296938	277.7632
20	0	109.2575326	353.089921
21	0	230.059501	415.505042
22	0	301.961849	568.923737
<b>Average</b>	<b>0</b>	<b>258.0395298</b>	<b>271.3570342</b>

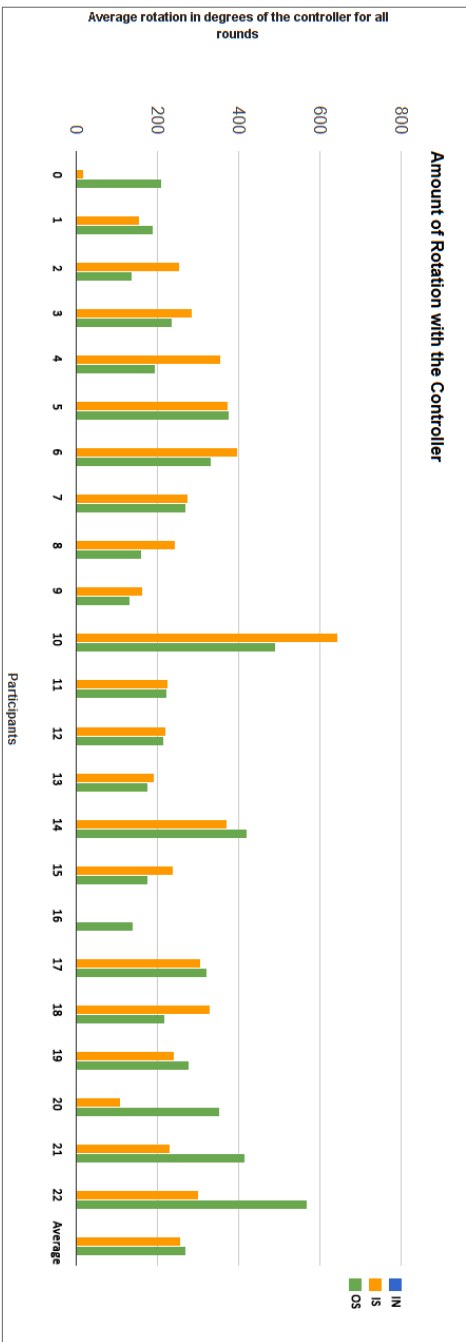


Figure 16: The average rotation in degrees each participant made using the controller for all rounds of the three interfaces. It also includes the average rotation in degrees of all participants combined.