



Pano: Design and Evaluation of a 360° Through-the-Lens Technique

Damien Clergeaud, Pascal Guitton

► To cite this version:

Damien Clergeaud, Pascal Guitton. Pano: Design and Evaluation of a 360° Through-the-Lens Technique. 2017 IEEE Symposium on 3D User Interfaces (3DUI), Mar 2017, Los Angeles, United States. pp.2 - 11, 10.1109/3DUI.2017.7893311 . hal-01505260

HAL Id: hal-01505260

<https://hal.science/hal-01505260>

Submitted on 11 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Pano: Design and Evaluation of a 360° Through-the-Lens Technique

Damien Clergeaud*
Inria
Université de Bordeaux - France

Pascal Guitton †
Université de Bordeaux - France
Inria

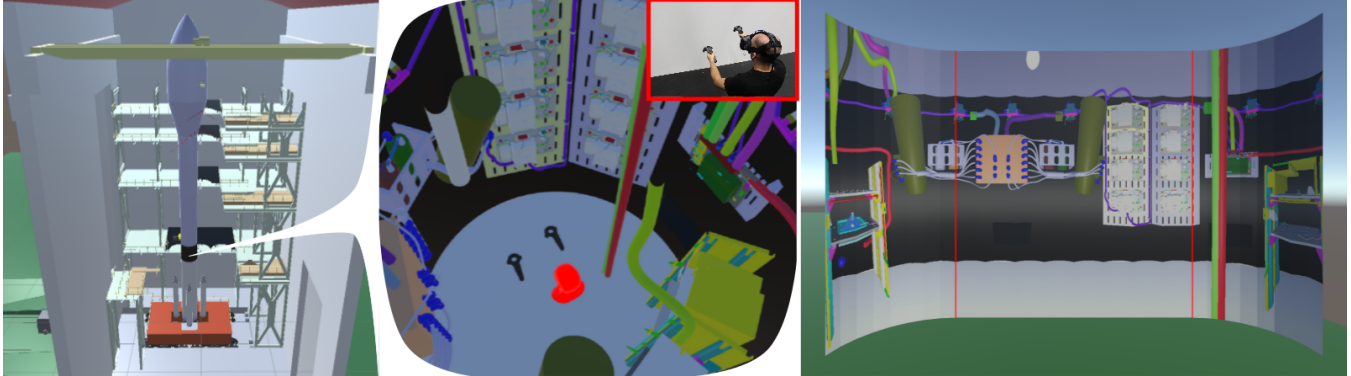


Figure 1: **Left:** Outside view of a launcher. **Middle:** In the little black cylinder of this launcher, an operator is working in virtual reality with an HMD. The red virtual object represents his avatar. **Right:** At the same time, remotely in the virtual environment, a second operator is using Pano to interact with the first operator. Thanks to Pano, he can see the entire interior of the launcher.

ABSTRACT

Virtual Reality experiments have enabled immersed users to perform virtual tasks in a Virtual Environment (VE). Before beginning a task, however, users must locate and select the different objects they will need in the VE. This first step is important and affects global performance in the virtual task. If a user takes too long to locate and select an object, the duration of the task is increased. Moreover, both the comfort and global efficiency of users deteriorate as search and selection time increase.

We have developed Pano, a technique which reduces this time by increasing the users natural field of view. More precisely, we provide a 360 panoramic virtual image which is displayed on a specific window, called the PanoWindow. Thanks to the PanoWindow, users can perceive and interact with the part of the Virtual Environment (VE) that is behind them without any additional head or body movement. In this paper, we present two user studies with 30 and 21 participants in different VEs. For the first study, participants were invited to perform object-finding tasks with and without Pano. The second study involved position-estimating tasks in order to know if the PanoWindow image enables users to build an accurate representation of the environment. First, the results show that Pano both reduces task duration and improves user comfort. Second, they demonstrate that good object-localization accuracy can be achieved using Pano.

Index Terms: K.6.1 [Management of Computing and Information Systems]: Project and People Management—Life Cycle; K.7.m [The Computing Profession]: Miscellaneous—Ethics

*e-mail: damien.clergeaud@inria.fr

†e-mail: pascal.guitton@inria.fr

1 INTRODUCTION

Virtual reality has long been used for different purposes (designing, training, understanding, etc.) and in different fields (industry, medicine, archeology, architecture, etc.). There is a common point between all these contexts: immersed users must first locate and select the objects they will need in the Virtual Environment (VE) before performing their virtual tasks. Once the objects have been identified in the 3D scene, users can then go on to perform the process at hand. The speed and comfort with which users perform this initial localization step are therefore essential to the quality and performance of the overall process since this step is often repeated throughout a VR session.

Different factors may hinder this performance. The first is the nature of the data in the VE. Searching for an object may be difficult because there are many similar objects and/or because the object is intrinsically complex. Imagine for instance VR experiments inside a space launcher or a manufacturing plant. Even if you know the environment, it may be time consuming to precisely locate a specific component. Moreover, if you are working with [visual representations of] abstract data (for instance financial or demographic information), you do not have any ecological landmark and the search time is therefore lengthened.

The second factor that may hinder performance is object occlusion. Two cases must be distinguished: objects placed in front of the users (within their field of view) and those placed behind them. In the first case, objects can be hidden if there is a barrier between the user and the object being searched for. Techniques have been developed which enable users to see "through" the barrier, using ray tracing and transparency for instance. In the second case, the object is not in the users natural field of view. Techniques have been developed to enhance awareness of the Virtual Environment located behind the user. Nevertheless, in almost all cases, immersed users must at least move their head in order to displace their field of view, and often they have to move their body as well. If we consider the time it takes a user to displace their field of view once, the temporal cost is not too high. But, if all the head and body movements per-

formed during a long work session are added together, the temporal cost becomes significant.

Moreover, such movements ruin the users perception of their initial scope of work and require them to refocus on the task afterwards, which calls for a great degree of precision. Indeed, after moving to locate an object, users have to return to their original position to complete their task. The accumulation of these movements degrades comfort and consequently reduces global performance.

In this article, we describe Pano, a method which enables immersed users to rapidly locate and select objects that are behind them without them having to move. After briefly describing the industrial context of our work in Section 2, we conduct a brief review of previous work in Section 3 and we present Pano, in Section 4. We then go on to describe our testing protocol, implemented in two user studies with 30 and 21 participants in different VEs. Finally, we present the results, which show both the time saved and comfort provided by Pano. We conclude by proposing extensions to this first version.

2 INDUSTRIAL CONTEXT

The aerospace industry is no longer composed of local and individual businesses. Due to the complexity of the products (their size, the number of components, the variety of systems and regulation constraints), the design of an aircraft or a launcher involves a considerable number of engineers with various fields of expertise. Furthermore, aerospace companies often have industrial facilities all over the world.

In such a complex setting, it is necessary to build virtual experiments that can be shared between different remote sites. Specific problems then arise, particularly in terms of the perception of other immersed users and of interaction tasks involving several immersed users. We are working with the Launcher Department of Airbus Group in order to design efficient collaborative interaction methods. These collaborative sessions allow multiple sites to be connected within the same virtual experiment and enable experts from different fields to be immersed simultaneously. For instance, if a problem occurs during the final stages of a launchers construction, it may be necessary to bring together experts who were involved in previous steps and on different sites, for the initial design and manufacturing processes for example.

The inside of a launcher is a very complex environment due to the large number of components, which limits the space where the operator can easily move around (figure 2). Even if users know the environment, it may be time consuming to precisely locate a specific component. This is the reason why immersed experts often have to help immersed operators to perform complex tasks by guiding them, as illustrated in figure 1. Concretely, an expert must indicate to the operator where objects are located inside the launcher.

During our preliminary tests with real users (i.e. members of the Airbus Group), we observed that object finding and, more generally, perception of the environment were crucial. This is not surprising since all real and simulated processes begin with the initial task of locating the objects needed to complete that task (components, tools ...). The speed and comfort with which users perform this initial localization task are therefore essential to the quality and performance of the overall process. This is the initial reason why we developed Pano. However, Pano is also useful in a simpler context with only a single immersed user. In this paper, we present two user studies to demonstrate the benefits of Pano in this simpler context.

3 PREVIOUS WORKS

In this paper, we focus on the first main challenge: perception of the VE (awareness). Gutwin et al. defined awareness as knowledge about the state of a particular environment [10]. As the VE changes

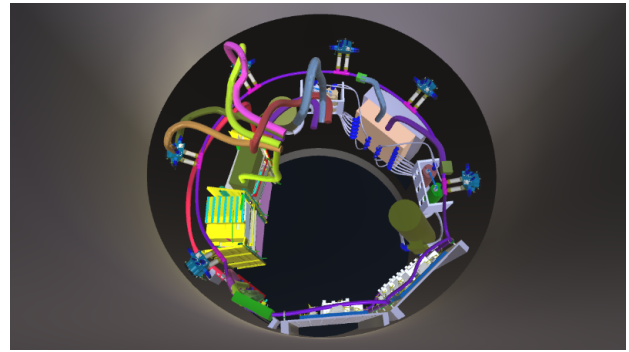


Figure 2: Inside-top view of the launcher with all the components.

over time, users have to maintain and update their awareness of this environment by interacting with it.

In the following, we consider awareness as comprising three aspects:

- Awareness of the components of the VE
- Awareness of other users in the VE
- Awareness of the global process

A lot of techniques have been developed to enhance users awareness of the Virtual Environment. In this section, we present previous works which we have organized around two main approaches. In the first subsection, we describe methods extending the users natural field of view (up to 360). Then, we present techniques based on the Trough-The-Lens metaphor which enables users to explore the VE using different windows [21].

Finally, readers may recall that the subject of the 3DUI Contest in 2012 was exactly the same as our industrial context (cf. Section 2): a guide had to give information through the virtual environment to another user whilst the latter was immersed in the VE. Several participants of the contest [3, 16] proposed that the guide could simultaneously have a world in miniature representation of the VE and the viewpoint of the other user. They suggested that the guide could also add navigation aids by interacting with the world in miniature representation.

3.1 Extending the Field of View to a 360 Image

Humans have been trying to represent the complexity of natural scenes for a very long time. 360 views have long been a popular choice for doing so amongst painters. For instance, 360 views were first used in China during the 10th century, and more recently to represent scenes at the imperial court or in Europe, where English painter R. Barker invented the word Panorama in 1792 and defined it with three main properties: a 360 circular shape, lighting from above the painting and spectators located at the center of the cylinder. One century later, French painter P. Philippoteaux invented cycloramas and painted The Gettysburg Battle which is still visible in the Gettysburg Battlefield visitor center (Pennsylvania). After being used in paintings, 360 views were used in cinematography. In 1897, R. Grimoin-Sanson patented Cineorama: spectators were located at the center of a cylinder onto which projected 10 cinema projectors. This concept has been reused and improved many times since its creation.

For instance, a few years ago, Virtual Reality used [partially-] cylindrical representations in the Reality Center systems popularized by SGI. Henrique Debarba al. [5] investigated extended field of view in an immersive VR setup, and from the perspective of embodied interaction. They noted that the quality of interaction was inferior and required more time.

Augmented Reality systems have used 360 visualization like FlyVIZ [2] which is an HMD enabling users to experience a real-time 360 view of their real surroundings. The panoramic image, which is captured by a camera positioned on top of the users head, is displayed on an HMD screen after having been subjected to certain alterations. The image displayed is monoscopic. Jason Orlosky et al. [18] built the same kind of Augmented Reality HMD which provided binocular vision thanks to two cameras placed in front of the HMD. The field of view could be of 180 or 238 depending of the lenses.

In a less embodied way, Mulloni et al. have presented two other ways of visualizing 360 panoramic images of our surroundings in order to provide clear cues for physical directions in the environment [14]. First, their Top-Down shape is circular and provides good readability of the panorama and direct mapping of the directions. However, since the shape is circular, the panorama is distorted near the center. The second shape is the Birds Eye. A cylindrical shape maps the panorama and the environment in a direct way. According to the 3D view, the panorama is warped and occluded on the sides. After performing several user studies, Mulloni et al. concluded that a good readability of the panorama is most important and that clear representation of the spatial mapping plays a secondary role.

3.2 Through-the-Lens Techniques

Many techniques have been identified in the literature which attempt to ease perception of the virtual environment (VE). These techniques are categorized in the taxonomy of Through-The-Lens techniques [21]. This taxonomy generalizes this metaphor which enables users to simultaneously explore a virtual environment from two different viewpoints. The situations which arise from the relations between these viewpoints are discussed in this paper.

In 1995, Stoakley et al. presented the World In Miniature metaphor [20]. This technique enables the user to handle a miniature copy of the VE. This copy helps the user to build a mental representation of the VE and to directly manipulate objects of the scene through the miniature copy.

Based on a similar concept, Evins et al. proposed Worldlet [6]. A Worldlet is a miniature representation of a particular location in a VE. The representation can be manipulated, and thus enables users to get a better perception of a location than a textual description or a thumbnail image.

In 1998, Fukatsu et al. [8] proposed an intuitive approach to control a bird's eye view while the user is immersed in the VE. This kind of approach is better suited for open world VE with no real boundaries.

Viega et al. presented 3D Magic Lenses that extend the metaphor of a see-through interface embodied in Magic Lenses to 3D environments [22]. There are two kinds of 3D Magic Lenses. The first is a flat lens defined in a 3D space that users can use like a magnifying glass. The second is a volumetric lens in which the 3D objects of the scene appear differently. This kind of tool is useful for discovering hidden properties of the VE, but is useless for getting more information about objects that are not in the users field of view.

The Magic Mirror of Grosjean et al. allows users to get information from locations that are occluded by other objects or to see the behind parts of an object [9]. The magic mirror works like a mirror which the user can position and orientate to access any hidden areas. This tool is suitable for getting specific information about an object, but it is inappropriate for accessing surrounding information.

Through-The-Lens techniques can also be used to build a virtual link between two distant locations. Schmalstieg et al. were the first to propose to use a viewport in the VE as a portal [19]. By walking through the viewport, users can teleport themselves to a remote location. On the one hand, this kind of mechanism can ease nav-

igation into the VE, but on the other hand, it can also disorientate users.

Kiyokawa et al. extended the idea in order to enable users to manipulate objects through a window located in the VE [12]. This kind of technique saves time since users can perform several actions without having to navigate between different sites.

More recently, Kunert et al. presented Photoportals, a mix between viewports and portals with a collaborative purpose [13]. Users can create flat or volumetric Photoportals. These can be used to get insight into a remote location in the VE, and can also be used to teleport users to that remote location, or to take an object from that location. Photoportals can also be used as a kind of camera in order to save a picture or a moment. Users can then use the saved Photoportals to teleport themselves back to the saved scene in order to relive it.

All these Through-The-Lens techniques offer collaborative features. Since different viewpoints can be shared with other users, this can facilitate communication between operators and experts. For instance, the WIM metaphor is well suited for collaborative sessions. If users avatars are displayed on the WIM, it enables one user to know the local context of another and how to reach them. What is more, viewpoint sharing enables users to directly guide other users or to know what other users are doing. That is why, during the 2012 3DUI Contest, many participants proposed using both techniques to guide the immersed user.

Through-The-Lens techniques are designed to facilitate collaboration. A better awareness of the VE implies a better awareness of the other immersed users, since the other users are an important part of the VE [15]. That is why Fraser et al. designed a Through-The-Lens technique for improving collaboration tasks. The Peripheral Lenses technique is designed to enhance awareness of other users [7]. Peripheral Lenses is a virtual navigation aid that simulates peripheral vision. Peripheral Lenses are panels on either side of the main display. They represent viewpoints with the same origin as the main viewpoint, but looking out towards either side. Fraser et al. noted that Peripheral Lenses are useful when users are discussing an object whilst standing side by side in the VE. Indeed, in this case Peripheral Lenses allow one user to know if another user is listening or if he is simply present. Globally, all of the aforementioned techniques could be used in collaborative sessions.

4 PANO

Pano is a Through-The-Lens interaction technique which is based on a window displaying a 360 panoramic image; we call it the PanoWindow (Figure 1, right). The first design of the PanoWindow uses a rectangular shape with rounded sides in order to help the user to understand the continuity effect of a 360 horizontal field of view image. In fact, if an object moves all the way to the right (resp. left) border of the PanoWindow, it disappears and reappears on the left (resp. right) of the same PanoWindow.

This window enables users to select and manipulate objects from the scene by selecting them and dragging their image directly on the screen. Since selection is achieved from the point of view of the panoramic camera, this technique doesn't suffer from the eye-hand visibility mismatch.

4.1 Implementation

The 360 panoramic virtual image is constructed using a set of classic perspective cameras. All cameras must be placed in the same position to create a cylinder view. Figure 3 shows the viewing frustums of the panoramic camera. If n is the number of cameras and A the aspect ratio of the screen, the horizontal and vertical fields of view in radians are:

$$FOV_{horizontal} = \frac{2\pi}{n} \quad (1)$$

$$FOV_{vertical} = 2\text{Arctan}\left(\frac{n}{A} \times \tan\left(\frac{\pi}{n}\right)\right) \quad (2)$$

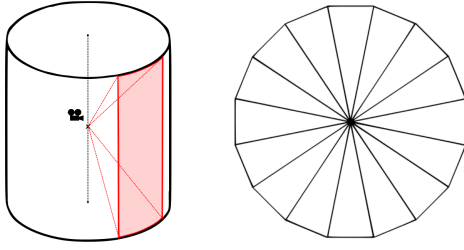


Figure 3: A top view of the cameras' frustums at the right, and diagonal view at the left with the frustum of one camera highlighted.

The number of cameras affects the image quality of the PanoWindow. A large number of cameras reduces the distortion effect between viewports, but it also increases the quantity of rendering operations. It is important to set a good compromise for real time applications.

In order to enable selection and manipulation of 3D objects in the VE directly through the PanoWindow, the collision point between the virtual ray and the screen controls the direction of a ray coming out of the panoramic camera. We use a classical ray-casting technique to perform the manipulation through the PanoWindow. Users must point their virtual ray at the image of the 3D object they wish to select on the PanoWindow. Users can then grab the object by holding a button on the wand controller when the collision point is hovering over the object. Then, users can drag the image of the object over the PanoWindow in order to move the object around the panoramic camera. A joystick on the wand controller allows users to control the distance between the manipulated object and the panoramic camera.

In order to simplify the use of Pano, a slider has been added to the screen which allows users to rotate the panoramic camera around its vertical axis without any real physical movement of the immersed user. This slider enables users to center their view on a specific object to avoid the cropping effect at the edge of the screen.

We added two vertical red lines to delineate a 180 field of view. As long as the panoramic camera has not been rotated, the view between the red lines corresponds to the view in front of the user. The right part of the image is on his right hand side and the left part is on his left. The red lines are effectively visual cues.

4.2 Uses

The most basic application for this technique is to enable users to access their surroundings. Since PanoWindow is a large virtual object, users are able to decide whether they want to display it or not.

Another advantage of this technique is that it can be shared by different users in a collaborative task (as in the industrial context described in Section 2), whether they are co-located or located in different remote sites. As in [3, 16], this technique can be used in asymmetric scenarios in which immersed users share their surrounding view with other users in order to guide them.

Thanks to the selection and manipulation features, users can interact with another users surroundings. These manipulations could be used to help users understand their tasks in the VE. Users can also use Pano to access their own surroundings.

Indeed, if the panoramic camera is attached to the users position, the PanoWindow enables them to visualize their surroundings. The

drawback in this situation is that the PanoWindow almost fills the users field of view. This is why users tend to choose to alternate between viewing the scene without the PanoWindow.

The PanoWindow can also be used by directly positioning the panoramic camera in the VE in order to get information from a distant location simply by checking the PanoWindow. The slider is particularly useful in such a situation. Manipulating objects directly through the PanoWindow is also beneficial in this case.

5 USER EVALUATION

Before evaluating Pano in a collaborative context, we needed to validate the general usability of Pano. Pano is designed to enhance the perception of the environment from a given viewpoint. By doing so, we hypothesize that Pano enables more information to be acquired with less effort, i.e. at a lower cost. Since Pano is a distorted representation of the environment, we also had to evaluate if Pano enables users to build a good mental image of the environment. Thus, we designed a first user study, based on an object-finding task, to test the efficiency of Pano for perceiving the VE. Next, we designed a second user study to test the accuracy of the spatial information provided by Pano. We designed a position-estimation task in order to evaluate the ability of users to build a mental representation of the environment displayed in Pano.

5.1 User Study 1

5.1.1 Participants

The first user study was designed to test the usefulness and the efficiency of the Pano technique. A total of 30 participants (19 males, 11 females) divided into two groups of 15 took part in this user study. 20 of the participants frequently used software which integrates 3D renderings (i.e., on a daily or weekly basis). The remaining 10 had already used software which integrates 3D rendering, but less frequently (i.e., on a monthly basis). 25 participants had already tried a virtual reality HMD. The remaining 5 had never tried a virtual reality setup. Participants were aged between 19 and 34 years old.

5.1.2 Apparatus

For the experiments, we used an HTC Vive in room scale mode. The 2 controllers of the HTC Vive were used. One enabled the participant to control the virtual ray, and the other enabled them to control a small virtual window which gave information about the tasks during the session.

We used a high performance computer (Intel Xeon E5-2680 v2 @ 2.8 GHz, 32 Gb RAM, 12 Gb Nvidia Quadro K6000 GPU). During the experiment, the framerate was constantly at 90 frames per second.

To develop the software, we used Unity 5.3 Personal Edition with the MiddleVR for Unity SDK. MiddleVR is a generic immersive virtual reality plugin which allows the VR system being used to be easily configured in order to adapt the simulation to the system.

5.1.3 Procedure

Participants were invited to perform four different tasks whilst being immersed in the VE using an HMD. Each task was an object-finding task:

- **C1:** Visual Object-finding task (without PanoWindow). Participants had to press a button as soon as they had visually located the target object in the VE.
- **C2:** Visual Object-finding task with PanoWindow. Participants had to press a button as soon as they had visually located the target object in the PanoWindow.

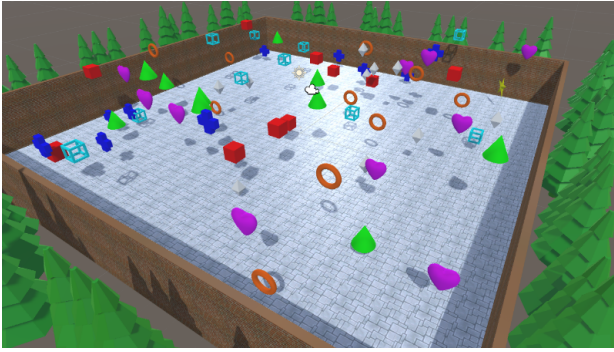


Figure 4: Top-view of the large scene during an object-finding task.

- **C3:** Object-finding and selection task with a ray-casted selection only. Participants had to point and select the target object in the VE with their virtual ray as soon as they had located it.
- **C4:** Object-finding and selection task with PanoWindow. Participants had to point and select the target object in the PanoWindow as soon as they had located it.

Participants were divided into two groups. The first group performed the tasks in the following order: C1-C2-C3-C4. The second group performed them as follows: C2-C1-C4-C3.

Conditions C3 and C4 could have sufficed to draw conclusions as to the effect of Pano, but we considered that the measurements obtained in conditions C1 and C2 could deepen our understanding of the effect of Pano in object-finding tasks. One drawback is that we were unable to verify whether users did indeed press the button only after seeing the correct target object. In order to minimize this source of error, participants were clearly asked to press the button only if they were sure to have seen the target object.

Participants were invited to perform several object-finding tasks in various VEs. We designed two different scenes:

- Inside a small cylindrical scene (Diameter 3 m; Height 4 m)
- Inside a large cubic scene (10 m x 10 m; Height 5 m)

A screenshot of the large scene can be seen in figure 4.

To begin with, participants were positioned in the middle of the scene. They could navigate throughout the scene only by walking in the HMDs tracked volume. When participants had to locate the object through the PanoWindow, they were unable to move it from the middle of the scene.

For each condition, participants had to perform 15 object-finding tasks in each of the two scenes in the following order: small large. For each scene, participants began the session by pressing a button on the controller when they were ready to start. Each of the 15 object-finding tasks was followed by a 5-second break. The task began as soon as objects appeared in the environment and ended when the participant had found the object being searched. During each task, a small window on the participants non-dominant hand showed the object to be found.

To summarize, each participant performed 15 object-finding tasks per scene, 2 scenes per condition in 4 different conditions: i.e. 120 object-finding tasks in total per participant. The evaluation took about 40 minutes in total, 30 minutes for the test and 10 minutes for discussions and questionnaires.

In the scenes, the objects were different colored abstract geometrical shapes (red cubes, green cones, white diamonds, pink hearts, orange tori, light blue hollow cubes, yellow stars and dark blue crosses). For each object-finding task, the scene contained just one

copy of the object to be found. All other objects were replicated 10 times in each scene, so that there were a total of 71 randomly positioned objects for each task.

Initially, object distribution for each task was randomly generated. The distribution was then slightly modified to comply with the two following rules:

- Target objects had to be uniformly distributed around the participant
- Target objects had to not be occluded by any other objects

The distribution of objects for each task was the same in conditions C1 and C2 and in conditions C3 and C4. A pre-study revealed that participants did not notice that they had accomplished all of the tasks twice in different conditions. Thus, all participants performed the tasks in all four conditions. Nevertheless, participants were divided into two groups, one of which began with the PanoWindow and the other without it.

Before the assessment, participants were invited to fill out a form about their experience with 3D software and virtual reality. Before each condition, participants went through a practice scene to make sure that they were familiar with the task at hand during the assessment.

At the end of the assessment, participants were once again asked to complete a final usability and comfort questionnaire. We used a System Usability Scale (SUS) [4] along with more specific items on 5-point Likert scales.

5.1.4 Design

During the assessment, time to completion and users movements were measured. The time it took to complete each object-finding task was recorded directly from the simulation. It was equal to the time that elapsed between the objects appearing in the scene and the moment when the participant completed the task by pressing the button or by directly selecting the object in the scene.

Movements were measured by recording, in each frame, all of the position and orientation coordinates of the head and the dominant hand (i.e. the hand that controls the virtual ray). For each object-finding task, we recorded what we call the angle of appearance. This angle corresponds to the minimal horizontal rotation that a participant has to effect to face the target object. Figure 5 illustrates how this angle was determined.

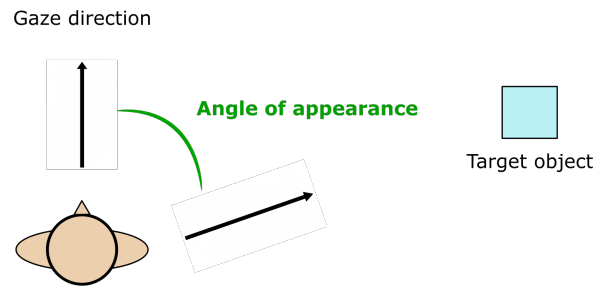


Figure 5: Top-view of the participant and the target object showing how the angle of appearance was calculated.

The formal hypotheses are:

- **H1:** Pano improves completion time for object-finding tasks where the objects are distributed all around the participant. Average completion time in conditions C2 and C4 are shorter than in conditions C1 and C3.

- **H2:**Pano minimizes the movements of the immersed user in conditions C2 and C4 vs C1 and C3. This reduction improves the users comfort for object-finding tasks where the objects are distributed all around the participant.

We are aware that comfort is a global concept and minimizing movements does not mean more comfort. But in the case of virtual reality, it is well known that head movement can induce cybersickness. Thus, avoiding cybersickness is part of the users comfort. In other words, if Pano helps to minimize movements, users could perform the task while sat on a chair, which could also increase comfort.

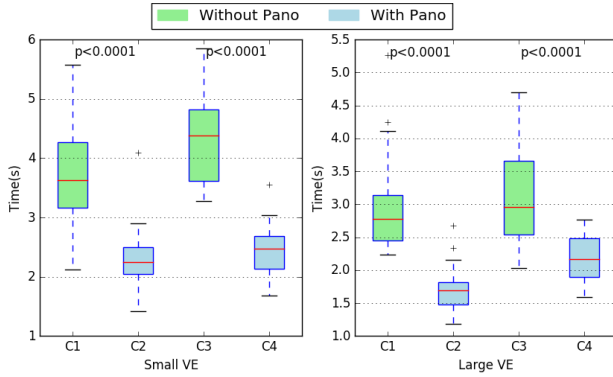


Figure 6: Time to complete the object-finding tasks (in seconds) for each condition in the two different scenes (*mean* \pm *SD*)

5.1.5 Results

The results presented in this section were considered statistically significant when $p < 0.05$, and were explicitly referred to as a trend if $0.05 < p < 0.1$. We performed a Wilcoxon Signed-rank test with all the measurements per scene, per participant and per condition. Then, we corrected for multiple comparisons using the false discovery rate (FDR) [17]. All the analyses were performed with Rkward, an integrated development environment for R.

In the following part of this section, we first compare time to completion depending on the condition and the scene. Then, we compare movements of the participants during the tasks with and without the PanoWindow.

Globally, participants performed the object-finding task faster with the PanoWindow than without. All of the following results can be seen in figure 6.

More specifically, for the small cylindrical scene, we found a significant effect on completion time. Results revealed a statistically significant difference between conditions C1 and C2 ($p < .0001$) and between C3 and C4 ($p < .0001$). Completion time for C1 (mean = 3.745, SD = 0.877) was higher than for C2 (mean = 2.298, SD = 0.521) and completion time for C3 (mean = 4.315, SD = 0.720) was higher than for C4 (mean = 2.455, SD = 0.422).

For the large cubic scene, we also found a significant effect on completion time. Results revealed a statistically significant difference between conditions C1 and C2 ($p < .0001$) and between C3 and C4 ($p < .0001$). Completion time in C1 (mean = 2.942, SD = 0.693) was higher than in C2 (mean = 1.696, SD = 0.322) and completion time in C3 (mean = 4.315, SD = 0.720) was higher than in C4 (mean = 2.455, SD = 0.422).

In Figures 11 and 10, mean completion time is plotted against the angle of appearance of the target object.

The higher the angle of appearance of the target object, the longer it takes to complete the task. Furthermore, there seems to be a breaking point around 60 for each scene. The fact that the field

of view of the HTC Vive is approximately 110 can help to explain this phenomenon. Globally, time to completion is shorter with the Pano technique (conditions C2 C4).

We defined gain as the ratio between the completion times for C1 and C2 and between those for C3 and C4. We can see that the gain is greater when the angle of appearance is between 60 and 140 degrees (small sc. C3-C4, large sc. C1-C2 and C3-C4).

Participants made fewer movements to complete the object-finding task with the PanoWindow than without. Part of the following results can be seen in figures 7 and 8. All boxplots have not been displayed to avoid overloading the figures.

More specifically, for the small cylindrical scene, we found a significant effect on hand movement, head movement, hand rotation and head rotation. For hand movement, results revealed a statistically significant difference between conditions C1 and C2 ($p < .0001$) and between C3 and C4 ($p < .0001$). Hand movement for C1 (mean = 19.02, SD = 5.82) was higher than for C2 (mean = 2.69, SD = 2.14) and hand movement for C3 (mean = 33.13, SD = 12.05) was higher than for C4 (mean = 6.47, SD = 3.57). For head movement, results revealed a statistically significant difference between conditions C1 and C2 ($p < .0001$) and between C3 and C4 ($p < .0001$). Head movement for C1 (mean = 13.57, SD = 2.82) was higher than for C2 (mean = 2.72, SD = 1.63) and head movement for C3 (mean = 16.11, SD = 5.10) was higher than for C4 (mean = 2.74, SD = 1.53). For hand rotation, results revealed a statistically significant difference between conditions C1 and C2 ($p < .0001$) and between C3 and C4 ($p < .0001$). Hand rotation for C1 (mean = 10300, SD = 5061) was higher than for C2 (mean = 4763, SD = 4455) and hand rotation for C3 (mean = 19032, SD = 5045) was higher than for C4 (mean = 11295, SD = 3941). Finally, for head rotation, results revealed a statistically significant difference between conditions C1 and C2 ($p < .0001$) and between C3 and C4 ($p < .0001$). Head rotation for C1 (mean = 23428, SD = 4784) was higher than for C2 (mean = 10043, SD = 4018) and head rotation for C3 (mean = 27653, SD = 6499) was higher than for C4 (mean = 9196, SD = 3616).

For the large cubic scene, we found a significant effect on hand movement, head movement, hand rotation and head rotation as well. For hand movement, results revealed a statistically significant difference between conditions C1 and C2 ($p < .0001$) and between C3 and C4 ($p < .0001$). Hand movement for C1 (mean = 18.34, SD = 5.21) was higher than for C2 (mean = 1.62, SD = 1.18) and hand movement for C3 (mean = 24.52, SD = 6.86) was higher than for C4 (mean = 5.11, SD = 3.84). For head movement, results revealed a statistically significant difference between conditions C1 and C2 ($p < .0001$) and between C3 and C4 ($p < .0001$). Head movement for C1 (mean = 10.21, SD = 2.75) was higher than C2 (mean = 1.68, SD = 0.97) and head movement for C3 (mean = 8.97, SD = 2.07) was higher than C4 (mean = 2.23, SD = 1.36). For hand rotation, results revealed a statistically significant difference between conditions C1 and C2 ($p < .0001$) and between C3 and C4 ($p < .0001$). Hand rotation for C1 (mean = 9644, SD = 4816) was higher than in C2 (mean = 3436, SD = 3802) and hand rotation for C3 (mean = 14604, SD = 2434) was higher than in C4 (mean = 8017, SD = 2151). Finally, for head rotation, results revealed a statistically significant difference between conditions C1 and C2 ($p < .0001$) and between C3 and C4 ($p < .0001$). Head rotation for C1 (mean = 16766, SD = 5399) was higher than in C2 (mean = 7425, SD = 2931) and head rotation for C3 (mean = 16239, SD = 4036) was higher than in C4 (mean = 7175, SD = 2946).

We now proceed to present the qualitative results gained from the surveys and from our discussions with the participants.

All the participants agreed that using Pano could improve completion time for object-finding tasks. The average score was 4.4 out of 5 (SD = 0.83).

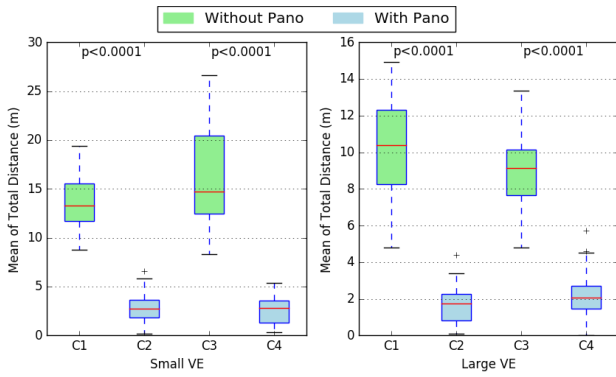


Figure 7: Head movement (translation, in meters) during object-finding tasks in each of the four conditions, in the two different scenes (*mean* \pm *SD*)

Generally, participants moved significantly less with Pano than without. Figures 7 and 8 show boxplots for the total distance travelled by the dominant hand and the head. Participants agreed that they moved less with Pano than without. The average score for this question was 4.0 ($SD = 1.42$). Participants also agreed that Pano rendered the tasks less tiring, with a score of 4.56 ($SD = 0.71$).

The score for the SUS was 85/100. It is important to bear in mind that the participants performed all the conditions.

To conclude, hypotheses H1 and H2 were verified for all the scenes.

5.2 User Study 2

5.2.1 Participants

The second user study was designed to test if the altered image provided by Pano can nevertheless enable users to build a good mental representation of the environment. A total of 21 participants (18 males, 3 females) took part in this user study. 13 of the participants frequently used software which integrates 3D renderings, i.e. on a daily or weekly basis. 4 of the participants had already used software which integrates 3D rendering, but less frequently, i.e. on a monthly basis. The remaining 4 had never or rarely used this kind of software. 9 participants had already tried a virtual reality HMD. The remaining 12 had never tried a virtual reality setup. Participants, aged between 18 and 37 years old, were invited to perform tasks whilst being immersed in the VE using an HMD.

Before the assessment, participants were invited to perform two paper tests in order to evaluate their cognitive spatial abilities. The first test evaluated mental rotation abilities [1], and the second evaluated the perspective-taking abilities of the participants [11]. The results confirmed that participants had varying abilities in these two areas. Participants were invited to perform a position-estimation task in two conditions.

5.2.2 Apparatus

For this experiment, we also used an HTC Vive in room scale mode. The 2 controllers of the HTC Vive were used.

We used a gaming performance computer (Intel Core I7 @ 4.0 Ghz, 16 Gb RAM, 6Gb Nvidia GTX 1060). During the experiment, the framerate was constantly at 90 frames per second.

To develop the software, we used Unity 5.3 Personal Edition with the SteamVR plugin for Unity.

5.2.3 Procedure

Participants were invited to perform a position estimation task in two conditions. For both tasks, participants were positioned in the

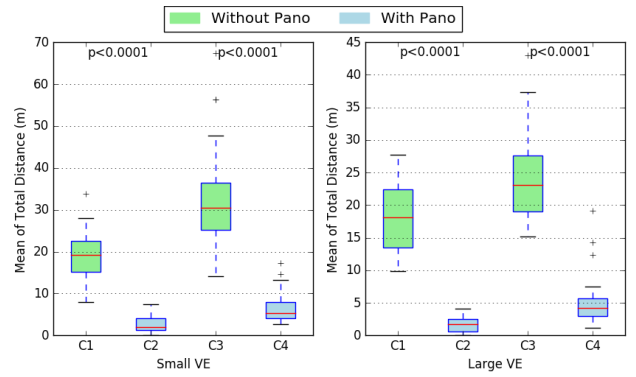


Figure 8: Hand movement (translation, in meters) during object-finding tasks in each of the four conditions, in the two different scenes (*mean* \pm *SD*)

middle of the scene with a PanoWindow placed right in front of them.

The position-estimation task was divided into 3 steps:

- A visualization step: participants had to locate the image of an object (that we will call the target) on the PanoWindow;
- An estimation step: participants had to estimate the position of the object seen in the PanoWindow by positioning a second object (that we will call the answer) in their surrounding environment. During this step, the PanoWindow is hidden to enable users to position the answer more easily;
- A feedback step: participants could see both objects (the target and the answer) in their surroundings and in the PanoWindow so they were able to appreciate the errors they made.

The two conditions were:

- Horizontal: the positions to be estimated were all at the same height (1 meter from the ground). Participants simply had to point in the direction of the target object.
- Horizontal and Vertical: the positions to be estimated were all around the participants on the surface of a cylinder, with a radius of 5 m.

Participants performed the Horizontal and the Horizontal Vertical conditions twice in the following order:

- Horizontal 1
- Horizontal Vertical 1
- 5 min break
- Horizontal 2
- Horizontal Vertical 2

We invited the participants to perform the same tasks twice in order to see if they were able to improve their performance just by repeating it. For each condition, participants performed 20 position-estimation tasks.

To summarize, each participant performed 40 positionestimations per condition: i.e. 80 position-estimation tasks in total per participant. The evaluation took about 40 minutes in total, 20 minutes for the test and 20 minutes for discussions and questionnaires.

Targets in the 20 position-estimation tasks were designed to be uniformly distributed around the participant.

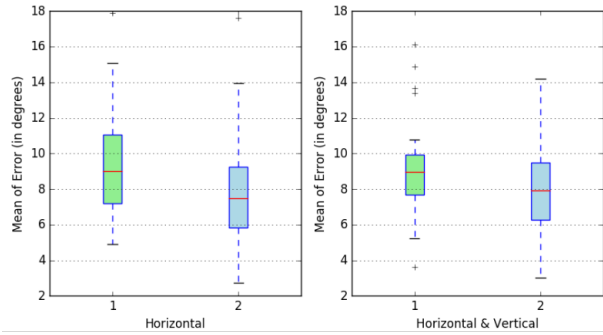


Figure 9: Mean Error in position-estimation task for each condition.

5.2.4 Design

For the Horizontal condition, the angle between targets and answers was measured in degrees for each position-estimation task. In the Horizontal Vertical condition, errors of height were measured in meters. The time to complete each position-estimation task was also recorded during the assessment.

The hypothesis is:

- **H:** Pano provides an image which enables users to estimate the position of the target objects with an accuracy of 30.

This hypothesis is inspired by our collaborative context. If users can estimate the right clock position of an object in order to guide someone else, then Pano can support communication efficiently. Because a clock is divided into 12 directions, users need to estimate the clock position of objects with a precision of 30 (since 30 is 1/12 of 360).

5.2.5 Results

In order to test our hypothesis, we computed the mean and the standard deviation of the error in degrees.

Globally, participants performed the position-estimation tasks with a horizontal error of less than 30 for each scene. Thus, the hypothesis is verified.

The results are:

- Horizontal 1, mean = 9.59 and SD = 3.45
- Horizontal 2, mean = 8.18 and SD = 3.31
- Horizontal Vertical 1, mean = 9.40 and SD = 2.99
- Horizontal Vertical 2, mean = 8.03 and SD = 2.49

These results are shown in figure 9.

6 DISCUSSION

Globally, these results demonstrate that the Pano technique fully meets our first expectations. Whatever the orientation and the location of the object in the scene, users are able to easily and quickly identify it when they use Pano.

Time savings: Both visual search time and physical motion were taken into account in the assessment of overall time savings. The results show statistically significant differences in completion time (mean, standard deviation) between using ray-casted selection alone vs Pano.

Comfort: The object-finding task requires participants to move their head around in all directions and Pano reduces these movements. Consequently, we hope that using Pano could contribute

to decreasing feelings of cybersickness. Before the session, 2 participants informed us of cybersickness problems they usually encountered when using VR. After the session, they reported feeling cybersickness when they were not using Pano, but conversely, they felt better when using the technique.

Selection: Globally, participants reported that they could easily select objects using Pano. Some of them said that selection in the PanoWindow was more difficult than direct selection in the VE. Classic virtual ray selection involves the whole arm, while selection in the PanoWindow just involves the wrist. Moreover, the image in the PanoWindow is generally smaller than the object in the scene. Thus, direct selection in Pano requires a greater degree of concentration and accuracy.

One participant said that he would prefer to interact with a tactile screen. Instead of using the virtual ray for pointing, this participant would have preferred to be able to use the virtual hand to touch the object on the PanoWindow.

Globally, the participants enjoyed using Pano. The SUS score was 85. This score allows us to assert that this first version of Pano is relatively easy to use.

Position-estimation task: Globally, participants performed well at this task although their spatial abilities were different.

This test shows that users can quickly learn to use Pano to build a mental representation of their environment. They achieve sufficient accuracy to be able to use Pano and guide someone else. In this case, the guide could use the clock position technique to describe the position of an object to another user. Furthermore, graduation on the PanoWindow could be added to make the clock positions explicit.

Figure 12 shows the error depending on the angle of appearance of the target. We can see that the error is minimal for the 4 angles that correspond to ahead, behind and the two sides. Precision fell between these positions because there are no references to help the users.

7 CONCLUSION AND FUTURE WORK

In this paper, we presented Pano, a visualization and interaction technique which reduces search and selection time by enhancing the awareness of the Virtual Environment for immersed users. We described the implementation of this technique and then presented two test protocols involving 30 and 21 participants using Pano. For the first user study, results demonstrate the benefits both for reducing the global duration of the task and improving the comfort of immersed users. Therefore, we conclude that Pano is readable and useful for sharing viewpoints. As for the second user study, results show that Pano enables users to construct a good mental representation of the environment, accurate enough to support communication.

We are currently working along three main axes. Firstly, we are working both on the implementation and the generic uses of Pano. For instance, what is the best shape for the window? What kinds of interaction features are better performed directly in the PanoWindow rather than in the full VE?

Secondly, we are currently working with experts from Airbus Group in order to introduce Pano into their real collaborative processes. Then, we would like to evaluate the benefits of the technique in terms of presence and communication between users in a real industrial context

Thirdly, we think that Pano could be useful outside the initial context of collaborative work for industrial scenarios. For instance, we think Pano could be used for 3D Data Visualization and Exploration. Furthermore, we also think that Pano could be used for video games. We could design an asymmetric and cooperative game using this technique.

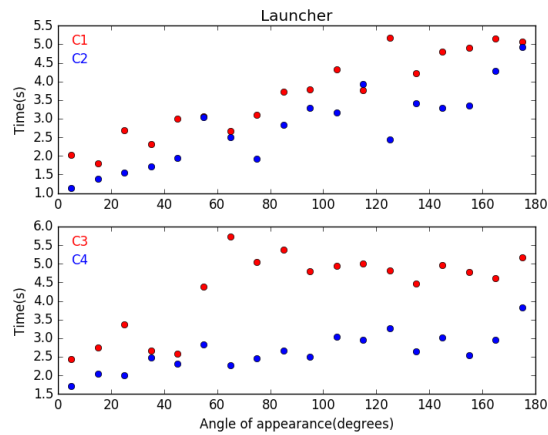


Figure 10: Average task completion time depending on the angle of appearance of the target object for each condition in the small VE.

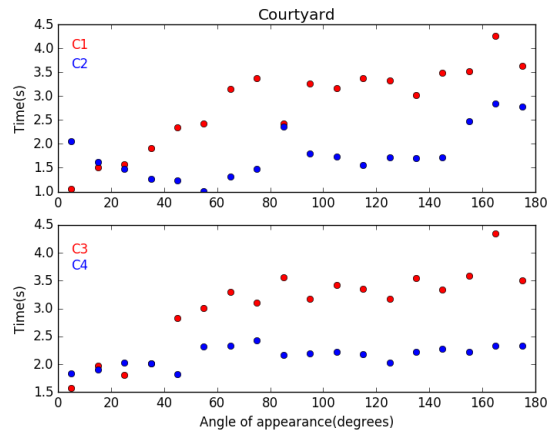


Figure 11: Average task completion time depending on the angle of appearance of the target object for each condition in the large VE.

REFERENCES

- [1] J. M. Albaret and E. Aubert. Etalonnage 15-19 ans du test de rotation mentale de vanderberg. *Evolutions psychomotrices*, pages 206–209, 1996.
- [2] J. Ardouin, A. Lécuyer, M. Marchal, C. Riant, and E. Marchand. Flyviz: a novel display device to provide humans with 360 vision by coupling catadioptric camera with hmd. In *Proceedings of the 18th ACM symposium on Virtual reality software and technology*, pages 41–44. ACM, 2012.
- [3] F. Bacim, E. D. Ragan, C. Stinson, S. Scerbo, and D. A. Bowman. Collaborative navigation in virtual search and rescue. In *3D User Interfaces (3DUI), 2012 IEEE Symposium on*, pages 187–188. IEEE, 2012.
- [4] J. Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.
- [5] H. G. Debarba, S. Perrin, B. Herbelin, and R. Boulic. Embodied interaction using non-planar projections in immersive virtual reality. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, pages 125–128. ACM, 2015.
- [6] T. T. Elvins, D. R. Nadeau, and D. Kirsh. Worldlets3d thumbnails for wayfinding in virtual environments. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*, pages 21–30. ACM, 1997.

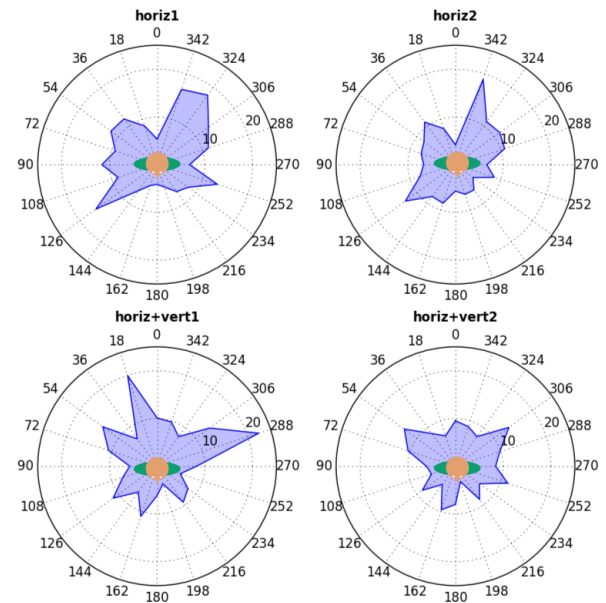


Figure 12: Error mean in degrees depending of the angle of the target for each condition. Participants were facing 180.

- [7] M. Fraser, S. Benford, J. Hindmarsh, and C. Heath. Supporting awareness and interaction through collaborative virtual interfaces. In *Proceedings of the 12th annual ACM symposium on User interface software and technology*, pages 27–36. ACM, 1999.
- [8] S. Fukatsu, Y. Kitamura, T. Masaki, and F. Kishino. Intuitive control of bird's eye overview images for navigation in an enormous virtual environment. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 67–76. ACM, 1998.
- [9] J. Grosjean and S. Coquillart. The magic mirror: A metaphor for assisting the exploration of virtual worlds. In *Spring Conference on Computer Graphics*, pages 125–129, 1999.
- [10] C. Gutwin and S. Greenberg. The importance of awareness for team cognition in distributed collaboration. 2001.
- [11] M. Hegarty and D. Waller. A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32(2):175–191, 2004.
- [12] K. Kiyokawa and H. Takemura. A tunnel window and its variations: Seamless teleportation techniques in a virtual environment. In *HCI International*. Citeseer, 2005.
- [13] A. Kunert, A. Kulik, S. Beck, and B. Froehlich. Photoportals: shared references in space and time. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1388–1399. ACM, 2014.
- [14] A. Mulloni, H. Seichter, A. Dünser, P. Baudisch, and D. Schmalstieg. 360 panoramic overviews for location-based services. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2565–2568. ACM, 2012.
- [15] T. T. H. Nguyen and T. Duval. A survey of communication and awareness in collaborative virtual environments. In *Collaborative Virtual Environments (3DCVE), 2014 International Workshop on*, pages 1–8. IEEE, 2014.
- [16] T. T. H. Nguyen, T. Duval, and C. Fleury. Guiding techniques for collaborative exploration in multi-scale shared virtual environments. In *GRAPP International Conference on Computer Graphics Theory and Applications*, pages 327–336, 2013.
- [17] W. S. Noble. How does multiple testing correction work? *Nature biotechnology*, 27(12):1135–1137, 2009.
- [18] J. Orlosky, Q. Wu, K. Kiyokawa, H. Takemura, and C. Nitschke. Fish-eye vision: peripheral spatial compression for improved field of view in head mounted displays. In *Proceedings of the 2nd ACM symposium*

- on *Spatial user interaction*, pages 54–61. ACM, 2014.
- [19] D. Schmalstieg and G. Schaufler. Sewing worlds together with seams: A mechanism to construct complex virtual environments. *Presence: Teleoperators and Virtual Environments*, 8(4):449–461, 1999.
 - [20] R. Stoakley, M. J. Conway, and R. Pausch. Virtual reality on a wim: interactive worlds in miniature. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 265–272. ACM Press/Addison-Wesley Publishing Co., 1995.
 - [21] S. L. Stoev and D. Schmalstieg. Application and taxonomy of through-the-lens techniques. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 57–64. ACM, 2002.
 - [22] J. Viega, M. J. Conway, G. Williams, and R. Pausch. 3d magic lenses. In *Proceedings of the 9th annual ACM symposium on User interface software and technology*, pages 51–58. ACM, 1996.