Pacific Graphics (2021)
M. Okabe, S. Lee, B. Wuensche and S. Zollmann (Editors)

Peripheral vision in simulated driving: comparing CAVE and head-mounted display

T. Tanoi and N. A. Dodgson

supplementary material — extended version of poster text

Abstract

Peripheral vision is widely thought to be important but is not provided in the majority of head-mounted displays (HMD). We investigate whether peripheral vision is important in a simulated driving task. Our hypothesis is that subjects will be able to complete the task more quickly if they use their peripheral vision. We compared subject performance in a CAVE environment, with 270° *field-of-view (so automatic peripheral vision) and in a HMD, with* 110° *field-of-view (so no peripheral vision but the ability to turn the head). Our results show almost no statistically significant differences between the two conditions. This contrasts with the opinions of our subjects: our expert users, in early tests, commented that peripheral vision helped in the task and the majority of our naïve subjects believed that the lack of peripheral vision in the HMD hindered them in the task.*

CCS Concepts

• ***Computing methodologies*** → *Virtual reality;* • ***Software and its engineering*** → *Virtual worlds software;*

1. Introduction

We compare a virtual reality head-mounted display (HMD) against a Cave Automatic Virtual Environment (CAVE) in a driving task to ascertain both the importance of peripheral vision in the task and the subjects' perception of the differences between the HMD and the CAVE. HMDs have limited or no peripheral vision and this has been thought to limit their attractiveness and usefulness sufficiently that researchers are developing wide field-of-view HMDs (e.g., StarVR's 210° FoV) or retro-fitting peripheral vision to existing HMDs [XB16]. We wished to investigate whether peripheral vision made any significant difference in a practical, simulated task. Our hypothesis is that the wider peripheral vision in the CAVE environment will produce better performance than the HMD. Our experimental results show that there was almost no significant difference in performance of the task between the two conditions, despite the majority of the subjects believing that the lack of peripheral vision in the HMD had hindered their performance.

Virtual reality HMDs have been in use for decades (Section 2), becoming popular and widely accessible in the last five years. However, they have drawbacks: the user is isolated from the real world, is disconnected from seeing their own body, and lacks peripheral vision. This work compares an HMD against a multi-screen CAVE, which suffers from none of these three drawbacks. We designed an experimental condition that could test our hypothesis in a task that had real-world application. We use a driving simulator task where we modify the visual display (between CAVE and HMD) but aim to keep as much of the rest of the environment as similar as possible to avoid confounding factors. For example, the user has a tactile interface (seat, steering wheel, pedals) that means that it can be used without the need for the user to see it, that is identical in both scenarios and that is familiar to anyone who has driven a car.

We used a driving simulator system (Section 3) to provide a common, static interface that gives the user tactile feedback and in which it is not necessary for the user to turn their body. The latter constraint is because our CAVE installation has no rear wall and therefore it is not possible to allow the user to turn their body without losing the illusion.

We put considerable thought into finding a task that would work within a driving simulator and that would use peripheral vision. The task (Section 4) was to drive around a computer-rendered city, searching for and collecting objects: we implemented a cityscape version of Pacman, with the driver having to find and pick up all of the pac-dots. Pac-dots are collected simply by the driver driving past them. These pac-dots are located on all streets within the city. The buildings are sufficiently tall that the driver cannot easily see what is down a side-street until they are at a junction. The intention is that the driver will need to be aware of what is down the side streets and that this will require them either to glance down side streets as they drive past or use their peripheral vision to identify that there is a pac-dot down a side street that remains to be collected.

In designing and refining the task, we conducted early informal tests (Section 5) with our computer graphics research team, who were aware of the aim of the project, to ensure that peripheral vision was a helpful factor. Once we were happy with the task, we then ran the actual experiment on naïve subjects (Section 6), who were unaware of the aim of the project. We present the formal results of the experiment (Section 7) and discuss their implications (Section 8).

2. Background and Related Work

Experimental virtual reality systems have existed since Ivan Sutherland's seminal work in the 1960s [Sut68]. More recently, there has been a flurry of interest as inexpensive, good quality, head-mounted displays (HMDs) have come onto the market.

Most commercially-available HMDs, such as the HTC Vive and Oculus Rift, have fields-of-view of up to 110°. The StarVR One is an exception, claiming up to 210° of vision. The human visual system has a horizontal field-of-view that is over 180°. The question is whether that extra peripheral vision helps in actual tasks or is simply there to provide context. While we do not attend to objects in our periphery, we are certainly aware of them [RCVD97]. As we navigate the world, our peripheral vision helps us to build and maintain a mental 3D model of the world. In our experiment we developed a task where it was important to build a model of the world, in order to keep track of objects yet to be collected, and where peripheral vision thus should be beneficial.

With regard to earlier work on whether peripheral vision is useful, we were inspired by Xiao and Benko, who extended an HMD to include a low-resolution peripheral display [XB16]. They found that the peripheral vision helped in a search task in the HMD.

There is considerable previous work that compares HMD and desktop [RPJ99, SSDP*09] and CAVE and desktop [DJK*06, PFK*08]. The results appear to depend on the task . For example, Prabhat et al. found that subjects preferred and performed better in the CAVE than in fishtank VR [PFK*08], while Demiralp et al. found the opposite [DJK*06]. Ruddle et al. found that subjects could navigate more quickly in an HMD than on the desktop [RPJ99], while Sousa Santos et al. found that subjects performed better with their desktop setup than in an HMD [SSDP*09].

In all work of the nature, it is challenging to get good comparisons between different display and interaction modalities. An example is Pausch et al.'s early study into the effectiveness of VR (comparing HMD against desktop-like display) where they controlled for the differences between the two display types by using the HMD to emulate a desktop display by bolting the HMD in place for the desktop context [PPW97]. Unfortunately, this led to a difference in interaction modality, where they provided an unusual 6-DoF manipulation device in the desktop context, the novelty of which may have affected their results [RCVD97]. In our experiment we aimed to ensure a common user interface between tasks, while acknowledging that the displays are different.

3. The Simulator

We created a driving simulator in Unity, gamified to provide a task which makes use of the user's peripheral vision. The simulator has five principal components: accepting input from a steering wheel and pedals, outputting to three different display contexts (HMD, CAVE, and a single monitor), a driving simulator, a city generator and renderer, and game mechanics to provide a task.

Equipment We chose to use a physical steering wheel, pedals and seat (Thrustmaster T500RS) to provide a common tactile physical interface between the different contexts. This would minimise any difference owing to different interfaces or to being unable to see the physical interface when in the HMD. Further, it provides an interface familiar to anyone who has driven a car and therefore we avoid problems with subjects who are unfamiliar with computer game control pads or VR controllers. We note that real drivers rarely look at their hands on the steering wheel and never look at their feet on the pedals therefore, once the subject is seated correctly, they should be able to control the virtual car as effectively in the HMD as in the CAVE despite the lack of visual feedback.

Display Contexts We required that the same visuals could be rendered in a HMD, in the CAVE, and on a single monitor. Our HMD was the HTC Vive (110° FoV). The HMD requires rendering the same scene twice (once for each eye) from a dynamically updating camera. Our CAVE was a 5 m × 5 m space, with the Thrustmaster setup centred in the space. There were three screens, each 2 m × 5 m with the imagery centred at eye height. There was no floor projection, which emulates being within a car (where the body of the car excludes being able to see the nearby ground plane). The CAVE requires rendering the scene from three cameras: one for each of left, front, and right screen (Figure 1). The single monitor requires rendering just a single scene, which we made identical to the front screen of the CAVE. The HMD is stereoscopic, with the other modalities being monoscopic. Given that the scenery is several metres from the viewer, stereoscopy was not considered to make any substantial difference. The HMD is head-tracked but there is no head-tracking in the CAVE: the viewer's head moves very little in position, owing to the viewer being seated and rotational movement is, of course, handled naturally by the CAVE.

Driving Physics We used an existing driving simulator, built for Unity, which handled acceleration, braking, collision detection, and generation of appropriate sound effects.

City Generator We wrote a city generator that created plausible looking buildings on a grid structure. The purpose of the city buildings is to occlude pac-dots from the user, forcing them either to use peripheral vision or to look to the side while driving. We needed multiple road networks to prevent the user memorising or predicting the layout, thus forcing them to rely on observation and peripheral vision. We used a simple input file format, where a small image file (which could be edited in any image editor) used specific colours to indicate which grid square contained buildings, empty road segments, road segments containing pac-dots, the car's starting position, and other features.

Gamification We gamified the driving simulator to provide a task for subjects to undertake. In this case, collecting objects lying in the street (akin to a 3D version of Pacman), which forces the user to look and drive down side streets to check that they have collected all the objects. Pac-dots were rendered as semi-transparent golden spheres, with a glittering arch above them to make it obvious where they are placed. This visual effect was developed through early tests in which we refined the presentation with the aim that it must be

Figure 1: *Equipment setup. Left: HMD. Right: CAVE.*

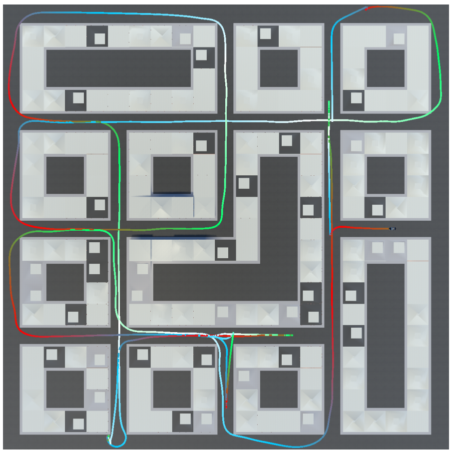

Figure 2: *One of the city layouts with the path taken by one of the subjects (colour changes indicate time cycles from start). Pac-dots are distributed only inside the grid, the outer ring road is free of pac-dots. Notice that this driver has exhibited two different behaviours: driving round the block to find new dots (left and top) and doing U-turns to go back to pick up dots (bottom).*

obvious in peripheral vision that there is something to collect. Pac-dots were collected by driving through or past them.

4. The Experimental Task

The task is to collect a series of 40 pac-dots uniformly distributed throughout the grid based city. The motivation behind the task is that the user will benefit from looking to their left and right as they come to junctions, in order to complete the task more quickly, but that the task can be accomplished without explicitly requiring this. Our hypothesis is that the subject can make better time in the task when they have peripheral vision (so do not have to turn their head explicitly to spot uncollected pac-dots) than when they have to remember to turn their head whenever they reach a junction.

We created two city layouts, A and B, that were designed to be as similar as possible, without being identical. Two layouts were needed to avoid learning the layout from one condition to the second. Each layout had 4×4 blocks, with some blocks combined to make double-sized blocks and L-shaped blocks (Figure 2). These combinations are to make it more challenging for the subject to find all the pac-dots and so that most streets do not extend across the full grid. Pac-dots are placed only in the interior of the grid, with the road round the outside free of dots.

5. Observations from Early Tests

We ran early tests on members of our computer graphics research group, all of whom were aware of the aim of the experiment. These early tests helped us to tune the maximum speed of the car, the length and width of the roads, the behaviour as the user left the 4×4 grid, the visual presentation of the pac-dots, and the level of discomfort in both conditions. Early tests were nauseating in both conditions but especially in the HMD. We corrected for this (slower car speed, better steering, increase the frame rate to 90fps) so that, in the actual experimental trials, none of our experimental subjects had to abort the experiment owing to discomfort and only one reported strong discomfort after their trial.

In the early tests, the expert members of the research group noticed that the peripheral vision afforded by the CAVE gave them an apparent advantage in completing the task. As they drove through a crossroads at full speed, they found it easier to spot uncollected pac-dots in peripheral vision in the CAVE than to whip the head from side to side to look down both side streets in the HMD. We took this to indicate that peripheral vision allows implicit memorisation of the location of uncollected pac-dots, leading to quicker completion times over the HMD, where you had instead to explicitly look down the side streets as you passed. These early informal observations indicate that the experimental task is a suitable test of our hypothesis.

Experts also noticed that, in either condition, it was often challenging to find the final two or four pac-dots owing to them being hidden down side-streets that had not been passed recently. Therefore we decided that it was important to measure how long it took to collect 90% of the pac-dots rather than 100%. In the experimental trials, we recorded time and distance to collect every one of the forty pac-dots, and did statistical analysis on time and distance taken to collect 90% (36 dots), 95% (38 dots) and 100% (40 dots).

6. Experiment

We recruited 17 naïve subjects (8 female, 9 male; age 18–40, average 22, std. dev. 5.5). Eight had no VR experience, 7 some, and 2 frequent use. Seven had no experience in a driving simulator before, with 10 having some experience. Fifteen had real-world driving experience. One had no gaming experience, 9 some, and 7 frequent.

Each subject was asked to complete five components to the experiment: a pre-test survey about their past experiences with driving, virtual reality, and gaming; an initial trial using only a single screen to familiarize them with the scenario; two recorded trials in first either the HMD or CAVE, followed by the alternate option; and finally a post-test interview about their perception of their performance as well as their perceived differences and discomfort.

A trial entails the user completing the previously described task of collecting the pac-dots over the city. The purpose of the single screen trial is two-fold: firstly to give the user a base line to compare their later experiences, and secondly to allow them to become familiar and comfortable with the driving simulator, thus preventing large discrepancies between first- and second-run results. We used a third city layout for this familiarisation trial to avoid subjects learning the layout of the virtual cities used in the recorded trials. The condition HMD/CAVE was changed within-subjects. We alternated each subject between using the HMD or CAVE first, which mitigated against a learning effect. We also alternated between using Layout A or Layout B in the HMD, with the other in the CAVE, to mitigate against one layout being easier than the other.

We recorded a variety of data including: time taken, position at 0.2 second intervals, time when pac-dots were obtained, and direction and duration of significant (45° away from the car's forward direction) head turns. These were automatically recorded in the HMD and manually observed in the CAVE using video recordings of the subject. From this data we can also calculate related data such as distance travelled to collect each item, and difference between a subject's first and second trial for time and distance (the primary metric used in the results).

After their trials, we interviewed each subject with questions that increasingly revealed information about the experiment's aim, to tease out their opinions of their experience (Table 1).

7. Results

We measured time taken and distance driven to collect the forty pac-dots. Our analysis is conducted on the time and distance required to collect 90% (36 dots), 95% (38 dots) and 100% (40 dots). Figures 3 and 4 show the distance driven and time taken for each of these, with multiple presentations of the same data split between the two contexts (HMD and CAVE), the two runs (first and second) and the two city layouts (A and B). As indicated by our early testers, the results confirm that there is a noticeable extra distance required to find the final two pac-dots (Figure 3).

	CAVE	HMD	Equal
In which context did you feel your performance was better?	9	6	2
	CAVE	HMD	Neither
Did you find your peripheral vision was limited in either context?	5	9	3
	CAVE	HMD	Neither
Which context did you find most comfortable?	12	2	3
	Yes	No	
The HMD has 110° field of view, but CAVE has 270° field of view. Do you feel this impacted your performance?	10	7	

Table 1: *Summary of responses to post-test interview questions. The questions were asked in the order shown, revealing increasing amounts of information about the purpose of the experiment.*

We ran a repeated measures three-way ANOVA on this data to see whether any of these distance or time measures (Figures 3 and 4) was significantly influenced by context, run and layout or any interaction of these. We removed outliers before running the ANOVA. For five of the six measures the ANOVA showed that *none* of these results is significant. For example, for 'Time to 36', there is no significant influence of context ($F(1,21) = 1.386$, $p = 0.252$), run ($F(1,21) = 0.139$, $p = 0.713$) or layout ($F(1,21) = 0.042$, $p = 0.840$). However, for 'Distance to 36', there is a significant influence of context ($F(1,21) = 5.253$, $p < 0.05$) but not of run ($F(1,21) = 0.358$, $p = 0.556$) or layout ($F(1,21) = 0.875$, $p = 0.360$). A post-hoc t-test on context shows that there is a significant difference at the 95% confidence level and that the HMD context has a shorter mean distance than the CAVE context.

We also measured the number of head turns made by subjects. This was done automatically in the HMD (checked by a manual count) and by manually counting head turns on a video of the subject taken in the CAVE context. Head turns are graphed in Figure 5. There is no significant difference between the number of head turns in any of context, run or layout. One subject failed to turn their head at all in either context. Having (CAVE) or not having (HMD) peripheral vision did not significantly influence the number of head turns made by subjects, with there being strong correlation ($r = 0.908$) between head turns in the two different contexts.

We conducted a post-test interview in which we asked a sequence of increasingly revealing questions about their experience. Table 1 summarises the subject's answers to the questions. Three of the subjects made unsolicited comments about the lack of peripheral vision in the HMD, when asked the first question "In which context did you feel your performance was better? And why?" Three other subjects made comments about the corners of the CAVE causing problems with their ability to complete the task: the issue appeared to be the rendering onto three walls rather than the continuous rendering of the HMD. We have anecdotal evidence from another CAVE installation that corners can be problematic and that users prefer a cylindrical screen. Finally, although only one subject reported strong discomfort from the HMD after their trial, six subjects reported mild discomfort: either that the HMD was uncomfortable or heavy or that it caused had mild nausea.

Notice, in Table 1, that five subjects commented that their peripheral vision was hindered in the CAVE rather than in the HMD.

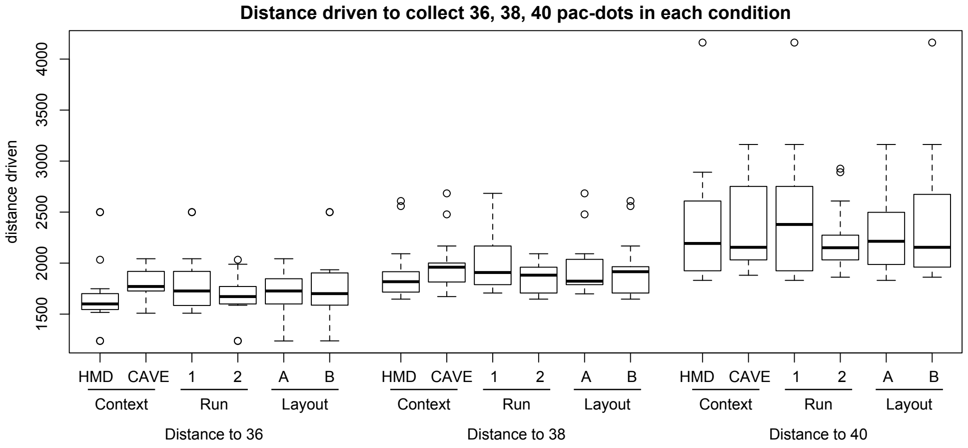

Figure 3: *Box plots of distance driven by subjects.*

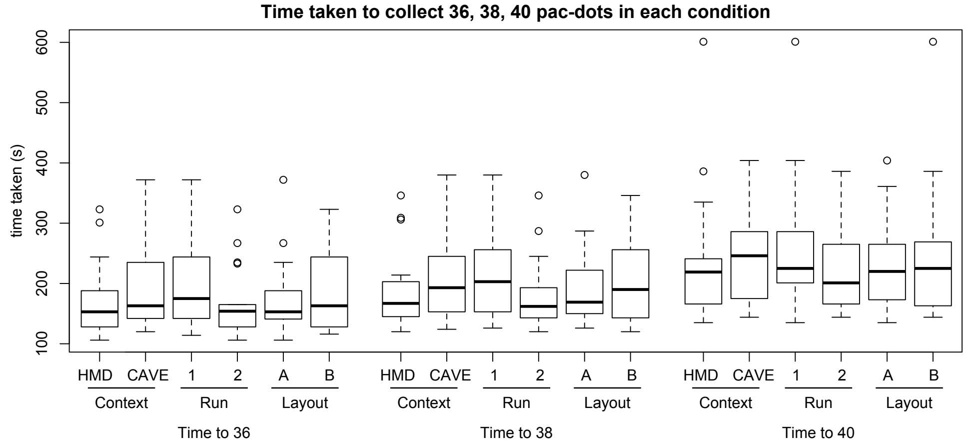

Figure 4: *Box plots of time taken by subjects.*

Some of these subjects then made a contradictory response to the follow-up question, saying that the smaller field of view of the HMD impacted on their performance.

8. Discussion

It is disappointing that the experiment produced only one statistically significant result. That result indicates that the HMD allowed participants to complete 90% of the task in shorter distance (but not shorter time) than in the CAVE.

What can we conclude from this? There are several possibilities. (1) There is no significant effect *generally*. That is, peripheral vision does not significantly improve a human's ability to perform tasks. (2) There is no significant effect for this particular task. (3) There is an effect but this task's confounding factors overwhelm the effect. We have insufficient data to distinguish between these possibilities, though there is likely good enough other evidence to reject option (1) [SRJ11].

With regard to (2), it is quite possible that, for this task, the extra benefit of peripheral vision makes little difference. The fact that there was no significant difference in the number of head-turns between the two contexts offers support for this conclusion. We constructed the task so that peripheral vision should have been useful. Indeed, our expert testers informally thought that peripheral vision was helping them and most of our naïve subjects thought that the lack of peripheral vision impaired their performance, once they were made aware of it. A new hypothesis, therefore, is that people can use their peripheral vision effectively in tasks if they are made aware that it is a useful thing to do (cf. training of the peripheral visual system [SRJ11]).

With regard to (3), there are a range of possible confounding factors. One factor is that many subjects simply took the car to maximum speed and then drove round the city collecting pac-dots, as you would if playing a game without peripheral vision being available. Few subjects made U-turns and informal analysis in the simulator indicates that there is a feeling of a considerable time penalty to turning around rather than driving round the block. Longer roads

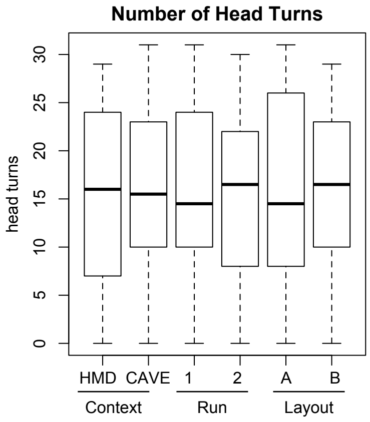

Figure 5: *Box plots of number of head turns made by subjects.*

may mitigate this problem and this is something that we discussed before running the experiment but abandoned on the grounds that it would make the trials much longer and we were concerned about nausea on long trials.

The HMD rendered the imagery in stereo, while the CAVE was monoscopic. This is unlikely to have a significant impact on results as stereopsis is of limited use in driving [BDK*01], owing to the majority of the scene being at more than a couple of metres from the subject.

The HMD is a sufficiently different experience to the CAVE that this could confound the peripheral vision results. The HMD isolates the subject from the real world. We also note that, in the HMD, we are able to render a realistic interior to the car, which is not possible in the CAVE. We had to make a decision of whether to draw the interior of the car projected onto the CAVE's screens (our final implementation does this) or to have the CAVE driver feel that they were essentially on an open trolley driving round the streets. An alternative approach for checking the importance of peripheral vision would be to do all of the experiments in the CAVE but to make the user wear a face mask that limits their peripheral vision to 110° for one of the contexts.

We subsequently developed two alternative tasks but were never able to convincingly remove everything that could be a confounding factor.

9. Conclusion

Having constructed a task in which peripheral vision should have been helpful, we found that the only statistically-significant difference between the 270° field-of-view CAVE context and the 110° field-of-view HMD context indicated that the HMD context was superior.

If we are to tease out whether peripheral vision is important in driving, then further tests need to be run, which may involve the following:

- running the same experiment modified to make U-turns more time-efficient than driving round the block;
- running a similar experiment entirely in the CAVE environment with a face mask to modify the subject's field-of-view;
- designing alternative experiments that dispense with the driving simulation and concentrate on artificial tasks designed to test use of head-turns and peripheral vision;
- running experiments in which subjects are randomly briefed or not briefed on the use of peripheral vision in the task, to see whether briefing affects the ability to use peripheral vision.

Acknowledgements

We thank *anonymous support staff* for their assistance in getting the equipment installed and operational; the computer graphics research group, especially *Anonymous* for useful discussions and for the early debugging tests; and all of our experimental subjects.

References

[BDK*01] BAUER A., DIETZ K., KOLLING G., HART W., SCHIEFER U.: The relevance of stereopsis for motorists: a pilot study. *Graefe's Archive for Clinical and Experimental Ophthalmology 239*, 6 (Jul 2001), 400–406. URL: `https://doi.org/10.1007/s004170100273`, `doi:10.1007/s004170100273`. 6

[DJK*06] DEMIRALP C., JACKSON C. D., KARELITZ D. B., ZHANG S., LAIDLAW D. H.: CAVE and fishtank virtual-reality displays: A qualitative and quantitative comparison. *IEEE Transactions on Visualization and Computer Graphics 12*, 3 (2006), 323–330. 2

[PFK*08] PRABHAT, FORSBERG A., KATZOURIN M., WHARTON K., SLATER M., ET AL.: A comparative study of desktop, fishtank, and CAVE systems for the exploration of volume rendered confocal data sets. *IEEE Transactions on Visualization and Computer Graphics 14*, 3 (2008), 551–563. 2

[PPW97] PAUSCH R., PROFFITT D., WILLIAMS G.: Quantifying immersion in virtual reality. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1997), SIGGRAPH '97, ACM Press/Addison-Wesley Publishing Co., pp. 13–18. URL: `http://dx.doi.org/10.1145/258734.258744`, `doi:10.1145/258734.258744`. 2

[RCVD97] ROBERTSON G., CZERWINSKI M., VAN DANTZICH M.: Immersion in desktop virtual reality. In *Proceedings of the 10th annual ACM Symposium on User Interface Software and Technology (UIST)* (1997), ACM, pp. 11–19. 2

[RPJ99] RUDDLE R. A., PAYNE S. J., JONES D. M.: Navigating large-scale virtual environments: what differences occur between helmet-mounted and desk-top displays? *Presence: Teleoperators and Virtual Environments 8*, 2 (1999), 157–168. 2

[SRJ11] STRASBURGER H., RENTSCHLER I., JÜTTNER M.: Peripheral vision and pattern recognition: A review. *Journal of vision 11*, 5 (2011), 13–13. 5

[SSDP*09] SOUSA SANTOS B., DIAS P., PIMENTEL A., BAGGERMAN J.-W., FERREIRA C., SILVA S., MADEIRA J.: Head-mounted display versus desktop for 3d navigation in virtual reality: a user study. *Multimedia Tools and Applications 41*, 1 (2009), 161. 2

[Sut68] SUTHERLAND I. E.: A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I* (New York, NY, USA, 1968), AFIPS '68 (Fall, part I), ACM, pp. 757–764. URL: `http://doi.acm.org/10.1145/1476589.1476686`, `doi:10.1145/1476589.1476686`. 2

[XB16] XIAO R., BENKO H.: Augmenting the field-of-view of head-mounted displays with sparse peripheral displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (2016), ACM, pp. 1221–1232. 1, 2