First-person Experience and Usability of Co-located Interaction in a Projection-based Virtual Environment

Andreas Simon
University of Applied Sciences Aargau
Bahnhofstrasse 102, 5000 Aarau, Switzerland
ansimon@gmail.com

ABSTRACT

Large screen projection-based display systems are very often not used by a single user alone, but shared by a small group of people. We have developed an interaction paradigm allowing multiple users to share a virtual environment in a conventional single-view stereoscopic projection-based display system, with each of the users handling the same interface and having a full first-person experience of the environment.

Multi-viewpoint images allow the use of spatial interaction techniques for multiple users in a conventional projection-based display. We evaluate the effectiveness of multi-viewpoint images for ray selection and direct object manipulation in a qualitative usability study and show that interaction with multi-viewpoint images is comparable to fully head-tracked (single-user) interaction. Based on ray casting and direct object manipulation, using tracked PDA's as common interaction device, we develop a technique for co-located multi-user interaction in conventional projection-based virtual environments. Evaluation of the VRGEO Demonstrator, an application for the review of complex 3D geo-seismic data sets in the oil-and-gas industry, shows that this paradigm allows multiple users to each have a full first-person experience of a complex, interactive virtual environment.

Categories and Subject Descriptors

I.3.7 [**Computer Graphics**]: Three-Dimensional Graphics and Realism – *Artificial, augmented, and virtual realities*, H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces – *Collaborative computing*

General Terms

Human Factors

Keywords

Co-located collaboration, Single Display Groupware, projection-based virtual environment, PDA interaction

1. INTRODUCTION

Large screen, stereoscopic projection-based display systems have become the prevalent display paradigm for virtual environments in most research labs and for industrial applications. Originally, these displays were developed as fully head tracked, immersive, single user environments [6], as an ergonomic, high-resolution, low-lag alternative to head mounted displays.

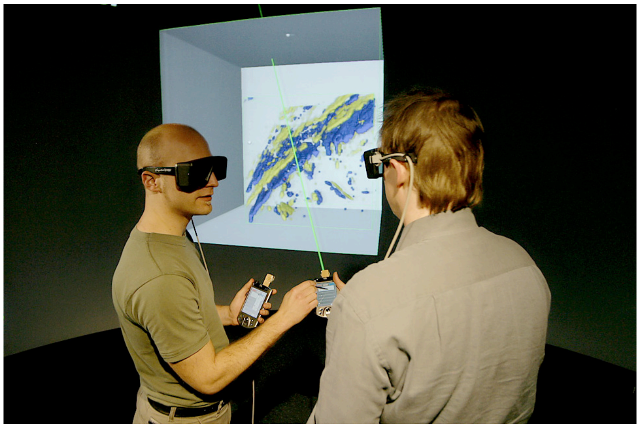

Figure 1. Two users sharing a projection-based virtual environment display

A key aspect of head-tracked immersive projection-based displays is their ability to transparently create a mixed reality environment, matching the viewed positions of virtual objects and the real space inside the working volume of the display system. This alignment of real and virtual space allows the user to naturally perceive his body, his hands, and the tracked interaction devices as part of the virtual environment. It is the basis for the efficient use of direct spatial interaction techniques such as object manipulation and ray casting selection in projection-based virtual environments.

In practice however, most projection-based displays are not used by a single user alone. Indeed, the ability of these displays to support ad-hoc collaboration is one of their key success factors, in particular in industrial applications. Typically, small groups use such a system together to directly view and share a virtual environment. Using a conventional, single-view display system, all viewers inside the display share exactly one and the same stereoscopic image. In this case, since the precise matching of real and virtual space requires fully viewpoint correct head-tracked images, only one user benefits from a completely accurate, undistorted, spatial image. In a typical scenario, this "master user" handles all the interaction in the application. All other viewers who share the display with the head-tracked master user necessarily view the stereoscopic image from a different viewpoint and with a different view direction. This results, from their perspective, in a distorted spatial image, exhibiting strong unwanted image motion (coupled to the motion of the head tracked viewer), and no longer guaranteeing a good match between real and virtual space. Consequently, in a projection-based display environment, only the head-tracked "master user" benefits from a full first-person interactive experience of the application; all other participants only see a secondary view and typically cannot use the same interaction tools effectively [19].

Based on omnistereo projection and multi-viewpoint images, we have developed a paradigm for co-located interaction in an immersive projection-based environment that aims to create a full *first-person experience for every user* in a projection-based display environment. Omnistereo projection allows good stereoscopic viewing of a single stereoscopic image for multiple users, independent of view direction, over a full 360° field of view [17]. Multi-viewpoint images let us "unwrap" the parallax for each individual user, projecting virtual interaction objects, such as intersection rays, in the correct position for each user [18]. This enables all users of a projection-based display system to use direct spatial interaction techniques in the virtual environment. To confirm that we can provide a good first-person experience for spatial interaction to multiple users, we have investigated ray casting selection and direct object manipulation with multi-viewpoint images in a qualitative and quantitative usability study. We introduce spatially tracked PDAs as a common interaction device for every user of the system, combining ray casting selection and direct object motion in the virtual environment with system control for menus, tools, and modes on the "private" interface of each user's PDA. This interaction paradigm extends the use of PDAs in Single Display Groupware (SDG) applications [21] by allowing full spatial interaction for co-located collaboration in a virtual environment.

The remainder of this paper is organized as follows: Section 2 briefly introduces multi-viewpoint images for interaction. Section 3 reviews previous work. Section 4 presents usability studies for ray casting selection and direct object manipulation. Section 5 presents the interaction paradigm implemented in the VRGEO Demonstrator and reports experiences with the VRGEO Demonstrator. Section 6 concludes and presents opportunities for future work.

2. MULTI-VIEWPOINT IMAGES

Viewing stereoscopic images from a viewpoint outside of the projection viewpoint introduces parallax, a skew distortion of the spatial image, resulting in a misalignment between real and virtual object positions. We use multi-viewpoint images [18], composing different image elements projected from multiple viewpoints into a single, consistent stereoscopic image, to overcome the parallax problem in non-head-tracked applications and to enable multi-user interaction in a shared projection-based display.

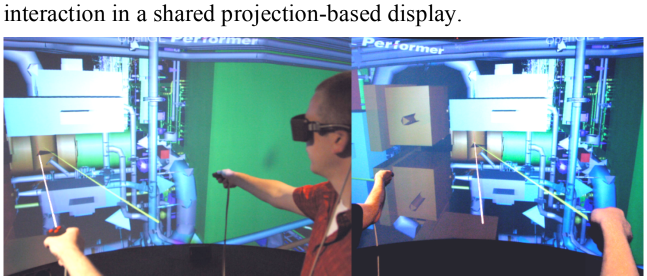

Figure 2. Left vs. right user's view of a single multi-viewpoint image: Each user's picking ray aligns correctly from his respective viewpoint

The multi-viewpoint image in Figure 2 is one and the same image. It combines three different viewpoint projections: One for the main scene and one for each of the two users. The main scene, containing engines and pipes, is rendered without head tracking from a static viewpoint centered in the middle of the display. For each of the two users, the user's picking ray is rendered from the respective user's head-tracked viewpoint. This places the picking ray, seen from that user's perspective, in correct alignment with his tracked interaction device.

With multi-viewpoint images, we can project different image elements from multiple viewpoints, corresponding to the viewing positions of multiple users, and combine them in a single image. This technique is independent of the number of users sharing the display. We use multi-viewpoint images to project interaction elements for each user in the correct position and depth, matching, from the user's point of view, the tracked real positions of interaction devices with the virtual position of visual and functional interaction elements such as pointers, menus or picking rays. The main scene in such a multi-viewpoint image is rendered from a fixed viewpoint, removing unwanted image motion.

3. RELATED WORK AND CONTRIBUTION

The benefits of virtual environments for collaboration and for providing a shared experience have been recognized for a long time. For co-located collaboration, the ability to simply share a large projection-based display with multiple people has been one of the major practical successes of these systems. This simplified approach, however, leads to a paradigm, where only one head-tracked user controls the interaction, taking the other participants on a ride with him. [19] has recognized the interaction problems that arise from the non head-tracked viewing in these displays, and has compared different object selection techniques for non head-tracked and head-tracked interaction.

True co-located collaboration in virtual environments, allowing two or more users to effectively share a common virtual environment as well as a real place, and to interact and to engage each other directly, has been approached from two mayor directions, depending on the respective display paradigm.

On one hand, there is research in mixed reality interfaces for head mounted displays (HMD). HMDs are inherently suitable for multi-user collaboration and interaction in virtual environments since they display an individual view for each user. Several such systems and application scenarios, typically using an augmented reality (AR) approach to viewing and interaction, have been proposed [23] [4] [3]. Studierstube [23] was one of the first systems to show the potential of an Augmented Reality approach for co-located collaboration. Regenbrecht [13] describes an HMD-based AR system, which allows multiple participants to interact with two- and three-dimensional data using tangible user interfaces. The system is based on a tabletop metaphor and uses camera-tracked markers on paper cards or props to provide a physical interface to virtual objects that can be naturally handled and shared by multiple participants. A PDA is used as a data entry palette, using a pick-and-drop [14] style interface to drag virtual objects onto the table.

Conventional projection-based displays only allow the display of only one head-tracked stereoscopic image. Several researchers have addressed this display problem by developing systems that allow the display of more then one stereoscopic image in a shared working volume, presenting a separate, fully head-tracked image for each user. The duo-responsive workbench [1] is a multi-view display system that supports two users by sequentially displaying four images on the screen of a responsive workbench. Other multi-user projection displays use some sort of spatial barrier to separate images for the different users [2][9][5]. Spatial barrier-type displays severely restrict the usable shared viewing volume and the field of view for each user. With the notable exception of Agrawala et al. [1], who present an interaction paradigm for the duo-workbench system, most of the work on multi-user projection-based display systems is solely concerned with technical aspects of display technology and rendering and does not develop suitable interaction paradigms.

For conventional (non-3D) scenarios, Stewart et al. [21] have coined the term Single Display Groupware (SDG) for co-located collaboration on a single shared screen. The Pebbles project by Meyers [12] connects multiple PDAs to a main computer displaying on a large projection screen in as SDG scenario. The PDAs are primarily used to control multiple mouse and keyboard input to whiteboard applications. Rekimoto has developed a system involving a shared display and private mobile devices [15]. Rekimoto introduces mobile computers and PDAs as common, spatially tracked interaction devices into his shared environment. At any time, with a special stylus, a user can pick-and-drop private information from the PDA and place it on the shared, public display.

While most researchers developing co-located interaction scenarios for projection-based virtual environments have concentrated on developing the technical aspects of multi-view display technologies, we present an interaction paradigm for co-located interaction using a conventional panoramic stereoscopic display system to allow multiple users to interact in a virtual environment with a full first-person experience. We evaluate the quality of the first-person user experience for direct interaction with a multi-viewpoint image by qualitative and quantitative usability testing. Our approach presents a suitable interaction paradigm using multi-viewpoint images [18] to allow multiple users to use direct interaction techniques for ray selection and object manipulation in an inside-out panoramic display. While the use of a PDA as interface device in a virtual environment is hardly new [25], we introduce spatially tracked PDAs in a co-located collaborative environment in the sense of Single Display Groupware systems [12], as a private tool and display to augment the shared display, combining 3D spatial and symbolic interaction.

4. USABILITY OF MULTI-VIEWPOINT IMAGES FOR INTERACTION

We have evaluated the usability of multi-viewpoint images for ray casting selection and for direct object manipulation in a simple docking task. The aim of this study is to establish the interaction quality of the multi-viewpoint technique, compared to conventional non head-tracked interaction, and compared to fully head-tracked (single-user) interaction. Since our interaction paradigm will use these direct spatial interaction techniques, performance of the multi-viewpoint technique comparable to the head-tracked single-user case will indicate a good first-person interaction experience for every user in the projection-based virtual environment.

The basic experimental setup can be seen in Figure 3. The subject is standing on a marked, fixed position 0.8m outside of the middle, the central projection viewpoint of our i-Cone stereoscopic display system [16]. The i-Cone display has a curved (5º conical section) projection screen with a radius of 3m. It uses four edge-blended Barco BR909 Reality CRT projectors to create a seamless 240º field of view active stereoscopic image. Stereo-Graphics ChrystalEyes3 shutter glasses are used for viewing. We use a Linux cluster with nVidia FX3000G genlocked graphics cards to drive the system with a total resolution of 8000x1600 pixel at 94Hz. The experiments are running with 47Hz constant frame rate. A Polhemus Fastrack system with a Polhemus Stylus as interaction device is used for 6DOF tracking. The non head-tracked and the multi view image conditions use omnistereo rendering with the static viewpoint in the center of the display at 1.7m; the fully head-tracked condition implements distortion correction for curved screen geometry. The experiments are implemented using the AVANGO [24] virtual environment framework.

4.1 Ray Selection

We have evaluated the effectiveness of multi-viewpoint images for a ray casting selection task in a qualitative and quantitative usability study with twelve unpaid subjects. Subjects were between 22 and 38 years of age, ten of the subjects were male, two female, eleven right-handed, one left-handed. Four of the subjects were expert users, eight subjects had little prior experience with immersive virtual environments.

The study is using a counterbalanced within subjects design to compensate for learning and order effects. Direct ray casting is used as interaction technique, independent variables of the experiment are the display condition – non head-tracked (NoHT), multi-viewpoint image (MV) with picking ray, and fully head-tracked image (HT) – and the distance to the selection target. Dependent variables of the experiment are the selection time and the number of target re-entries per trial.

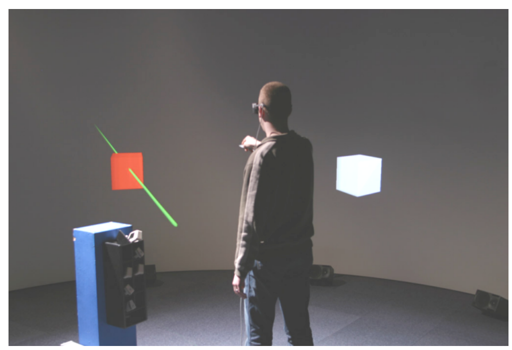

Figure 3. Performing ray selection in the i-Cone

Following the experimental task design of experiments for evaluation of desktop pointing devices [10], participants were asked to point at targets appearing in different positions in the display system. The targets, 0.1m cubes, were placed at a distance between 1m and 1.5m from the participant in the near condition, and, using 0.3m cubes, between 5m and 7m in the far condition. We have used larger targets for the far object distances, to approximately match task difficulty for the two conditions. The angular distance between start position and selection target typically was about 60º, resulting in an index of difficulty of 3.5 bit for the near, and 4.5 bit for the far condition.

When a subject points at a target, the target would highlight; pointing at a target with 0.7s dwell time completes a selection. In contrast to other pointing studies, we did not require a button press for selection to prevent unwanted pitching motion of the stylus. After selection, a new target would appear at a new position within 0.3s. Selection time is measured from leaving the start position to completing the new selection (excluding dwell time); the target re-entry count would be measured as the number of target entries before dwell time was reached and selection was completed. A run would consist of 15 selections. Subjects would do two runs per condition, for a total of 180 task repetitions.

Subjects receive brief oral instructions and were instructed to complete the selection as accurately and as fast as possible. Subjects had about 3 minutes to familiarize themselves with the task in the near, fully head-tracked condition. After executing the test, we would conduct a 5-minute interview. In the interview, subjects were asked to rate the "ease of task execution", the "naturalness of mapping" and the "quality of stereoscopic viewing" for each of the display conditions on a five point Lickert scale (1–poor to 5–excellent). Total test time per subject was about 25 minutes.

4.2 Direct Object Manipulation

We have also evaluated multi-viewpoint images for a simple docking task with direct object manipulation in a quantitative usability study with eight unpaid subjects. Subjects were between 22 and 38 years of age, seven of the subjects were male, one female, six right-handed, two left-handed. Three of the subjects were expert users, five subjects had little prior experience with immersive virtual environments.

The study is using a counterbalanced within subjects design. Direct isotonic object motion (dragging) is used as interaction technique; independent variable of the experiment is the display condition (multi-viewpoint image with tracked picking ray, and fully head-tracked). Dependent variable of the experiment is the docking time.

Subjects were asked to drag a 0.3m long L-shaped reference object into a target object by selecting the reference object and dragging it (with pressed Stylus button) into the target. Reference and target objects are placed at about 0.5m distance, well within the volume of reach of the subject.

When the reference object and the target are aligned (within 0.05m and 10º) the reference object highlights. Releasing the aligned reference object completes the task. A run consists of 10 repetitions; subjects would do two runs per condition.

Subjects receive brief oral instructions and were instructed to complete the docking as accurately and as fast as possible. They have about 5 minutes to familiarize themselves with the task in the fully head tracked condition. After executing the test, we would conduct a 5-minute interview. Total test time per subject was about 20 minutes.

4.3 Results and Discussion

A paired samples t-test was run to determine significant within subjects differences in performance between the three display conditions for the selection task. Table 1 shows mean values for selection time and target re-entry per trial for all twelve subjects (lower values are better). Table 2 shows the paired samples test results with mean value ratios comparing non head-tracked (NoHT) to head-tracked (HT) and multi-viewpoint image (MV) to the head-tracked condition, t-values and (2-tailed) significance.

Mean task completion time for selection without head tracking in the near condition is 0.75s while mean task completion time for multi-viewpoint and head-tracked condition are 0.43s and 0.46s respectively. Paired differences for the near condition show non head-tracked selection to be 63% slower than head-tracked, target re-entry is over 3x higher (both results are significant with t(11)=8.07 and t(11)=5.99, p<0.01). Multi-viewpoint selection is 6% faster than head-tracked (significant with t(11)=-2.87, p<0.05), while target re-entry per trial is slightly (0.06) higher (NS with t(11)=2.17, p>0.05).

Table 1. Ray Selection: Mean performance

	NoHT	MV	HT
Near			
Selection Time	0.75	0.43	0.46
Target Re-entry	0.44	0.19	0.13
Far			
Selection Time	0.85	0.53	0.59
Target Re-entry	0.35	0.21	0.18

Table 2. Ray Selection: Paired samples difference

	NoHT vs. HT			MV vs. HT		
	Mean	t(11)	Sig.	Mean	t(11)	Sig.
Near						
Selection Time	0.29	8.07	.000	-0.03	-2.87	.015
Target Re-entry	0.31	5.99	.000	0.06	2.17	.052
Far						
Selection Time	0.27	6.21	.000	-0.06	-2.88	.015
Target Re-entry	0.17	2.54	.027	0.03	0.54	.603

Mean task completion time for selection without head tracking in the far condition is 0.85s while mean task completion time for multi-viewpoint and head-tracked condition are 0.53s and 0.59s respectively. Paired differences for the near condition show non head-tracked selection to be 58% slower than head-tracked, target re-entry is 2x higher (again both results are significant with t(11)=6.21, p<0.01 and t(11)=2.54, p<0.05). Multi-viewpoint selection in the far condition is 10% faster than head-tracked (significant with t(11)=-2.88, p<0.05), while target re-entry per trial is not significantly different (t(11)=0.54, p>0.05).

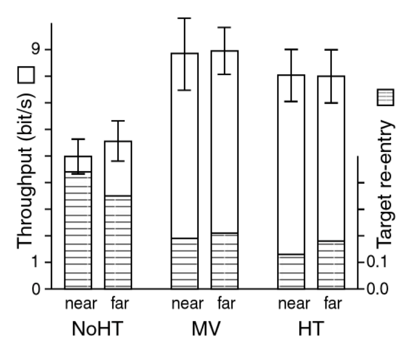

Figure 4. Ray Selection: Throughput and target re-entry

It is remarkable that while the difference between non head-tracked selection and multi-viewpoint as well as head-tracked selection is slightly lower in the far target condition, the basic difference in performance between near and far condition is fully explained by the respective index of difficulty. Figure 4 shows mean performance and 95% confidence interval as throughput in bit/s for the near and far condition, as well as the corresponding target re-entry rate (dashed). It is interesting to note, that throughput and target re-entry rate improve significantly for non head-tracked viewing (NoHT) in the far condition. This agrees with the general observation that closer targets are increasingly difficult to reach without head-tracking.

Compared to non head-tracked interaction, all subjects performed better with multi-viewpoint rendering (MV) and with head-tracking (HT). Subjects were about 45% to 60% slower for ray selection without head-tracking than with multi-viewpoint rendering or fully head-tracked. Although measured times and task setup are different, these results agree with results by Steed and Parker [19] who found ray selection of small objects with non head-tracked interaction to be 40% slower than full head-tracking in a CAVE.

During initial pre-testing of the docking task, we would get a very high number of incomplete task executions (defined as docking times far outside of standard distribution or explicit aborts of the task by the subject) in the non head-tracked condition. It turned out that the docking task was *impossible to execute*, since the parallax offset in the non head-tracked condition would place the point of interaction far outside of the user's volume of reach, making it impossible to properly rotate the reference object and to reach certain positions of the target object. In other cases, the task was very hard to execute and participants would become frustrated quickly. This result agrees with observations of [19]. We have therefore limited the study to comparing multi-viewpoint and the fully head tracked conditions.

Table 3. Docking Task: Mean performance

	NoHT	MV	HT
Docking Time	N/A	2.81	2.64

Mean task completion time for docking in the multi-viewpoint image condition was slightly slower at 2.81s compared to 2.64s for the fully head-tracked task (Table 3). This difference in task performance is not significant ($t(7)=0.68$, $p>0.1$).

When working with the non head-tracked condition for the first time, most subjects would make spontaneous negative comments, like "this is broken", "oops!" or "this is so weird".

In the interview, subjects rate the "ease of task execution" at 2.2 for the non head-tracked, and at 4.8 (on a five point Lickert scale) for both, multi-viewpoint and head-tracked condition. The "naturalness of the mapping" was rated at 1.4 for non head-tracked, 4.8 for multi-viewpoint and head-tracked condition. The "quality of stereoscopic viewing" was rated at 3.6 for non head-tracked, 3.9 for multi-viewpoint and 4.1 for head-tracked viewing (not significant with $t(19)=-1.49$, $p>0.1$). From these ratings it is clear that non head-tracked interaction is not delivering good performance for first-person interaction. The small, not significant rating difference for "quality of stereoscopic viewing" between multi-viewpoint images and fully head-tracked stereo is encouraging, considering that multi-viewpoint images use omnistereo rendering and do not respond to user head motion. This indicates that omnistereo rendering provides good stereoscopic image quality to non head-tracked viewers in a panoramic projection-based display system.

Usability testing confirms that non head-tracked interaction is inferior to multi-viewpoint and head-tracked interaction and does not allow full spatial interaction in a conventional projection-based virtual environment. We can substantiate that multi-viewpoint images allow comparable and sometimes even slightly better performance than single-user head-tracking and are therefore a highly suitable for multiple-user interaction in a projection-based display. We speculate that the superior performance of multi-viewpoint images for ray selection is due to lag-induced "swimming" of the scene in the head-tracked image, missing in the static multi-viewpoint image. At close range however, in particular for the docking task, performance with full head-tracking may benefit from motion parallax. Multi-viewpoint images enable each user in a projection-based display to execute direct object manipulation tasks with a similar performance then a fully head tracked user. This allows us to effectively use direct object manipulation techniques in our interface paradigm.

5. VRGEO DEMONSTRATOR

Based on multi-viewpoint image rendering, we have developed and evaluated a demonstrator application that allows multiple users to work together in a conventional projection-based virtual environment. The co-located VRGEO Demonstrator serves as a demonstrator and research tool for the VRGEO project, a cooperation between a consortium of companies from the oil-and-gas industry and Fraunhofer IMK. The demonstrator supports the collaborative analysis of complex geoscience surfaces and volume data. 3D seismic cubes are analyzed with volume rendering lenses and by placing and annotating 2D texture slices and clipping planes in the 3D volume.

We introduce spatially tracked PDAs to the interface of the VRGEO Demonstrator, to implement a common private interface for each user. The primary motivation for introducing PDAs as an interface into our collaborative virtual environment is the same as for Myers et al. [12] who have introduced PDAs into Single Display Groupware (SDG) systems: The PDA as personal device allows us take advantage of the fact that users are familiar with the device and have already learned the interface paradigm outside of our environment, enabling all users to handle the same interface. The PDA serves as an additional individual and private display for each user, introducing the separation of public and private data into our virtual environment (Figure 1). It also solves the problem of separating the representation of the application state and the individual contexts and modes for each user [12] by allowing us to put all the individual application state information on each user's PDA interface.

5.1 Co-located Multi-user Interface

Multiple workspaces–in the case of the VRGEO Demonstrator boxes, each containing one geoscientific volumetric data set–enable users to spread out the data over the whole display and to make better use of the large display surface (Figure 5). Inside a box visualization tools like volumetric rendering lenses or texture slices allow to view and analyze different aspects of the data set, set markers and take snapshots. The boxes work as spatial separators and allow users to arrange and partition different visualizations, allowing users to easily arrange and grab a coherent part of the scene and move it next to another for comparison.

Figure 5. A group of users using multiple workspaces in a 240° i-Cone™ display

The tracked PDA is used as a pointing device for 3D object selection by ray casting, extending a picking ray from the tip of the PDA. We use multi-viewpoint rendering for the selection ray, aligning, from each individual user's perspective, the position and pointing direction of the ray with the tip of each user's PDA.

For selection, the user points his PDA at a virtual object (Figure 6) and clicks the top-left PDA button. As a result of 3D object selection, the context on the 2D PDA interface switches, and displays the selected object with the corresponding interface pane acting as a context menu. The 3D object selection has exactly the same effect as selecting the corresponding tabbed interface pane on the PDA and making a selection of the object by name in a list box.

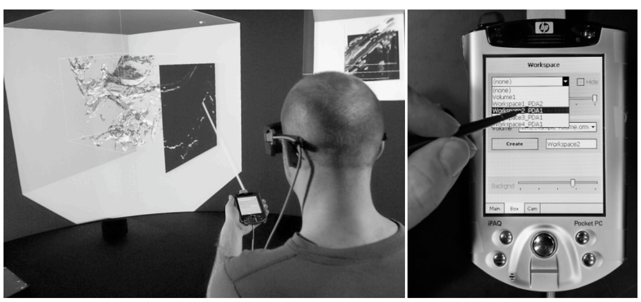

Figure 6. Ray selection vs. list box selection.

In the interface, we want enable the user to place objects at a comfortable viewing distance and spread them out over a large field of view. In a collaborative application, we cannot use travel to individually move larger distances inside the virtual environment since this would disturb other users (similar to the experience of collaboratively browsing a rotating postcard stand). Therefore, users need to be able to perform interaction and object motion at a distance with minimum effort.

Whenever the selected object is beyond the user's immediate reach (more than 0.6m from the user), we switch from direct object motion to a virtual motion technique we call Scaled Grab. Scaled Grab combines image plane selection and motion techniques and is similar to world-in-miniature (WIM) object manipulation [22]. Unlike Mine's Scaled-world Grab [11], which scales down the world to bring the selected object within reach of the user, Scaled Grab scales up the user's range of hand motion, to extend to the selected object. In this respect it behaves like a WIM, but without introducing an explicit miniature representation of the object, using the PDA as a handle on the selected object instead.

Using the multi-viewpoint technique we can warp the tracked position and orientation of each user's PDA to determine the view-frustum of a virtual camera in the virtual scene. This introduces the PDA as a spatial display into the virtual environment, similar to a Chameleon-type handheld display interface [7]. In our application, we use this technique to render an image to the PDA, using it as a virtual camera to provide a natural and direct interface to take snapshots of the virtual environment. The PDA screen now acts as the finder, reacting to the orientation and position of the PDA in the same way as a real camera would.

5.2 Experiences

We have presented the VRGEO Demonstrator on numerous occasions to groups of three to eight visitors. In two 60-minute evaluation sessions, four members of the VRGEO consortium, representing several mayor oil companies, have used the demonstrator. These evaluations sessions have retuned the most valuable feedback. In the current set up, because of limitations with the Polhemus Fastrack tracking system (chaining two units), we are limited to four fully tracked PDAs. We have experimented with using an additional non-tracked PDA, but users did reject the non-tracked device because it "is not the full interface".

There is practically no need to explain the interface of the application at all. Most visitors would grab the PDA and immediately start exploring the interface on their own. We would only demonstrate the basic GUI and explain the use of the top-left PDA button as the select/execute button. As expected, learning of a new interface in a co-located environment is much more relaxed than in a single user environment. New users would take their time to look and browse the interface, not feeling rushed even in a demo situation. We would frequently observe users discussing functionality and helping each other with the interface.

We did not receive any negative feedback on the ray-based object selection and direct object manipulation using the multi-viewpoint technique. Scaled Grab has proven to be very effective and was completely transparent to the users. Most users were completely unaware that there was something special going on for object manipulation at a distance, until we switched the scaling of the user's hand motion off. Typically, users would handle the PDA in their non-dominating hand, to be able to use the PDA GUI with the pen in a normal fashion. For some users this would lead to problems with the 3D PDA interface since they had to handle ray selection and object motion with their non-dominating hand. Although we have not seen severe problems with this issue, the interface seems to favor ambidextrous users.

Originally, with all users of the application handling an identical interface, we expected problems for users to identify without doubt, which interaction elements in the shared display were his, and which elements belonged to other users. For each user, only the interaction elements directly associated with him will align with his interaction device. From his point of view, because of parallax, the interaction elements belonging to other users will not align with their respective devices. Initially, we experimented with color-coding the interaction elements and colored marks on the devices to establish a connection between the device and the representation. However, for two reasons we have seen no irritation about the mapping between interaction elements and devices, making this color coding obsolete: Users are able to observe the alignment between their device and the virtual representation of the device from their point of view. Only the associated interaction elements will be aligned correctly, other user's interaction elements are not aligned. Second, since the interaction device is dynamically tracked, the motion of a user's hand, his interaction device and the visual interaction elements are immediately the same. This is a very strong and natural cue for a connection between a device and a visual representation that is immediately picked out.

In our evaluation scenarios, it was difficult to actually observe active collaborative behavior. With the oil-and-gas experts we could see situations where one user was moving and turning the data set around, while another user would adjust the color palette of the same volume to segment out new structures. With non-experts we would observe more individual viewing of the data and exploration of the interface and less interaction.

Overall, using the i-Cone in a collaborative fashion delivers a very different experience than the conventional single-user paradigm. Feedback about having the identical interface for each user in the system was enthusiastic, with many visitors remarking that they like to feel in charge over the whole application.

6. CONCLUSION AND FUTURE WORK

We have introduced an interaction paradigm for co-located collaboration in large projection-based display systems. Usability testing shows that multi-viewpoint images are an effective technique to give every user in a large projection-based display a first-person interactive experience for spatial interaction tasks. Interaction performance for ray selection and direct object motion in a docking task with multi-viewpoint images is comparable to interaction with full head-tracking.

Based on the concept of SDG systems, we introduce tracked PDAs as personal interface device for each user. Informal observations show that the introduction of co-located multi-user collaboration improves the overall interactivity of the virtual environment and delivers a full first-person experience to every user in the system. Despite some possible ergonomic problems with the use of the tracked PDAs, the introduction of common devices and common device metaphors, together with the ability of co-located collaboration has a very positive effect on the learning experience of new and casual users.

In the future we will use a wireless optical tracking system, allowing us get rid of all the wires and to support a larger number of active users. With a clip-on mechanism for the optical tracking target, users will be able to bring their own PDAs into a virtual environment session. We would like to develop a more complex application scenario that encourages more immediate collaboration between users.

7. ACKNOWLEDGEMENTS

Part of this work was conducted while the author was with Fraunhofer IMK in Sankt Augustin, Germany. The author would like to thank Sascha Scholz for his work on the implementation of the usability test and the VRGEO consortium for financial support of this research.

8. REFERENCES

[1] Agrawala, M., Beers, A., Fröhlich, B., Hanrahan, P., McDowall, I., and Bolas, M.: The Two-user Responsive Workbench: Support for Collaboration through Individual Views of a Shared Space. Proc SIGGRAPH '97. ACM Press (1997) 327–332

[2] Arthur, K., Preston, T., Taylor, R., Brooks, F., Whitton, M., and Wright, W. Designing and Building the PIT: a Head-Tracked Stereo Workspace for Two Users**.** Proc. 2nd International Immersive Projection Technology Workshop (1998), May 1998

[3] Butz, A., Höllerer, T., Feiner, S., MacIntyre, B., and Beshers, C. "Enveloping users and computers in a collaborative 3D augmented reality." In *Proc. IWAR '99 (IEEE and ACM Int. Workshop on Augmented Reality),* San Francisco, CA, October 20–21, 1999, pp. 35–44.

[4] Billinghurst, M., and Kato, H., "Collaborative mixed reality" in *Mixed reality: Merging real and virtual world*s, Y. Ohata and H. Tamura (Eds.), Co-published by Ohmsha and Springer-Verlag, 1999, 261–284.

[5] Bimber, O., Fröhlich, B., Schmalstieg, D., and Encarnação, L.M. The Virtual Showcase. IEEE Computer Graphics & Applications, vol. 21, no.6, pp. 48-55, 2001

[6] Cruz-Neira, C., Sandin, D., DeFanti, T., Kenyon, R., and Hart, J. The CAVE Audio-Visual Environment. ACM Trans. on Graphics 35,1 (1992) 65-72

[7] Fitzmaurice, G. Situated Information Spaces and Spatially Aware Palmtop Computers. Communications of the ACM 36,7 (1993) 38-49

[8] Greenberg, S., Boyle, M. and LaBerge, J.: PDAs and Shared Public Displays: Making Personal Information Public, and Public Information Personal. Personal Technologies, 3, 1 (1999)

[9] Kitamura, Y., Konishi, T., Yamamoto, S., and Kishino, F.: Interactive Stereoscopic Display for Three or More Users. Proc SIGGRAPH 2001. ACM Press (2001) 231–240

[10] MacKenzie, I., Kauppinen, T., and Silfverberg, M. Accuracy measures for evaluating computer pointing devices. Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI 2001, 9-16.

[11] Mine, M., Brooks, F., and Sequin, C. Moving Objects in Space: Exploiting Proprioception in Virtual Environment Interaction. Proc Siggraph'97 (1997)

[12] Myers, B., Stiel, H., and Gargiulo, R.: Collaborations using multiple PDAs connected to a PC. Proc ACM CSCW'98. ACM Press (1998) 285–294

[13] Regenbrecht, H., and Wagner, M.: Interaction in a collaborative augmented reality environment. Proc CHI 2002. ACM Press (2002) 504–505

[14] Rekimoto, J.: Pick-and-Drop: A Direct Manipulation Technique for Multiple Computer Environments. Proc ACM UISF'97. ACM Press (1997) 31–39

[15] Rekimoto, J.: A Multiple Device Approach for Supporting Whiteboard-based Interactions. Proc CHI'98. ACM Press (1998) 18–23

[16] Simon, A., and Göbel, M., "The i-Cone™ – A Panoramic Display System for Virtual Environments," Pacific Graphics '02, Conference Proceedings, September 2002, pp. 3-7.

[17] Simon, A., Smith, R., and Pawlicki, R., OmniStereo for Panoramic Virtual Environment Display Systems, Proc. IEEE Virtual Reality 2004 (VR'04), Chicago, Illinois, March 27 - 31, 2004

[18] Simon, A., Scholz, S. Multi-Viewpoint Images for Multi-User Interaction. Proc IEEE Virtual Reality'05 (2005)

[19] Steed A., Parker, C., 3D Selection Strategies for Head Tracked and Non-Head Tracked Operation of Spatially Immersive Displays, Proc. 8th International Immersive Projection Technology Workshop, May 2004

[20] Stewart, J., et al.: When Two Hands Are Better Than One: Enhancing Collaboration Using Single Display Groupware. Proc SIGCHI'98. ACM Press (1998) 287–288

[21] Stewart, J., Bederson, B. and Druin, A.: Single Display Groupware: A Model for Co-present Collaboration. Proc ACM CHI'99. ACM Press (1999) 286–293

[22] Stoakley, R., Conway, M., Pausch, R. Virtual reality on a WIM: interactive worlds in miniature. CHI'95 (1995), 265-272.

[23] Szalavari, Z., Schmalstieg, D., Fuhrmann, A., and Gervautz, M., Studierstube: an environment for collaboration in augmented reality, Virtual Reality, 3(1), pp. 37–48. 1998.

[24] Tramberend, H.: Avocado: A Distributed Virtual Reality Framework. Proc. IEEE Virtual Reality '99, 14-21, 1999

[25] Watsen, K., Darken, R., and Capps, M., "A Handheld Computer as an Interaction Device to a Virtual Environment," presented at Third Immersive Projection Technology Workshop, 1999.