EUROGRAPHICS 2007 / D. Cohen-Or and P. Slavík
(Guest Editors)

Volume 26 (2007), Number 3

Consistent Viewing and Interaction for Multiple Users in Projection-Based VR Systems

Gerwin de Haan[†], René Molenaar, Michal Koutek and Frits H. Post

Delft University of Technology, The Netherlands

Abstract

In projection-based Virtual Reality (VR) systems, typically only one headtracked user views stereo images rendered from the correct view position. For other users, who are presented a distorted image, moving with the first user's head motion, it is difficult to correctly view and interact with 3D objects in the virtual environment. In close-range VR systems, such as the Virtual Workbench, distortion effects are especially large because objects are within close range and users are relatively far apart. On these systems, multi-user collaboration proves to be difficult. In this paper, we analyze the problem and describe a novel, easy to implement method to prevent and reduce image distortion and its negative effects on close-range interaction task performance. First, our method combines a shared camera model and view distortion compensation. It minimizes the overall distortion for each user, while important user-personal objects such as interaction cursors, rays and controls remain distortion-free. Second, our method retains co-location for interaction techniques to make interaction more consistent. We performed a user experiment on our Virtual Workbench to analyze user performance under distorted view conditions with and without the use of our method. Our findings demonstrate the negative impact of view distortion on task performance and the positive effect our method introduces. This indicates that our method can enhance the multi-user collaboration experience on close-range, projection-based VR systems.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computing Methodologies]: Computer GraphicsVirtual Reality;

1. Introduction

Large screen projection displays, such as the Virtual Workbench, invite direct collaboration between a small group of people. The shared, interactive viewing of information on a single screen with multiple users facilitates natural communication on the virtual environment. The various 3D depth cues used in rendering a correct 3D image of the information are only optimal when observed from a single viewpoint. Users that are not close to this optimal viewpoint perceive a distorted graphical representation, which has a negative influence on system usability and collaboration.

Two general approaches have been proposed to enhance collaborative use of projection-based Virtual Reality (VR). First is the use of additional tracking and display hardware to generate correct (stereo) images for each user individually [BMC04]. This approach involves costly extensions to standard projection-based systems and image quality often suffers from limitations in image separation techniques. The second approach applies extra tracking hardware and special rendering techniques to cope with image distortion [Sim07]. Image rendering and depth cues are adapted to reduce the overall perceived distortion.

Although this solution does not provide a completely correct image for all users, we still expect it to be very usable in real collaborative scenarios. Furthermore, as up-scaling tracking hardware for more users is easier and more cost-effective than additional display hardware, it would provide an attractive option for extending new and existing projection-based VR systems. For these reasons, we chose to analyze, extend and evaluate this solution for use on close-range VR systems, and, more specifically, on our Vir-

---

[†] e-mail:g.dehaan@tudelft.nl

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

Figure 1: *Demonstration of our method in a two-user scenario at the Virtual Workbench. The scene is rendered from the viewpoint of the left user (left image), the right user (right image) or a dynamic, average viewpoint (middle image). Although objects may not appear in the correct perspective for either user in the averaged viewpoint, interaction with the distorted environment is still possible, as each user has personal, corrected interaction tools.*

tual Workbench system, where users stand relatively further apart, have a very different, off-axis viewing perspective and use direct and co-located interaction. Such a collaborative two-user scenario at the Virtual Workbench is illustrated in Figure 1.

The contribution of this paper consists of three main parts. First, we analyze of the distortion effects and their influence on 3D interaction and co-location. Second, we present a method to minimize view-distortions of individual users while maintaining co-location for effective interaction. Finally, we present a user study performed with our Virtual Workbench, in which test subjects performed 3D interaction tasks using both direct and remote techniques. The experimental results quantify the negative impact of view distortion and the corrective effects of our method. This paper is constructed as follows: We first describe related work in section 2. Then, in section 3, we describe the causes of distortion effects in projection-based VR and our approach for a solution. This is followed by a technical description in section 4. Our method is evaluated during a series of user experiments, of which relevant results and an analysis are presented in section 5. Finally, we draw conclusions and discuss possible future extensions in section 6.

2. Related Work

A general challenge for multi-user 3D stereo displays is to provide a separate, correct image for each eye of each user [BMC04]. Besides tracking each user's head, the key challenge here is the separation of all the images. For projection-based VR system, solutions include optical filtering (e.g. polarization filters), time multiplexing (e.g. shutter glasses), multiple screens, or a combination of these. The techniques can be used to extend existing, single-user systems, or to construct new multi-user systems from scratch.

The use of time multiplexing is demonstrated in the *Two-User Workbench* system [ABM*97] and in a multi-screen immersive environment [BLCN02]. Limitations included decreased image brightness, crosstalk and image flicker. A hybrid solution is demonstrated in [FHH*05], where polarization filtering and time multiplexing was carefully combined to achieve better image quality for more users. Separate screens, masks and mirrors can be used to create a shared physical space for all users, for examples the *PIT* [APT*98] and the *Virtual Showcase* [BFSE01]. Technical limitations in display hardware and physical installation issues limit the applicability and scalability of extending existing VR systems for collaborative use for multiple users.

Simon et al. describe *multi-viewpoint merging*, an alternative approach to cope with multi-user interaction on a single display [SS05] [Sim07]. Instead of generating a completely correct 3D image of the entire virtual environment for each user, the scene is rendered from a static, central viewpoint. Only some elements in the scene, like interaction tools, are displayed correctly for its corresponding, headtracked user by correcting the visible projection of these objects. They report that, although many objects are visually distorted, this solution still remains usable for collaborative work in VR systems. At the cost of image distortion and object deformation, existing systems can be extended for multiple users by tracking the extra users. These interesting properties were our main motivation to extend this work at three points.

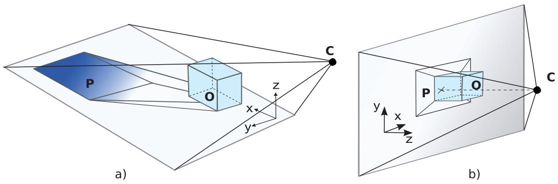

Figure 2: *Difference in projection transformations: The Virtual Workbench (a) has an asymmetric off-axis frustum, while a Virtual Wall (b) has a symmetric on-axis frustum.*

First, as the Virtual Workbench differs greatly from the panoramic screen used in Simon's study, we suffer from larger distortion effects. In close-range, tabletop VR systems users stand relatively further apart and have a very different, off-axis viewing perspective (see Figure 2), which is sensitive to head motion. Also, virtual objects are mostly within hand reach, and direct, co-located manipulation techniques are preferred. Therefore, we perform a problem analysis to investigate the source and effects of this distortion on interaction in more detail, as [WHR99] did for eye-separation. Second, roles among users [HSAW05] as well as proper division of tasks are important factors that influence the success of collaboration and task performance. We strive for a flexible solution to match the quality of viewing and interaction with the task and roles in mind for each user. For this we investigate the advantages and disadvantages of different solutions and camera options. Third, the question arises how much distortion can be tolerated in various collaborative VR scenarios. Only a few user studies exist concerning multi-user setup and distributed multi-user VR systems, for example [HSS05]. With the benefits and limitations of different quantitative and qualitative evaluation in mind, we conducted a user study on our Virtual Workbench.

3. Analysis and Approach

We performed a usability analysis to investigate current problems and possible solutions [Mol07]. Figure 3 illustrates the basic problems of a collaborative two-user scenario on a regular workbench system, while the usability claims and the most important advantages and disadvantages are summarized in a table.

3.1. Problem Description

In the classic, single-user scenario, only one user (User1) experiences both the stereo and motion parallax depth cues correctly. For another user (User2) viewing the same screen, interaction can become much more difficult or even impossible.

First, the stereo fusion may be lost. When the User1 tilts his head, the stereo parallax remains correct only for him, but becomes very unpleasant for User2. Depending on the head tilting of User1 it may even become impossible for User2 to perceive depth, as there will be no intersection between the lines of sight to the two projection points that are to be fused.

Second, if User2 does perceive stereo images, he still experiences both point of view distortions and motion distortions. Point of view distortions are the result of seeing the projections from a wrong point of view (see Figure 3); an object may appear sheared and at the wrong location, and interaction tools do not appear co-located either. Also motion distortions influence interaction and are the result of the absence or presence of parallax effects. Head movements of tracked users result in image changes, while non-tracked users do not experience the motion parallax effect when they should.

	General claims if User1 - headtracked; User2 not
+	Face to face and non-verbal communication possible
+	Shared use of system resources
+	Several users are able to add value to the cooperation
-	Hardware supports only one (active) user
-	Tracked glasses and interactions devices may need to be switched, this interrupts the session
-	Some movements of User1 cause eye strain for User2
-	Some movements of User1 cause in-fusible images for User2 (see Figure 3)
	Point of View Distortion for User2
-	Objects will appear sheared
-	Objects will appear at a wrong location
-	Users need to compensate for distorted view during interaction
	Motion Distortion for User2
-	User1 head movement will cause the objects to move, as seen from User2's point of view
-	Head movement User2 will not update the world
-	Interaction is disturbed by User1's head movements

In our research we explored several solutions with alternate forms of display and headtracking as well as direct, close-range manipulation. In this paper, we only concentrate on one approach which focuses on the interaction problems that occur when looking at a single scene rendered from an incorrect viewpoint.

In our effort to find a suitable solution for multi-user use of a single rendering, the overall quality of usage and the performance of additional tracked users is taken into account. One goal is to provide equal opportunities for all users, causing two users to be able to collaborate equally and also have better sense of what the other user is seeing [HSAW05]. To support multiple users on one system we use alternative camera models, which are presented in this section. Besides choosing a camera model we can also solve some interaction problems by calculating proper locations of interaction elements of other tracked users.

3.2. Alternative Camera Models

Some viewing problems can be eliminated by altering the effects of headtracking on the rendering viewpoint. In some scenarios, eliminating headtracking completely can make the system more accessible for multiple users. A main advantages is that adding additional viewers will not be limited by hardware limitations. Images remain stable and fixed, so no user will experience the motion swimming distortion and there will be no in-fusible images caused by the head movements of a single user. Of course, there is only one specific position to have the correct perspective, and the elimination of motion parallax takes away important depth information. In general, this solution is best for larger audience

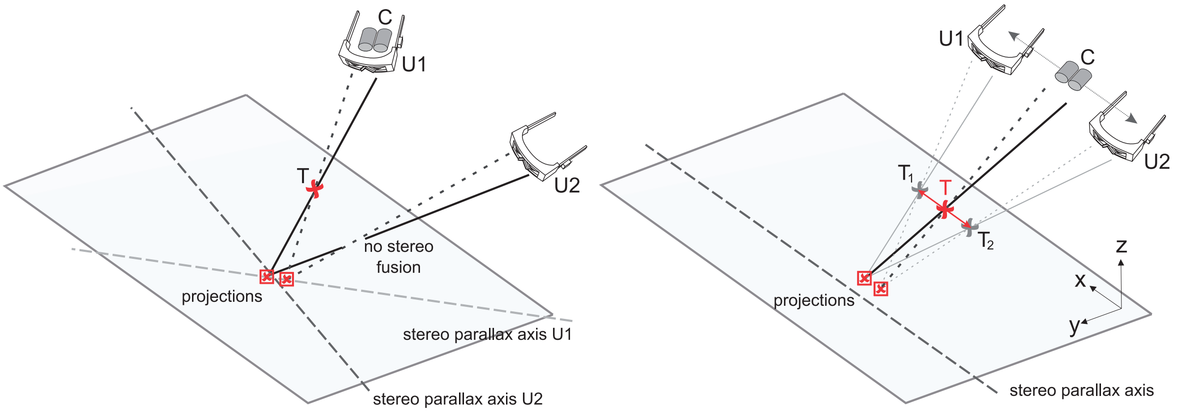

Figure 3: *Two-user viewing problems. Images are rendered for User1 ($U1$). Depending on the head pose of both users, the stereo images may be difficult or impossible to fuse for User2 ($U2$) (left image). This is caused by $U2$ having a different stereo-parallax axis. To avoid parallax differences between users, the cameras can be constrained to be parallel to the x-axis. $U2$ will now perceive stereo, although point T is perceived at location T_2 (right image)*

presentations (e.g. a panoramic screen [SS05]), and not for close-range, interactive work. For the Virtual Workbench, the elimination of all motion parallax takes away much of the depth perception. Also head tilting is a greater issue here.

Instead of totally eliminating all headtracking we can also partly constrain the headtracking. By only tracking head position and not the rotation, the headtilting problem can be reduced. Since the stereo images will be rendered with horizontal parallax only (some systems only allow horizontal parallax due to the passive filter technique used). Additional viewers will have less fusion problems when the headtracked user rotates his head. But the headtracked user is now unable to tilt head and keep images fusible for himself. Users will quickly learn however how to hold their head. As a result the images will remain more stable.

The distortion can be split equally between users, by placing the viewpoint at an average position between the multiple headtracked positions, see Figure 3. By using this *dynamic average viewpoint*, two users will experience half of the distortions, but also half of the motion parallax, which can still serve as an important depth cue. Both users will experience some motion swimming effect. Head rotations will not be taking into account, but there will be reduced possibility of eye strain and in-fusible images.

3.3. Viewpoint Compensation

When choosing an altered camera model, users will observe distortions in both perspective and motion parallax of the projections. The need for users to manually compensate for this when using interaction devices can be counteracted. By using the headtracking of a user, interaction can be based upon his point of view. This is accomplished by predistorting or *warping* the geometry of some (interaction) elements of the scene in such a way that they seem correct from his specific viewpoint. While these elements can be shown correct only for that user, the rest of the scene can be updated according to the active camera model, such as the dynamic average viewpoint. We call this *viewpoint compensation*. This is similar to the *multi-viewpoint merging* solution [SS05], where selection rays are displayed correctly for a specific user while the overall scene is projected for a static viewpoint. In our dynamic average viewpoint approach however, we also need to apply continuous compensation to counteract additional motion swimming distortions. In this way, perception and action can be made consistent for every tracked user even though the overall scene may appear distorted by perspective and motion parallax effects, see Figure 4(left). This warping process and its implementation is described in the following section.

4. Method

In this description of our method and its implementation in our VR system, we distinguish between the consistent viewing of the scene and consistent interaction with objects in the scene. We define consistency as the agreement between the user's actions and his perceptions. To clarify this in the context of distorted spaces, we distinguish three different coordinate spaces. Figure 5a illustrates the three coordinate systems used. In the *tracker space*, the tracker hardware registers the actions performed by the users. Correct alignment and calibration allow a direct translation of these measurements to the *scene space*. This scene space is rendered and projected onto the screen. Based on these projections, each user makes a mental 3D reconstruction of the scene, which we call the *perceived space*.

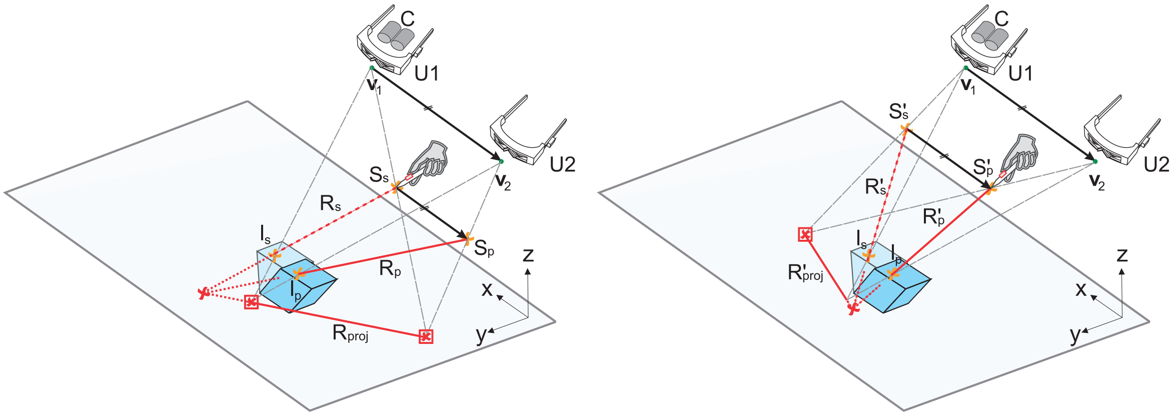

Figure 4: *View distortion during ray-casting (left) and with viewpoint compensation (right). The camera is at* $U1$*'s viewpoint, subscripts* $S, P, Proj$ *stand for in scene, perceived and projected. User* $U2$ *holds a stylus and tries to shoot a ray on the cube. In the left image,* $U2$ *experiences inconsistency: based on* R_{proj}*, he perceives* R_P *as the ray coming from* S_P*. In the right image, we correct for viewing distortion by pre-transforming the ray to* R'_S*. From the new* R'_{proj}*,* $U2$ *now correctly perceives the ray* R'_P *as shooting out of his stylus* S'_P*, and can consistently point at the perceived cube.*

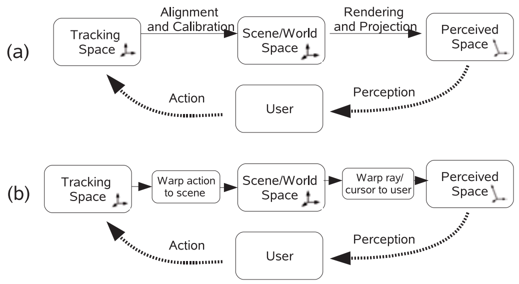

Figure 5: *Relations between coordinate spaces: (a) the user observes the scene in perceived space, while performing actions in tracking space. (b) When perceived space and tracking space are not in agreement, corrections on both visible geometry and the performed actions are needed.*

In a normal situation, the rendering viewpoint directly corresponds with the user's head position. Therefore, the tracking space and the perceived space directly agree, providing a co-located virtual environment. When these projections are observed from a different position, the mental 3D reconstruction of the scene or perceived space does not agree anymore with the tracker space. Since we know how this tracked user perceives the scene, we can apply appropriate calculations to make actions and perception consistent again. These calculations consist of corrections for viewing and corrections for interaction, see Figure 5b.

4.1. Consistent Viewing

Any element of the scene that is directly related to a certain user can be transformed to match the perspective of the user. We will describe necessary algebraic calculations for this *warping* process.

We first transform the interaction rays, cursors and other private elements to the perception space of the user, see Figure 4. In this way, we can display these elements co-located with users interaction devices. For example, we can make the interaction ray seem to shoot out of the stylus, or correctly display widgets on our *Plexipad*, a handheld transparent prop.

When using this method, all interacting users need to be headtracked. In our Virtual Workbench system, we use six 6-DOF sensors, to fully support two users. Users are headtracked and they have a stylus and a Plexipad. The headtracker is used to register the users viewpoint positions $\mathbf{v}1$ and $\mathbf{v}2$, see Figure 4. The current positions of $\mathbf{v}1$ and $\mathbf{v}2$ are used to calculate the amount of perspective distortion for each user. When for example the images are rendered from a camera at viewpoint $\mathbf{v}1$, $U1$ correctly perceives the scene while $U2$ observes a sheared scene.

We use a *warping matrix* to properly transform those elements in the scene that are associated with a specific user. This warping matrix allows us to transform points and geometry between the perceived space and the scene space. Figure 6 illustrates this situation. User $U1$ has a correct view and observes the Z axis as being perpendicular to the ground plane, pointing upwards. His perception of the scene space can be described by an orthonormal coordinate system matrix, which is in this case equal to the identity matrix, as all the axis base vectors are orthogonal and have a unit length. The view of the second user $U2$ however, suffers from a shearing distortion. He experiences this axis as pointing into

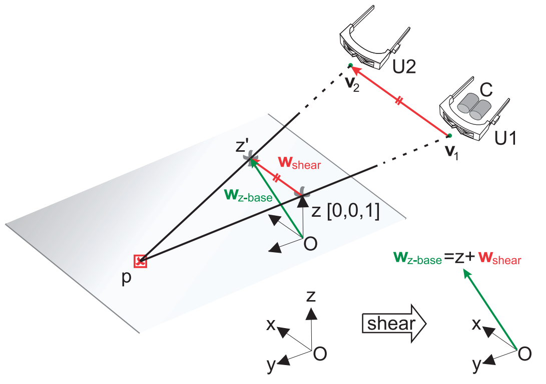

Figure 6: *Construction of the sheared coordinate system: The axis vectors X and Y are not distorted. The Z axis suffers from shear distortion. Point Z is perceived at Z'.*

a sheared direction. The X and Y base vectors remain unaffected, but the Z base vector is affected.

The vector $\mathbf{w}_{shear}$ defines the magnitude and direction of the shearing and can be calculated from any point above the ground plane. Here, we use the base vector $\mathbf{z}[0,0,1]$ for convenience. As shown in Figure 6, its projection from viewpoint $\mathbf{v}1$ is the point $\mathbf{p}$. We calculate $\mathbf{w}_{shear}$, which is parallel to the vector $\mathbf{v}1\mathbf{v}2$, see equation 1. Then, we calculate position of the point $\mathbf{z}'$ in the perceived space of the user $U2$, see equation 2.

$$\mathbf{w}_{shear} = \frac{|\mathbf{z}-\mathbf{p}|}{|\mathbf{v}1-\mathbf{p}|}(\mathbf{v}1-\mathbf{v}2) \tag{1}$$

$$\mathbf{z}' = \mathbf{z} + \mathbf{w}_{shear} \tag{2}$$

We use the point $\mathbf{z}'$ to calculate the vector $\mathbf{w}_{z-base}$, which is the z base vector of the coordinate system of the second user, see equation 3.

$$\mathbf{w}_{z-base} = \mathbf{z} + \mathbf{w}_{shear} \tag{3}$$

The amount of the distortion for the whole scene can be encapsulated in the shearing matrix $\mathbf{M}_{shear}$, as shown in equation 4. Basically, we insert the vector $\mathbf{w}_{z-base}$ at the place of the normal z-base vector.

$$\mathbf{M}_{shear} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \multicolumn{3}{c}{\mathbf{w}_{z-base}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{4}$$

It is clear that in this approach, the shearing distortion affects the rendering of the individual vertices of geometry in the scene. The amount of shearing for each vertex is a function of the camera position $\mathbf{v}1$, the observer position $\mathbf{v}2$ and the z coordinates of the vertex. In general: the higher a point, the larger the shearing distortion, the larger the distance between $\mathbf{v}1$ and $\mathbf{v}2$, the larger the shearing distortion. The correction matrix $\mathbf{M}_{warp}$ for counteracting the distortion, is obtained by inverting $\mathbf{M}_{shear}$, see equation 5.

$$\mathbf{M}_{warp} = \mathbf{M}_{shear}^{-1} \tag{5}$$

This matrix can be used to *pre-distort* or *pre-warp* geometry in scene space. We implemented the warping function as a dynamically updating scene graph node, which can be used to easily correct distortion effects on various parts of the scene.

4.2. Consistent Interaction with Scene Objects

Having only a proper visual perception is not yet sufficient for consistent interaction from a different viewpoint. Although the perceived space is made to match the scene space, it does not directly match the tracking space that registers all interactions. Theoretically, these spaces can also be made to agree by applying a correction, similar to the previous section. However, implementation details of interaction techniques and scene graph hierarchies prevent this, and would result in unexpected deformations of objects. Interaction techniques typically use transformation matrices and quaternion algebra, but assume orthonormality of the input data. As soon as we would apply distortion corrections on tracker data, orthonormality would be lost. As a result, shearing aspects would interfere with normal calculations and would introduce unexpected deformation in objects in scene space. To avoid these distortions we describe the adjustments needed to use interaction techniques under distorted viewing conditions.

Consistent Selection

For object selection, we use both direct and remote (ray-based) techniques. The *direct selection* technique uses a single 3D point provided by the stylus. On account of possible viewing distortions, we need to multiply that point in percieved space S'_p by $\mathbf{M}_{shear}$ matrix to obtain the corresponding point in scene space S'_s. The *ray-based selection* technique uses both a point and a direction vector provided by the stylus, both of which have to be transformed from tracking space into scene space. The (unit) direction vector can be extracted from the orientation quaternion of the stylus $\mathbf{q}_{stylus}$. To avoid distortion effects that could occur by warping the quaternion directly, we divide the calculation of the ray in the transformation of two points. For this, we calculate the position of an arbitrary second point along the stylus ray. The two points, the stylus tip and the second point, will be converted into scene space. Based on these two points, the correct selection ray can be reconstructed in scene space.

Consistent Manipulation

For object manipulation, we also use both direct and remote (ray-based) techniques. Once the object has been selected, both techniques use the same algorithm for manipulations. The important adjustment here is the separate treatment of translation and rotation transformations. The reason for this is that the translation matrices remain pure translation matrices despite of shear distortion, for rotations this is not the case. A more detailed technical discussion on maintaining consistency during manipulation, and especially rotations, is given in [KMdHP07].

The pose of an object consists of a position (x,y,z) and orientation (quaternion $\mathbf{q}_{object}$). After object selection, we calculate the so-called *interaction point* I_s, see Figure 4. We store the distance vector R'_p between I_p and stylus position S'_p and the vector $\mathbf{v}_{OIs}$ between the objects origin and I_s. For direct manipulation, distance R'_p is zero, while for ray-based manipulation this corresponds with the length of the selection ray. During manipulation, we combine the stored $\mathbf{v}_{OIs}$ and R'_p with the new stylus position S'_p to translate the object to the new positions in the scene. To calculate the new orientation of the object during manipulation, we first store the quaternion difference $\mathbf{q}_{delta}$ between the initial object orientation and stylus orientation $\mathbf{q}_{stylus}$. Then, every manipulation step the current quaternion $\mathbf{q}_{stylus}$ is warped to match the intended rotation. In this way, we obtain $\mathbf{q'}_{stylus}$ in scene space. To calculate the new object rotation around the point I_s, $\mathbf{q'}_{stylus}$ is multiplied with $\mathbf{q}_{delta}$.

During manipulation two extra distortion effects become noticeable. First, in normal ray-based manipulation, the length R_s of the ray remains constant in scene space. However, distortion within the perceived space and head movements cause the amount of distortion in $\mathbf{M}_{shear}$, and thus the perceived length R_p of the ray, to constantly change. The correct length is calculated by taking the distance between the current stylus position S'_s and the unwarped interaction point I_s. Second, extra side effects occur when applying rotations on objects. Rotations affect the observed shape of the object, caused by the shearing that affects the object. Also, we can not directly apply rotational movements made with the interaction tools, since rotation angles change in sheared space. This is especially noticeable when the $\mathbf{M}_{shear}$ matrix is changed significantly and when upward and tilting movements are made at the same time. A detailed technical description of these extra distortion effects and its compensation are beyond the scope of this paper and are given in [KMdHP07].

5. Evaluation

A series of user experiments was performed for a qualitative and quantitative evaluation of our method. These experiments were performed on our Virtual Workbench system, a large tabletop stereo display combined with 3D tracking hardware. The table surface is slightly tilted at 10 degrees, measures 180 by 110 cm and provides room for multiple standing users. Stereo is provided by two projectors equipped with *Infitec* passive stereo filtering. We used two 6-DOF electromagnetic tracking sensors for each active user, one for headtracking and one for stylus tracking. The software for performing the experiments and evaluation was implemented in Python on top of our *iVR* software framework.

5.1. Experiments

The experiments of interest here were part of a larger set of two-user experiments, consisting of evaluation of various display configurations and the evaluation of interaction performance. Each experiment took about an hour for each user to complete. The first part investigates the activities of pairs of users working side by side at the Virtual Workbench in different roles. The second part analyzes the impact of distortion on the task performance and accuracy of co-located selection and manipulation interaction techniques for a single user.

The first part consisted of informal discussions on various two-user display, tracking and interaction scenarios. Participants were asked to interact with the VR scene, to think aloud and to give their opinions about the distortions and proposed solutions. The scene consisted of randomly placed colored cubes. Users explored interaction possibilities and communicated through questions like “What color is this cube?” and “Can you place that cube over here?”. In each scenario, the quality of interaction and communication was discussed and the different techniques were explained. After these explorations, users were asked to fill out questionnaires consisting of Likert scale questions and short answer questions.

The quantitative part of the experiment was performed by each participant individually. Here, the direct and remote selection and manipulation skills of the test subjects were measured under various conditions. A series of four tasks was performed: Object selection with the stylus cursor (*SC*), object selection with the stylus ray (*SR*), objects manipulation with the cursor (*MC*) and object manipulation with the ray (*MR*). Target objects were placed pseudo randomly and not yet visible in the VR scene, just above the tabletop of the Virtual Workbench. For the selection task, 12 small spheres were placed in a grid, while for the manipulation task six poles with a ring around them were used. Two poles were placed axis-aligned, while the remaining poles were placed diagonally over two or three axes. The selection task was performed as follows: First, the user placed the stylus inside a ‘stopwatch sphere’. When the user moved the stylus out of the sphere, one of the twelve target spheres appeared and the stopwatch started running. The stopwatch stopped when the target was selected and the stylus button pressed. Then, the user returned the stylus to the stopwatch sphere to continue with the next target. The manipulation task was performed in a similar fashion. Here, the user had to pick the ring on the start of the pole first, then hold the button and move the ring along the pole to its ending, see Figure 7

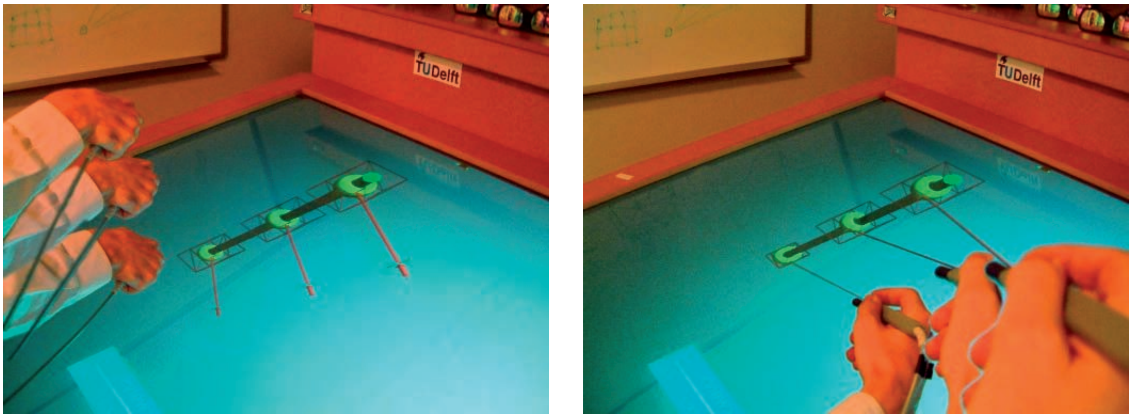

Figure 7: *Sequence of ray-casting manipulation task. The scene is rendered from User1's viewpoint and photographed from User2's viewpoint. In the left image, User1 sees the ray shoot from his stylus (although User2 does not) and moves a ring upwards along a vertical cylinder. In the right image, User2 performs exactly the same task from a different viewpoint. By using warping, the ray shoots from his stylus and he can move the ring diagonally along the distorted cylinder.*

5.2. Conditions

Twelve individuals participated in the experiments. For many subjects, this was their first VR experience. For the individual quantitative measurements, each of the four tasks was performed with eight repetitions under eight different conditions. The eight test conditions were constructed from variation in viewpoint configuration and the use of warping compensation: normal headtracked situation (*ht*), non-headtracked without viewpoint offset (*not ht*), non-headtracked with x cm offset from the correct viewpoint (*25*,*50*,*75*), and these same offsets again while viewpoint compensation for interaction was enabled (*25+w*, *50+w*, *75+w*). The offset distances roughly match the distances in the positioning behavior of the secondary user with respect to the primary user. These situations are: looking *over the shoulder (25cm)*, standing directly *shoulder-to-shoulder (50cm)*, and standing comfortably *side-by-side (75 cm)*. The targets for selection and manipulation appeared in random order in the scene and the sequence of sub-conditions was also chosen randomly each time. For each of the four tasks, the participant had a training period of two minutes under various, random viewing and interaction conditions. This training period was introduced to reduce the strong learning effect, which influences the performance of the individual tests.

5.3. Results

In this section we summarize the most important results on viewpoint compensation. A complete transcript of all experiments can be found in [Mol07]. The main quantitative results are summarized in Table 1 and Figure 8. Important results of the qualitative evaluation are given here.

Users' reactions were very positive during the qualitative evaluation. We see from questionnaire results that users acknowledged that the system should allow multiple users to be active, and that proper alignment of their private tools created a more workable and pleasant situation. The possibilities for head rotations for one user was indeed a source of discomfort for the other user. They did not find the introduced constraining of stereo-parallax very discomforting. Some users needed to get used to not tilting their head too much. Users strongly agreed that the possibilities to communicate about the objects displayed and work together were enhanced, and that viewing compensation is essential when using interaction devices under distorted viewing conditions.

Users responded extensively in the short answers and confirmed that we made good improvements for collaborative situations. They remarked that the uncorrected situation was sometimes frustrating, and that viewing compensation made tool behavior predictable again, making tasks easier. One user stated the results clearly: “Seeing your cursor in the right perspective gives a feeling of control, that the space around it is distorted is not such a problem”. Another user remarked that working with a wrong viewpoint is not that bad if it remains static, allowing users to adapt to the situation. A correct cursor made it a lot easier to adapt to a distorted viewpoint, and could also enhance the feeling of immersion. A steady image is more important than a correct image, and correct feedback is very important. Another user remarked that, without viewpoint compensation, it is nearly impossible to correctly collaborate, because one of the users will have difficulty operating interaction tools properly. Users did report some other, unexpected issues such as image blur at screen edges, light reflections and focus problems. For some, shorter users, targets would sometimes be outside the working area, preventing them from accomplishing a task nor-

Table 1: *Experiment results. Selection time mean (M), standard deviation (SD), and percentage offset of mean to the headtracked (ht) situation are given for eight different viewing conditions. Columns SC and SR are for selection, MC, MR, MCE and MRE for manipulation. Columns MCE and MRE indicate manipulation accuracy measurements in centimeters.*

	Cursor (SC)			**Ray-Based (SR)**			**Cursor (MC)**			**Ray-Based (MR)**			**Cursor (MCE)**			**Ray-Based (MRE)**		
	M(s)	SD	%	M(s)	SD	%	M(s)	SD	%	M(s)	SD	%	M(cm)	SD	%	M(cm)	SD	%
ht	2.0	1.16	+0	1.5	0.50	+0	2.1	0.85	+0	2.4	0.76	+0	0.9	0.49	+0	1.0	0.44	+0
not ht	2.3	1.24	+15	1.5	0.59	+3	2.3	0.88	+7	2.5	0.97	+0	1.0	0.54	+6	1.0	0.57	-1
25	2.1	0.85	+4	1.4	0.40	-3	2.5	1.00	+17	2.7	0.89	+9	1.2	0.82	+26	1.1	0.62	+17
50	2.6	1.47	+29	1.9	0.92	+30	2.8	1.17	+33	3.0	1.06	+22	1.2	0.85	+33	1.4	0.84	+43
75	3.1	1.99	+55	2.3	1.04	+55	3.4	1.49	+57	3.4	1.24	+39	1.4	0.91	+48	1.8	1.19	+88
25+w	2.0	1.16	-1	1.4	0.58	-5	2.4	1.20	+12	2.5	0.82	+0	1.0	0.68	+9	1.0	0.53	+5
50+w	2.2	1.51	+9	1.4	0.43	-6	2.7	1.19	+24	2.7	1.22	+10	1.1	0.51	+15	1.2	0.64	+24
75+w	2.0	1.04	-1	1.7	0.64	+15	2.8	1.05	+32	2.8	1.12	+15	1.2	0.58	+31	1.2	0.66	+26

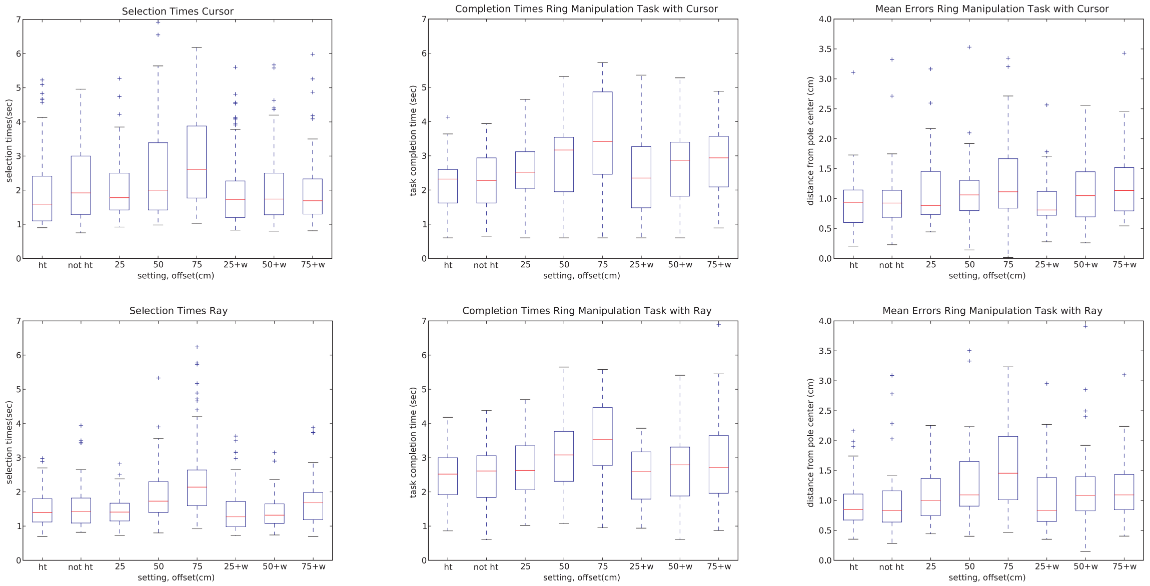

Figure 8: *Experiment results. Top row: cursor (a) selection time (SC), (b) manipulation time (MC), (c) manipulation mean error (MCE). Bottom row: ray-based (d) selection time (SR), (e) manipulation time (MR), (f) manipulation mean error (MRE). Each sub-figure has eight box plots for viewing conditions, ht: headtracked, 25, 50, 75 cm offset from correct head position, and 25+w, 50+w, 75+w cm offset with viewpoint compensation.*

mally. This shows that when large distortions occur, tasks can become difficult for other reasons.

From the quantitative results given in Table 1 and Figure 8, we observe the expected, negative effects of increasing distortion on both mean task completion times and errors when no viewpoint compensation is applied. Also, the standard deviation increases, indicating an increase in task difficulty. When compared to the performance of the classic, headtracked solution, only an offset of 25cm (looking over the shoulder) appears to be a comparable, workable solution. When standing shoulder-to-shoulder (50cm) or at comfortable distance (75cm), the mean selection time and mean error quickly raise up to and over 50%.

When viewpoint compensation is used, the negative influence of increasing distortion on task performance is not strongly visible. When compared to the headtracked situation, there is a slight increase of standard deviation. We must note that the manipulation task itself was more difficult under increased viewpoint offset conditions. The distortion would cause the images to be projected at screen edges or almost off-screen. Also the trajectories became longer and slanted as the distortions increase, making the task more difficult. In some tasks, such as ray-based selection, task performance under distorted view conditions is comparable to the normal headtracked situation. These results are in agreement with the results presented in [Sim07].

An interesting observation is the positive effect of viewpoint compensation on direct interaction, which is essential on close-range VR systems. Under distorted viewing conditions, co-location between the physical space and the sheared-perceived scene is lost. Without correct co-location, direct interaction tasks would be almost impossible to complete without any visual feedback. We clearly see this from the large negative impact of view distortion on the cursor-based selection measurements. In this situation, the user solely relies on visual feedback of the cursor. When viewpoint compensation is applied, co-location is maintained. The experimental results indicate that, even though the world is sheared, the selection performance is brought back to the level of the normal headtracked situation.

6. Conclusions and Discussion

In this paper we extend earlier work on multi-user interaction on projection based VR. We analyzed the problems of multi-user viewing and interaction on the Virtual Workbench, a close-range VR system. We first discussed possible solutions and describe the camera models and compensation techniques used to ensure visual and interaction consistency for the interacting users. We implemented these techniques in our VR development framework and performed both a quantitative and qualitative user study to evaluate them. We evaluated effects of various amounts of view distortion on task performance and accuracy, and the impact of our view compensation approach. For this, we focused on basic, close-range object selection and manipulation tasks on the Virtual Workbench.

Motion parallax is essential for a usable work experience on the Virtual Workbench. Our dynamic average viewpoint is an alternative camera model that introduces motion parallax and retains stereo for multiple users. The stereo parallax is in this case restricted to the horizontal axis to avoid overall fusion problems. Evaluations show that such an alternative camera model can improve the overall collaborative user experience. At the same time, by using the viewpoint compensation approach, usability of interaction tools can be maintained, also in direct, close-range situations. It removes confusion and discomfort during interaction in distorted spaces by making interaction consistent with the user's perception of the scene. In many cases, interaction with view compensation can be as effective as in the single-user case.

We experienced that complex rotation and docking tasks in distorted spaces are more difficult to make consistent, since rotations can change the shape of the objects. We describe our recent approach to solve this issue in [KMdHP07]. Image quality and usability can also be further improved by taking into account individual eye-separation and lighting effects [WHR99]. Furthermore, we plan to design a set of generic solutions and guidelines to extend existing interaction techniques. Based on our current experiments, we expect our method to enhance the effectiveness of collaborative work on the Virtual Workbench. We feel our approach is an attractive solution for many applications because it is easy to implement and accessible. We are currently extending our VR-interaction framework to provide a flexible configuration of viewing and interaction scenarios. Finally, we want to evaluate the new possibilities in real-world VR applications for collaborative data exploration and visualization.

References

[ABM*97] AGRAWALA M., BEERS A. C., MCDOWALL I., FROEHLICH B., BOLAS M., HANRAHAN P.: The two-user responsive workbench: support for collaboration through individual views of a shared space. In *Proc. ACM SIGGRAPH* (1997), pp. 327–332.

[APT*98] ARTHUR K., PRESTON T., TAYLOR R., BROOKS F., WHITTON M., WRIGHT W.: *Designing and Building the PIT: a Head-Tracked Stereo Workspace for Two Users*. Tech. rep., Chapel Hill, NC, USA, 1998.

[BFSE01] BIMBER O., FRÖHLICH B., SCHMALSTIEG D., ENCARNACAO L. M.: The virtual showcase. *IEEE Computer Graphics and Applications 21*, 6 (2001), 48–55.

[BLCN02] BLOM K., LINDAHL G., CRUZ-NEIRA C.: Multiple active viewers in projection-based immersive environments. In *Proceedings, Seventh Annual Symposium on Immersive Projection Technology (IPT 2002)* (March 24 - 25, 2002).

[BMC04] BOLAS M., MCDOWALL I., CORR D.: New research and explorations into multiuser immersive display systems. *IEEE Comput. Graph. Appl. 24*, 1 (2004), 18–21.

[FHH*05] FROEHLICH B., HOCHSTRATE J., HOFFMANN J., KLUEGER K., BLACH R., BUES M., STEFANI O.: Implementing multi-viewer stereo displays. In *Proc. WSCG '05* (2005), pp. 139–146.

[HSAW05] HELDAL I., STEED A., AXELSSON A.-S., WIDESTRÖM J.: Immersiveness and symmetry in copresent scenarios. In *Proc. IEEE VR '05* (2005), IEEE Computer Society, pp. 171–178.

[HSS05] HELDAL I., STEED A., SCHROEDER R.: Evaluating collaboration in distributed virtual environments for a puzzle-solving task. In *Proc. HCI International* (2005).

[KMdHP07] KOUTEK M., MOLENAAR R., DE HAAN G., POST F. H.: Visual consistency in rotational manipulation tasks in sheared-perceived virtual environments. In *Proc. of Immersive Projection Technology and Eurographics Virtual Environments '07* (july 2007).

[Mol07] MOLENAAR R.: Viewing and Interaction for Multiple Users in Projection-Based VR Systems. In *MSC. thesis, Delft University of Technology, Data Visualization Group* (april 2007).

[Sim07] SIMON A.: Usability of multiviewpoint images for spatial interaction in projection-based display systems. *IEEE Trans. Vis. Comput. Graph. 13*, 1 (2007), 26–33.

[SS05] SIMON A., SCHOLZ S.: Multi-viewpoint images for multi-user interaction. In *Proc. IEEEVR '05* (2005), pp. 107–113.

[WHR99] WARTELL Z. J., HODGES L. F., RIBARSKY W.: The analytic distortion induced by false-eye separation in head-tracked stereoscopic displays, 1999.