

Immersion in Desktop Virtual Reality

George Robertson, Mary Czerwinski, and Maarten van Dantzich

Microsoft Research
One Microsoft Way
Redmond, WA 98052, USA
Tel: 1-206-703-1527
E-mail: ggr@microsoft.com

ABSTRACT

This paper explores techniques for evaluating and improving immersion in Desktop Virtual Reality (VR). Three experiments are reported which extend findings on immersion in VR reported by Pausch et al. [9]. In the current experiments, a visual search paradigm was used to examine navigation in Desktop VR both with and without navigational aids. Pausch et al. found that non-head tracked users took significantly longer than predicted when the search target was absent, which was interpreted as indicative of a loss of sense of immersion. Our first experiment extended the Pausch et al. experiment to a desktop display. Our findings differ in that search times matched prediction when the target was absent, indicating that the Pausch et al. study does not transfer to Desktop VR. In the second and third experiments, our visual search task was performed while navigating a set of 3D hallways. We introduce a new navigation aid called *Peripheral Lenses*, intended to provide simulated peripheral vision. Informal studies suggested that Peripheral Lenses decrease search time, indicating an enhanced sense of immersion in Desktop VR. However, formal studies contradict that, demonstrating the importance of formal usability studies in the development of user interface software. We also gained evidence that visual attention findings transfer to Desktop VR.

KEYWORDS: Virtual Reality, Immersion, Evaluation, Visual Search Paradigm.

INTRODUCTION

Ivan Sutherland implemented the first virtual reality (VR) system in 1968 [14]. By wearing a head-mounted display (HMD) on which wire-frame graphics were displayed, users perceived that they were occupying the same space as virtual objects. The goal of VR is to place the user in a

three dimensional environment that can be directly manipulated, so the user perceives interaction with the environment rather than the computer [8].

It is commonly believed (but not yet proven) that VR attains its power by captivating the user's attention to induce a sense of *immersion*. This is usually done with a display that allows the user to look in any direction, and that updates the user's viewpoint by passively tracking the user's head motion. HMDs and CAVEs [3] are examples of these kinds of displays. However, there are other forms of VR where immersion occurs. Fish Tank VR [5, 20] uses passive head tracking, but uses a desktop stereo display rather than surrounding the user visually. Desktop VR [10] is the use of animated interactive 3D graphics to build virtual worlds with desktop displays and without head tracking.

It is important to be clear about the meaning of the term immersion. Webster defines it as "the state of being absorbed or deeply involved." Clearly, immersion can occur while we are watching a movie or television, or while playing video games (many of which are examples

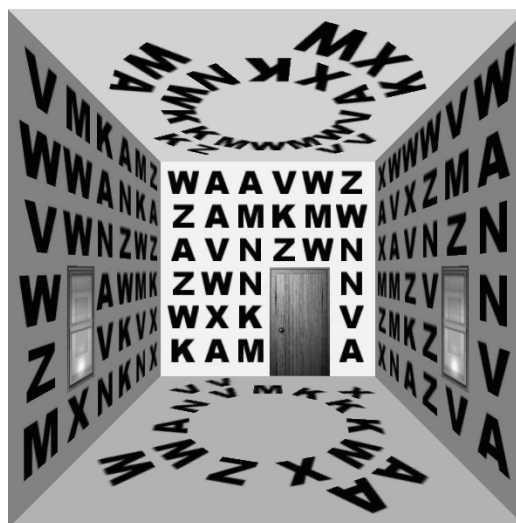


Figure 1: Rotation Study Room. The missing wall has only letters on it.

of Desktop VR). Experience suggests that proper 3D cues and interactive animation aid immersion, and that the act of controlling the animation can draw the user into the 3D world. In other words, immersion should not be equated with the use of head-mounted displays: mental and emotional immersion does take place, independent of visual or perceptual immersion [4]. In addition, current HMD-based VR techniques suffer from poor display resolution, display jitter, and lag between head movement and the resulting change to the display. These problems tend to inhibit the illusion of immersion, and are not a problem in Desktop VR systems.

A common criticism of Desktop VR focuses on the lack of peripheral vision afforded by a desktop display, claiming that users are not aware of their surroundings or of their location in the virtual space. This is closely related to the issue of how immersive a Desktop VR system can be. In this paper, we will describe some attempts to measure immersion. We will also describe some Desktop VR navigation aids that increase the user's awareness of what surrounds them, hence presumably increase immersion.

VISUAL SEARCH

When playing a game in 3D or navigating through a 3D environment, it is often difficult to get a sense of "place", or know what is behind you. Many of these navigational experiences involve tasks in which the user is searching for a key item or piece of information. Among everyday tasks that rely heavily on visual search capabilities are searching for a certain product brand on a grocery shelf and looking for a known reference on a newspaper page. When these common tasks are transferred to a virtual world, they may become more difficult because of the lack of sense of place. Little is known about searching for stimuli in a 3D environment, even on the desktop, and this will be the focus of this paper.

Much research on visual search has been carried out in 2D, however. It will be interesting to see what, if any, differences there are between searching with an added visual dimension. Most, if not all, of the theories of attention and visual search today assume search through a display to be limited in capacity unless the target item can be defined by a unique, separable feature, such as the color red [2, 18, 19], or unless specific training situations are included in the testing. Typically, researchers have used the visual search paradigm to explore cognitive capacity issues in attention and search. In visual search tasks, the user is instructed to search a display of one or more visually presented items for the presence of one or more targets currently held in working memory. The general finding in such tasks is that performance is best characterized as limited in capacity (i.e., relatively slow, and sensitive to attentional demands) [16, 17]. The typical visual search experiment would also consist of several different display set sizes (e.g., 4, 8, or 12 stimuli

present). When such experiments are run, there is usually a linear relationship between the reaction time and display set size function (typically around 38 msec. per item for target absent functions, and about half of that for target present functions [13]). This limited capacity visual search has been found to be especially resistant to the effects of training when targets and distracters are chosen randomly from the same set [11, 12]).

Thus, the visual search paradigm is an excellent context in which to investigate cognitive factors influencing 3D navigation. A concern in traditional (2D) visual search studies is whether there are some search tasks that can be executed automatically, that is, whether there are tasks that can be performed in a seemingly effortless fashion, bypassing more time-consuming processes and allowing attention to be allocated elsewhere. Automatic visual search can be, in certain cases, a pre-attentive process that does not rely on slower, cognitive processes. As noted above, evidence for this type of automaticity is found in some studies of visual search for basic, easily discriminable, features of target stimuli such as unique coloring (see [2, 18]).

In the experiments reported in this paper, however, highly confusable letter sets were used intentionally in order to examine effortful, limited-capacity search. Pausch et al. [9] reported that searching for letters in a simulated 3D desktop environment has indeed proved problematic. Experiment 1 investigates extending that result to a Desktop VR environment. Experiments 2 and 3 examine the use of *Peripheral Lenses*, a new technique intended to alleviate problems navigating and searching in Desktop VR. Along the way it will be observed how well traditional findings about attention and visual search in 2D studies can be leveraged as we attempt to design for search and navigation tasks in Desktop VR.

Previous Quantification of Immersion in VR

Pausch et al. [9] have provided one of the first attempts to quantify immersion in VR, using a visual search task to measure immersion. The task involved looking for a specific target letter in a set of similarly shaped letters distributed around the walls, floor and ceiling of a virtual room (see Figure 1, which is our implementation of Pausch et al.'s study). The users had to find the letter or determine that the letter was absent in the room. Using similarly shaped letters made this a cognitively demanding task rather than a pre-attentive task, so the limiting factor on completion time was not the speed at which the user's viewpoint could be moved, but rather attentive examination of the stimulus. Assuming random placement of the target letter in the room, on average the user will have to search half the room when the letter is present. Determining that a letter is absent requires searching the entire room, and thus would be predicted to take twice as long as the average search time when the letter is present.

Pausch et al. studied two conditions, which will be referred to as HMD VR and Fixed-HMD VR, respectively. In both cases, the user wore an HMD and was positioned in the center of the virtual room, controlling rotation of the body and head (i.e., the user could look around the room, but not move around in it). In the first case (HMD VR), the HMD was free to move and the virtual viewpoint was updated by tracking the user's head movements with a Polhemus tracker. In the second case (Fixed-HMD VR), the HMD was fixed in place and the subject used a two-handed device with an embedded Polhemus tracker to control the camera. One part of the device was held with the non-dominant hand while the other part was moved around with the dominant hand. The two rotational degrees of freedom offered by the device resemble those of the human head: pitch and yaw. The protruding tracker wire made it difficult to turn the device more than 180 degrees in either direction. In Pausch et al.'s study, the Fixed-HMD VR condition was referred to as Desktop VR. We believe that Fixed-HMD VR is so different from Desktop VR that we use a different term.

Pausch et al. reported the following results.

1. When the target letter was present, average search times for the two conditions were the same.
2. When the target letter was absent, HMD VR matched the predicted search time, but Fixed-HMD VR was 41% slower than predicted. Pausch et al. suggest that HMD VR users had a better mental model of the space, hence avoided redundant scanning, or rechecking the display for the target letter.
3. There was positive transfer of training from an initial HMD VR session to the Fixed-HMD VR.
4. There was negative transfer of training from an initial session in Fixed-HMD VR to an HMD VR. Pausch et al. suggest that this demonstrates that HMD VR users build a better mental model of the space.

Boyd [1] has also studied the effects of immersion. In his study, three different kinds of virtual environments are compared. The first used an HMD with head tracking. The second used a tracker the way a puppet is used. The third used a tracker to control a virtual vehicle. The second and third used desktop displays with the tracker being held in the dominant hand. Subjects were placed in a virtual space in which they could move around in a small area. The task was to locate a virtual object that looked like a telescope, walk up to it and look through it. Boyd's results show that the HMD with head tracking was superior to the other two systems.

Issues Raised by Previous Studies

The Pausch et al. and Boyd studies raise more questions

than they answer. Pausch et al.'s Fixed-HMD VR is quite different from Desktop VR systems. First, it involves the user wearing an HMD fixed in one position. They used this configuration to balance the display resolution variable. The result was that users received the low resolution of HMDs without their advantage of head-centric camera control. It has been known for some time [5, 10] that both Desktop VR and Fish Tank VR have an advantage over HMD VR by providing higher resolution. Second, Pausch et al.'s system involves using an unfamiliar two-handed input device.

Hence, the first issue is whether Pausch et al.'s results transfer to a more traditional Desktop VR, with a desktop display and a mouse. This includes determining how the task completion times compare, whether one sees the same slowdown in the target absent case, and the same kind of rescanning. If this is not the case, can those measures really be used for measuring immersion?

Furthermore, landmarks may account for differences in the Pausch et al. study versus other studies of immersion. This issue applies to Pausch et al.'s study because the floor and ceiling have no landmarks, making those a likely area where users will become disoriented. The walls do have landmarks, just to avoid this problem. Obviously, if the visual space has no landmarks, it will pose a more difficult navigation and search task, but one that could easily be improved by adding simple landmarks.

Do Pausch et al.'s results transfer to navigation tasks? The Pausch et al. study involves a user fixed in one location, able only to turn around and look up and down. Boyd's study limits the user to navigation in a very small area. What happens when the user needs to move around a large or complex environment, which is common in Desktop VR systems?

Finally, even if Desktop VR is shown to be consistently slower or more demanding than HMD VR, there may be simple navigation aids that can be added to Desktop VR to eliminate the performance difference.

EXPERIMENT 1: ROTATION STUDY

In this first experiment, we test the Fixed-HMD VR condition in Pausch et al.'s [9] study using a desktop display rather than a fixed HMD (hence changing it from Fixed-HMD VR to Desktop VR). We started with the same virtual room (four meters on a side), the same stimuli, and the same input device. In spite of having a different display, we chose the vertical focal angle (75 degrees) and aspect ratio (1.2:1) so that a given view displayed approximately the same number of letters as seen in Pausch et al.'s HMD. As in Pausch et al.'s study, the room had a number of landmarks to help the user, including a door, windows on two of the walls, and slightly different colors for the walls. In one condition,

we used the same two-handed input device described above. In another condition, we used a mouse for input by mapping mouse movement forward and backward to head up and down, and mouse movement left and right to rotation of the body left and right. In Pausch et al.'s experiments, the user indicated task completion by speaking to the experimenter. To control for this extra lag, our subjects responded by pressing one of two buttons to indicate their responses. Since the two-handed device had no buttons, we accompanied it with a rocking foot switch. Subjects using the mouse did the experiment once using the foot switch, and once using the mouse buttons.

This experiment ran on a Gateway 2000 200 MHz Pentium Pro with an Intergraph Intense 3D Pro graphics card, was implemented in C++ using OpenGL, and ran in Windows NT 4.0. Frame times were around 37 frames per second.

Method

Subjects. Thirty-two subjects (twenty-three males) participated in the study. All had normal or corrected-to-normal vision. Age of the participants ranged from 18-50 years old.

Stimuli and Design. For both experiments, the principal dependent variables were response time from onset of display and response accuracy. All subjects were instructed to maintain a high degree of accuracy.

The design was a 2 X 2 X 2 repeated measures factorial. Three main independent variables were manipulated. First, the input device utilized was Pausch et al.'s custom-made two-handed device containing a Polhemus tracker, or a 2-button Microsoft serial mouse, and this factor was run between subjects. Second, half of all displays contained the target item; half did not. Third, two sets of confusable letter sets were used: an angular set (AKMNVWXYZ) and a vertical set (EFHILT). Each letter set item was presented as a target twice—once for a target present trial and once for a target absent trial. Location of the target item was randomly chosen in the room every trial.

Procedure. All experiments were self-paced, the sequence of events being as follows: Subjects fixated their eyes on the target letter above the door on the front wall of the four-sided room (see Figure 1), and said the target letter out loud to facilitate its recall during the trial. Subjects indicated readiness to begin a trial by either depressing a mouse button or by rocking their foot forwards or backwards on the footpedal. After this, the room display with letters on all four walls, ceiling and floor was presented. Subjects were to search for the target item and respond to its presence or absence as quickly as possible by buttonpress or footpress, without making mistakes. An incorrect response was signaled by a discouraging audio cue, after which the target item for

the next trial was presented above the door on the front wall again. Sessions were blocked into two sets of either 18 or 12 trials, and at the end of each block, subjects were presented with a dialog box congratulating them on their progress. Subjects were encouraged to rest at the end of each block or at the beginning of each trial if needed.

As a precaution for the different amounts of training subjects have experienced between clicking with the mouse buttons and with the footpedal, mouse subjects were run twice through the study—once using the mouse to navigate but responding with the footpedal and once using the mouse only. Order was counterbalanced. No significant effects were found when responding with the footpedal versus the mouse buttons. Therefore, only the first training condition for mouse subjects (whether responding was carried out via footpedal or mouse-button) will be used in the analyses reported.

Results and Discussion

A 2 (input device) X 2 (letter set: angular vs. vertical) X 2 (target present or target absent) Analysis of Variance (ANOVA) with repeated measures was carried out on the reaction time data for all 32 subjects. Because there was evidence of a speed-accuracy tradeoff in the error data for the target present condition (average percent correct of 85% vs. 96% in the target absent condition), only correct trial data is used in the analysis of the reaction time data. Results showed significant main effects of letter set, $F(1,29)=21.23$, $p<.001$, and target presence, $F(1,29)=186.07$, $p<.001$. Average search times are presented in Table 1.

Device	Set 1 Target Present	Set 2 Target Present	Set 1 Target Absent	Set 2 Target Absent
Mouse	15.33 (6.1)	13.2 (6.8)	35.05 (11.8)	29.14 (12.04)
Two-handed Device	14.97 (4.3)	13.47 (5.3)	34.06 (9.3)	28.2 (8.9)
Total	15.15 (5.2)	13.3 (6.0)	34.57 (10.5)	28.7 (10.5)

Table 1. Average search times (in seconds) and standard deviations (in parentheses) for Experiment 1.

In addition, we saw the following significant results:

1. Average search times for target letter present were about two times faster than reported by Pausch et al. We believe this effect can be explained by the fact that the angle subtended by the display from the eye is about 30 degrees for the desktop display compared to 90 degrees for the HMD. The smaller visual angle means that the time to scan the display is significantly lower. Thus, we have observed a significant performance advantage when searching for information in Desktop VR (supporting others who have reported performance benefits for both Desktop and Fish Tank VR [10,18]).

2. Search times for target letters absent were almost exactly two times the average time for target present, as visual search theories predict. This suggests that the Pausch et al. findings do not apply to Desktop VR. It also suggests that research using visual search in 2D displays may generalize to Desktop VR.
3. We did observe rescanning behaviors about one-third of the time, both for target present and absent cases, for subjects that did not develop systematic search strategies. However, most of the rescanning was taking place on the floor and ceiling where no landmarks were present. From our study, it is not clear that rescanning time could be effectively used as a measure of immersion.
4. We observed no significant difference between the mouse and the two-handed input device.
5. One could argue that the balanced nature of this study made it possible for subjects to cheat. By noticing that each letter was present once and absent once, a subject could respond to the absent case faster once the letter had been seen. We analyzed the first versus second presentations of target letters and found no evidence of a performance advantage for the second presentation of a target letter.

PERIPHERAL LENSES: A NEW NAVIGATION AID

Desktop VR is often criticized for not providing a sense of where one is in the virtual space, or what is in the periphery, particularly when navigating through the space. One might ask how we are able to navigate through physical space and maintain such a good model of where we are. Several factors should be considered. First, we seldom magically appear somewhere in the real world, but rather we move into a location from another. As we enter a new space, we see what is in that space and continually build a mental model of that space and our position in it. When we are just placed in a new environment (like when landing at an airport never before visited), we initially experience a sense of disorientation that remains until we learn to rely on cues and landmarks to tell us where we are. For example, we see other people heading for baggage claim, or we see a sign pointing to baggage claim. Spaces in VR systems suffer from the magical appearance syndrome: they are seldom places into which we initially navigate [15]. Effective maps of a virtual space that are shown to a user before entering the space might help solve this problem.

Second, when driving a car, we use rear view mirrors to help us build and maintain a mental model of where other cars are with respect to us. Our use of the rear view mirror is like a flying spot scanner. It is periodically updated with a glance. It works well for spotting movement toward or away from us (like other cars rapidly

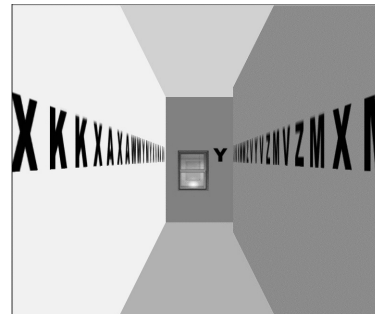


Figure 2: Navigation Study Hallway

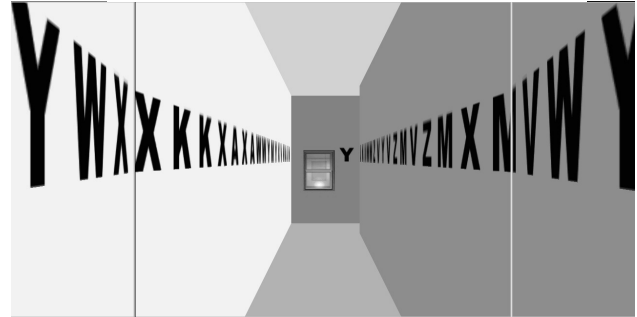


Figure 3: Navigation Hallway with Peripheral Lenses

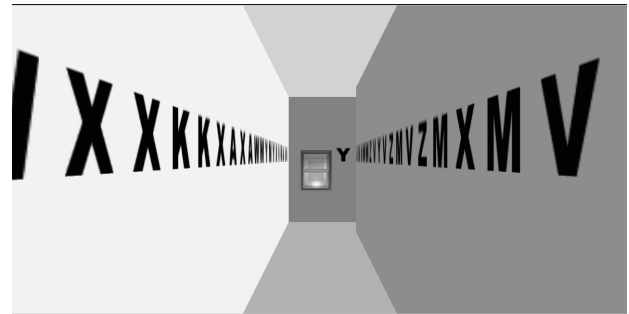


Figure 4: Navigation Hallway with wide aspect ratio

approaching). However, it does not really help in building a model of the space itself. We tried using virtual rear view mirrors as a navigation aid, but they added no significant value. We believe this is because in most VR systems, it is a model of the space rather than movement of objects in the space that is needed. Such mirrors might be useful for spotting movement of other users in Doom-style games or for finding other users moving around in shared virtual environments.

Third, when we move around in the real world, we make significant use of our peripheral vision. While we do not attend directly to objects in our periphery, we definitely know they are there. In Desktop VR and Fish Tank VR, we are missing these peripheral cues, while HMD VR does provide them to some extent.

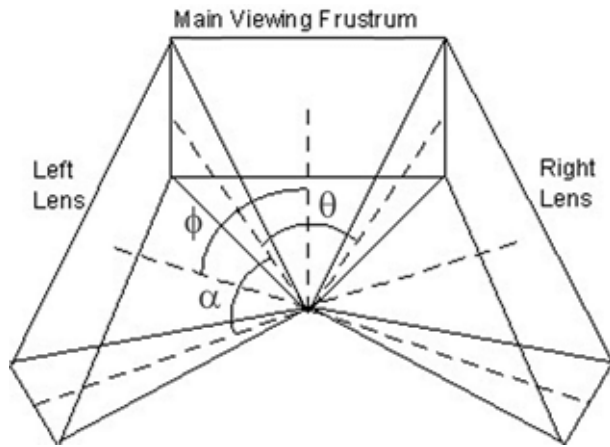


Figure 5: Peripheral Lens Configuration

Is it possible to simulate peripheral vision? We propose *Peripheral Lenses* as a virtual navigation aid for simulating peripheral vision. Peripheral Lenses are panels on either side of the main display. They represent viewpoints (or cameras) at the same origin as the main viewpoint, but looking off toward either side.

Figure 2 shows a virtual hallway without Peripheral Lenses. Figure 3 shows the same scene with Peripheral Lenses. Each lens is the same height as the main display, and 1/3 the width of the main display. Figure 5 shows the configuration of the lens. With the default vertical focal angle of 35 degrees and a default aspect ratio of 1.2:1, the horizontal focal angle (θ) is 42 degrees. The user is able to adjust the lens angle (ϕ) between 30 and 60 degrees, with default value of 42 degrees. As that angle is adjusted, the aspect ratio of each lens must be changed, so that the effective horizontal focal angle of the lens (α) is maintained as follows: $\alpha = 2\phi - \theta$. The lens aspect ratio is therefore the main display's aspect ratio times α/θ .

Do Peripheral Lenses help in a navigation task? Informally, they seem to help considerably while turning corners in a virtual space. One might expect to get the same added value by changing the aspect ratio of the window and using the same screen space with one window. In fact, this adds only one-third as much information as Peripheral Lenses do (note the two additional letters visible at the edge of each lens when comparing Figure 3 and 4). Also, one wide window does not assist the user in turning corners, while the lenses do, because the user can see around the corner. Another issue is the additional cost of rendering the scene in the two Peripheral Lenses. Displaying the lenses requires the scene to be rendered three times from different cameras, but does not actually cause a threefold increase in rendering computation because of the smaller pixel count in the lenses (the rendering is fill-limited). We observed roughly a factor of two slowdown with both lenses active. The next study addresses the issue of the value of

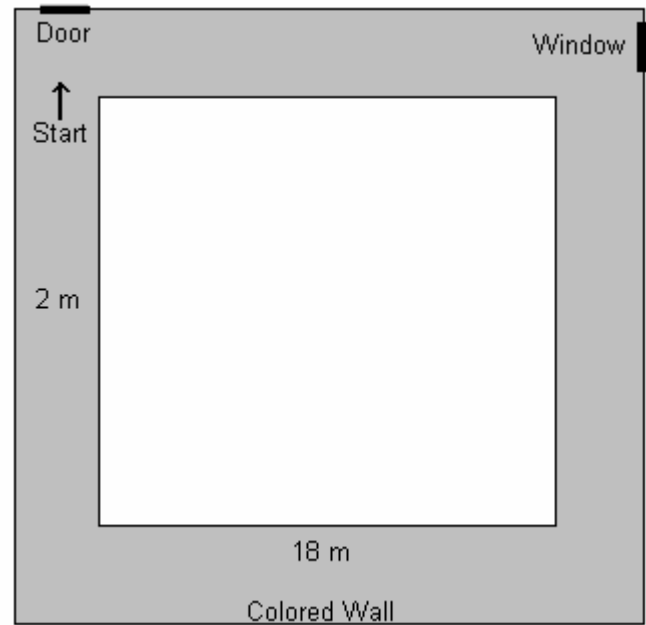


Figure 6: Top Down View of Hallways

Peripheral Lenses formally.

EXPERIMENT 2: NAVIGATION STUDY

In the second experiment, we wanted to understand the effects of navigation on the earlier results. We also wanted to introduce and evaluate *Peripheral Lenses*.

In this study, the mouse is used to control navigation around a set of virtual hallways (see Figure 2). The halls are 18 meters long, 2 meters wide, and 4 meters high. The 172 search letters are placed half way up the sides of the walls. Landmarks included a door, a window, and a wall of a significantly different color. Mouse movement forward and backward is used to move forward and backward, with the distance from the starting point (the center of the display) determining the speed. Mouse movement left and right controls turning left and right. In order to make movement through this space similar to movement through real space, the viewpoint is constrained to approach no closer than 18 inches from any wall. Note that this means that while turning a corner, the constraint implements a quarter-circle around the inside corner. Implementing these constraints provided a more natural feel to the navigation. Without these constraints the user could easily approach the wall too closely and lose orientation.

Experiment 2 used the same hardware and software base. It ran at about 37 frames per second without Peripheral Lenses, and about 18 frames per second with them.

Method

Subjects. Twenty-six subjects (twenty-four males) participated in this experiment. All subjects had normal or corrected-to-normal vision and were 18-50 years of

age. None of the subjects in Experiment 2 had participated in the first experiment.

Stimuli and design. Experiment 2 consisted of a visual search task in which the 3D room was converted into a series of connected hallways (see Figure 2), comprising a square when seen from the “top down” (see Figure 6). In this new display configuration, Peripheral Lenses either were or were not provided as a navigational aid (see Figure 3). Each Peripheral Lens was 200 by 500 pixels, subtending approximately 7.5 degrees of horizontal visual angle. The main window was 600 by 500 pixels (23 degrees of horizontal visual angle). When peripheral lenses were not included, the main window remained the same size. Only the mouse was used for navigating and responding in Experiment 2. Also, the more confusable letter set (letter set 1, angular) was the sole set of stimuli from which targets and distracters were chosen. Letters were the same size and are detailed in Experiment 1 above. Every item in the letter set was used two times as target present (in a random location) and two times as target absent. The design was therefore a 2 (Peripheral Lenses or not) x 2 (Target Presence) factorial with repeated measures.

Procedure. Experiment 2 was self-paced, the sequence of events being as follows: subjects fixated their eyes on the target letter above the door on the front wall of the first hallway. Subjects indicated readiness to begin a trial by depressing a mouse button. After this, the letters were displayed randomly on all of the hallways, lined up on both the left and right-hand walls about mid-way up the wall. Subjects were to search for the target item and respond to its presence or absence as quickly as possible by mouse button press without making mistakes. An incorrect response was signaled by a discouraging audio cue, after which the target item for the next trial was presented above the door on the first hallway again. Sessions were blocked into two sets of 18 trials, and at the end of each block, subjects were presented with a dialog box congratulating them on their progress. Subjects were encouraged to take rests at the end of each block or at the beginning of each trial if warranted.

Results and Discussion

A 2 (Peripheral Lenses present or not) X 2 (target present or target absent) Analysis of Variance (ANOVA) with repeated measures was carried out on the reaction time data for all 26 subjects. Due to evidence of a speed-accuracy tradeoff in the target present conditions (average percent correct of 78% vs. 98% in the target absent condition), only correct responses were included in the reaction time data analysis. Results showed significant main effects of target presence, $F(1,24)=266.27$, $p<.001$. No other main effects or interactions were significant. The results are shown in Table 2 below.

We observed the following effects in Experiment 2:

1. Search times were, on the average, slightly lower with Peripheral Lenses (average RT=32.16 seconds versus 35.29 seconds), although this result was not significant at the .05 level.
2. Search times were similar to those reported in Experiment 1, indicating that our findings about visual search tasks from the previous experiment also apply to a more realistic navigation task in Desktop VR.
3. Search ratios for target absent versus target present trials were approximately 2:1, as predicted by visual search theories and supporting point #2 above.

Condition	Target Present	Target Absent
Without Lenses	24.15 (4.06)	46.4 (9.56)
With Lenses	22.4 (7.6)	41.9 (8.3)
Total	23.3 (6.04)	44.1 (9.07)

Table 2. Average search times (in seconds) and standard deviations (in parentheses) for Experiment 2.

Our informal analysis of Peripheral Lenses suggested that much of the advantage was coming from increased ease in turning corners. To test this hypothesis, we ran a pilot study on corner turning behavior. In this study, the subjects were instructed to move about the four corridors of Figure 2 or Figure 3 as quickly as possible. Each trial was terminated automatically as soon as the subject turned the fourth corner. There was no search task, so presumably any difference could be attributed solely to corner turning speed. The results of the pilot study indicated no significant difference between the lens and no lens conditions, hence are not shown in detail.

EXPERIMENT 3: MODIFIED NAVIGATION STUDY

In Experiment 2 the aspect ratio was confounded between the lens and no lens conditions (the window for the no lens condition was smaller), and our subject population was fairly heterogeneous with respect to their experience with 3D environments. For those reasons, we repeated the experiment with the following changes:

1. The window viewing sizes were held constant between the lens and no lens conditions (compare Figure 3 with Figure 4).
2. Subjects were strictly screened so that no participants with 3D computer environment experience were evaluated.
3. Subject gender was balanced.
4. Target presentation was not fully counterbalanced. In other words, targets were chosen from the letter set in a completely random fashion, with replacement.

Experiment 3 used the same hardware and software base. It ran at about 25 frames per second without Peripheral Lenses, and about 18 frames per second with them.

Method

Subjects. Twenty-four subjects (11 males) participated in this experiment. All subjects had normal or corrected-to-normal vision, and ranged in age from 18-55. No subject in Experiment 3 had any experience with 3D user interfaces, or participated in the first two experiments.

Stimuli, design and procedure. Experiment 3 consisted of the same visual search task in Experiment 2. However, the size of the window in the no Peripheral Lens condition was altered to be as wide as the main window plus the two lenses (1000 by 500 pixels, subtending approximately 38 degrees of horizontal visual angle, see Figure 4). Also, subjects only carried out 26 trials (6 of which were considered training) in each lens condition. All other details of the experiment were the same. The design was again a 2 (Peripheral Lenses or not) x 2 (Target Presence) within subjects factorial with repeated measures.

Results and Discussion

A 2 (Peripheral Lenses present or not) X 2 (target present or target absent) Analysis of Variance (ANOVA) with repeated measures was carried out on the reaction time data for 23 subjects. (One subject was eliminated from the analysis because she used incorrect response buttons for a large percentage of one block of trials, which resulted in a large amount of missing data). Due to a speed-accuracy tradeoff in the target present condition (average percent correct of 84% vs. 92% in the target absent condition), only correct trial data was used in the reaction time analysis. Results showed significant main effects of lens condition, $F(1,22)=8.04$, $p<.01$, and target presence, $F(1,22)=175.47$, $p<.001$. The interaction was not significant. The results are shown in Table 3 below.

We observed the following effects in Experiment 3:

1. Search times were, on the average, slightly slower with Peripheral Lenses (average RT=32.16 seconds versus 35.29 seconds), and this result was significant.
2. Once again we observed target absent search times to be at least twice as slow as target present search times, although the ratio was slightly higher than 2:1 in this study. The higher ratio was still small compared to the ratio seen in the Pausch et al. study.
3. Target present results without lenses are very similar to the results in Experiment 2, suggesting that the wide aspect ratio of Figure 4 did not make a noticeable difference.

Condition	Target Present	Target Absent
Without Lenses	22.95 (5.91)	51.2 (15.1)
With Lenses	26.7 (8.1)	59.7 (19.0)
Total	24.8 (7.3)	55.5 (17.5)

Table 3. Average search times (in seconds) and standard deviations (in parentheses) for Experiment 3.

CONCLUSIONS

Proponents of VR have held that immersion is critical to VR and can best be attained by fully visually immersing the user with an HMD or CAVE, and by driving the user's virtual viewpoint by tracking the user's head movements. Some have argued that Fish Tank VR is just as immersive, by using desktop stereo displays and head tracking [20]. A few have argued that immersion is also possible in Desktop VR [10], using desktop displays and driving the virtual viewpoint with more common input devices, like the mouse. In this paper, we explored methods for measuring or observing immersion in Desktop VR, and enhancing it with a new navigation aid called Peripheral Lenses.

Pausch et al. [9] purported to quantify immersion in VR using a visual search task. We extended that study in several ways. First, we attempted to apply the result to Desktop VR. Our results suggest that the findings of Pausch et al. do not apply to Desktop VR, because we observed predicted times for target absent conditions instead of the slower times that Pausch et al. observed. From this it follows that users of Desktop VR need not suffer a worse sense of space or a lower level of immersion. We observed much faster search times than Pausch et al. We believe this can be explained because the desktop display uses a smaller visual angle, thus is faster to scan. This suggests a significant advantage for Fish Tank VR and Desktop VR over HMD VR. We also observed that the search times were the same for a mouse driven interface compared to the two-handed input device used by Pausch et al. We did observe some rescanning behavior as reported in Pausch et al., but only in areas without landmarks, and with no significant increase in search times overall.

Second, we extended the study to a more realistic navigation task. We found that search times for the target absent condition matched the predicted search times, as in the first experiment. Future studies should take our results for navigation and apply them to HMD VR. We also gained support for the generalizability of the large body of knowledge about 2D visual search and attention into Desktop VR domains.

Third, we introduced Peripheral Lenses as a technique for simulating peripheral vision in Desktop VR. Our intuition and informal studies indicated that Peripheral Lenses would make certain navigation tasks (e.g., turning corners) considerably easier. However, the formal studies do not support this. Experiment 2 shows that search times were slight faster with Peripheral Lenses, but the results were not statistically significant. Experiment 3 shows statistically significantly slower results with Peripheral Lenses, for subjects with no experience with 3D interfaces and for a balanced gender mix. Further studies are needed to understand exactly when Peripheral Lenses

are effective. The fact that the formal studies contradict the informal studies confirms the importance of formal usability studies in the development of user interface software.

We believe that visual search tasks are a good method for evaluating navigation aids intended to increase immersion, and for comparing different approaches to VR.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the help of Randy Pausch and George Williams in understanding the details of their earlier study. In particular, we would like to thank them for loaning us the two-handing input device so that we could come as close as possible to replicating their study. We would like to acknowledge the help of Dan Robbins for discussions about rear view mirrors and the Peripheral Lens. We would like to offer special thanks to Kevin Larson and Nancy Mickel for their assistance in running the experiments.

REFERENCES

1. Boyd, C. (1997), Does immersion make a virtual environment more usable?, CHI'97 Conference Companion, 325-326.
2. Carter, R. C. (1982). Visual Search with Color. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 127-136
3. Cruz-Neira, C, Sandin, T., DeFanti, R (1993), Surround screen projection-based virtual reality: the design and implementation of the cave. *Proceedings of SIGGRAPH'93*, 135-142.
4. Csikszentmihalyi, M. & Kubey, R. (1981), Television and the rest of life: A systematic comparison of subjective experience. *Public Opinion Quarterly*, 45, 317-328.
5. Deering, M. (1992), High resolution virtual reality, in *Computer Graphics*, 26, 2, 195-202.
6. Dutta, A., Walker, B.N., Czerwinski, M.P. & Feldman, E.M. (1996). Some fundamentals of training and transfer: Practice benefits are not automatic. Chapter in Dutta, A. & Quinones, M.A. (Eds.), *Training for 21st Century Technology: Applications of Psychological Research*, APA Books.
7. Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 32-39.
8. Foley, J. (1987), *Interfaces for advanced computing*, Scientific American, October, 1987.
9. Pausch, R., Proffitt, D., and Williams, G. (1997), Quantifying immersion in virtual reality, SIGGRAPH'97 (to appear).
10. Robertson, G., Card, S., and Mackinlay, J.(1993), Nonimmersive virtual reality, *IEEE Computer*, Feb. 1993.
11. Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1-66.
12. Shiffrin, R. M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127-190.
13. Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist*, 57, 421-457.
14. Sutherland, I. (1969), A head-mounted three dimensional display, Fall Joint Computer Conferences, AFIPS Conference Proceedings, 33, 757-764.
15. Thorndyke, P. W. & Hayes-Roth, B. (1982). Differences in spatial knowledge acquired through maps and navigation. *Cognitive Psychology*, 14, 560-589.
16. Townsend, J.T. (1990). Serial vs. parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science*, 1, 46-54.
17. Townsend, J.T. & Ashby, F.G. (1983). *The stochastic modeling of elementary psychological processes*. NY: Cambridge.
18. Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 194-214.
19. Triesman, A. M. & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97-136.
20. Ware, C., Arthur, K., and Booth, K.(1993), Fish tank virtual reality, CHI'93 Proceedings, 37-42.