Photoportals: Shared References in Space and Time

André Kunert Alexander Kulik Stephan Beck Bernd Froehlich

Bauhaus-Universität Weimar
Bauhausstrasse 11, 99423 Weimar, Germany
<first name>.<last name>@uni-weimar.de

Figure 1. Collaborative interaction with Photoportals. Left: A user takes a virtual photo of a local and two remote collaborators in a virtual city model. Note that the local user (standing on the left) also appears in the virtual photo. Right: Two local users (on the left side) share a box-shaped Photoportal with two remote collaborators (on the right side). A remote user operates a virtual pointer inside the portal to manipulate the ship model in the background at a convenient scale and distance while the other users provide guidance from their individual viewpoints.

ABSTRACT
Photoportals build on digital photography as a unifying metaphor for reference-based interaction in 3D virtual environments. Virtual photos and videos serve as three-dimensional references to objects, places, moments in time and activities of users. Our Photoportals also provide access to intermediate or alternative versions of a scenario and allow the review of recorded task sequences that include life-size representations of the captured users.

We propose to exploit such references to structure collaborative activities of collocated and remote users. Photoportals offer additional access points for multiple users and encourage mutual support through the preparation and provision of references for manipulation and navigation tasks. They support the pattern of territoriality with configurable space representations that can be used for private interaction, as well as be shared and exchanged with others.

Author Keywords
Interactive Systems; Multi-User Interaction; Collaborative Virtual Environments; 3D Interaction; 3D User Interfaces

ACM 978-1-4503-2540-0/14/02...$15.00.
http://dx.doi.org/10.1145/2531602.2531727

ACM Classification Keywords
H.5.2 Information interfaces and presentation: User Interfaces

INTRODUCTION
Virtual reality allows the exploration of simulated places and objects as if they exist in real space and real time. Many 3D interaction techniques are correspondingly based on the experiences and skills we have acquired in the real world – with the restriction that the displayed objects cannot be touched and felt with bare hands. Instead, mediating tools and references are used to manipulate the virtual matter.

We suggest photography as a unifying metaphor for reference-based interaction techniques. We have learned to use photos as visual representations of places, objects and moments in time and to exchange these with others. In this sense, photography allows the appropriation of inaccessible objects and locations for review and presentations. Photos also capture viewpoints and serve as mnemonics for relations and past experiences. Scaling is an implicit parameter of the recording process. Scenes captured with contemporary digital cameras can be instantly reviewed on the built-in screen.

Our virtual photos capture scenes and actions not just as images, but in 3D and, therefore, they become portals to remote locations or alternate realities which can be entered or manipulated at any time. Our virtual camera also supports the recording and playback of 3D videos that involves animated objects, as well as actions and body representations of local and remote users (Figure 1 left). We show how this metaphor of virtual photography supports a multitude of reference-based 3D interaction techniques while maintaining the direct manipulation paradigm on the basis of a congruent input-output space.

The work presented in this paper is a first approach to provide groups of collocated and remote users acting in a truly shared 3D space with advanced collaborative tools and techniques. The main contributions of our work are:

- A novel interface concept that builds on the metaphor of digital photography for the immediate creation and the management of sharable references in space and time.
- A coherent set of interaction techniques that exploit scene references for efficient collaboration in shared virtual environments.
- A user study demonstrating the usability of Photoportals.

We designed a powerful mixed-reality system to explore the future of local and remote 3D collaboration in immersive virtual environments. We argue that reference-based interaction techniques are an important tool in such systems since they support territoriality, exchange of information and complementary activities of users. Our Photoportals demonstrate the applicability of this interface concept to address the challenges of interaction in 3D collaborative systems.

DESIGN CHALLENGES OF COLLABORATIVE SYSTEMS

Mutual awareness of participating users and their actions is a primary requirement of collaborative user interfaces. Gutwin and Greenberg defined workspace awareness as the "up-to-the moment understanding of another person's interaction with the shared workspace" and emphasized its three main elements: 1. the presence, identity and authorship of participants (who), 2. the involved artifacts, actions and intentions (what) and 3. the location, gaze, view and reach of users (where) [18]. The framework also considers the history of artifacts and events (how and when). They noted that the ease of people maintaining workspace awareness in real-world collaborative settings is based on the continuous gathering of this information through consequential communication (the observation of each other's activity), feedthrough (the sensory perception of involved artifacts) and intentional communication (verbal and gesturing).

In collocated settings, real-world workspace awareness can directly be exploited for computer-supported collaborative work. Multi-touch tabletops, for example, provide such a shared space for joint interaction with 2D data. More elaborated technologies are required for multi-user 3D applications, as the appearance of 3D geometries depends on the users' individual viewing perspectives. In 1997 Agrawala et al. presented the two-user responsive workbench [1], the first 3D display system that supported more than one user. Since then, various alternative approaches and improvements have been proposed (e.g. [3, 37, 23, 16, 26]).

In the realm of distributed collaborative systems, workspace awareness is often limited by low bandwidth and several explicit notification options have been proposed [11, 10]. Novel sensor technology and increasing network bandwidth, however, promote the introduction of high-fidelity telepresence systems that support whole-body 3D capturing and reconstruction of remote participants in real time [15, 28, 5]. The quality of remote user representations is not yet perfect and factors such as noise, missing data and latency clearly affect usability but, like in collocated settings, the systems already offer similar levels of workspace awareness. They create a seamlessly shared 3D interaction space for users at different physical locations.

The high level of workspace awareness provided in collocated and tightly coupled distributed settings supports the coordinated interaction of multiple users. However, conflicts still occur, specifically if users try to manipulate the same virtual objects simultaneously [30, 35, 22, 33]. Grossman and Balakrishnan [16] suggested the locking of manipulated objects while highlighting which user is operating on it. Riege et al. instead proposed merging the input of multiple users and visualizing the simultaneously-induced forces [35]. In both cases, only a mapping for the conflicting input is suggested. The actual conflict is not prevented.

The design space of combining the input from multiple users received particular interest. Distributing degrees of freedom among participants, for example, can enforce the users' cooperation [8, 34]. However, Pinho et al. also found that this approach negatively affected the comprehensibility of the interface [34]. Ruddle showed that merging the input from different users is likely to increase the cooperation overhead [36].

Instead of enforcing joint efforts, Benford et al. suggested encouraging collaboration with complementary tools that can be combined for additional functionalities [6]. We followed this example of complementary interfaces. Photoportals can be used to prepare location references for the manipulation input of other participants or as targets for group navigation.

Scott et al. observed the pattern of territoriality in collocated collaboration. It refers to the common behavior of individuals in a group to establish three distinct interaction areas: personal, group and storage territories [38]. Supporting this user behavior can reduce input interferences [43]. Several implementations of the concept have later been reported in the context of collaborative tabletop applications. They generally divide the screen space into individual areas (e.g. [24]). Grossman and Balakrishnan applied this approach to multi-user interaction with a volumetric display [16]. Alternatively, additional displays have been proposed for private views of details or independent content manipulation [42, 29, 39].

We adopted the pattern of territoriality. Our Photoportals support the ad-hoc creation of configurable spacelets that can be used as private or storage spaces. Users can instantly create secondary scene representation and use them for the individual examination of objects and locations from various perspectives. Manipulation conflicts are effectively prevented, as this does not affect the shared scene. If users actually want to change the configuration of the shared scene, they can apply their manipulations in individual Photoportal versions, compare their proposals and enter the one they agree on.

Avoid Input Interference	Exploit Simultaneous Input	Avoid Fragmented Visibility	Exploit Multiple Views	Avoid Unexpected Motion	Exploit Collaborative Planning
Known Techniques					
• Object locking [16] • Explicit notifications [10, 16]	• Input merging [8, 35, 36, 34] • Territoriality [16, 38, 43, 24, 42, 29, 39] • Complementary input [6]	• Extended field of view [20] • Show Through [2]	• Specialized views [1]	• Shared control [26] • Augmented group navigation [26]	• Shared WIM [5]
Novel Strategies for Collaboration Support as Provided by Reference-Based Interaction with Photoportals					
Secondary scene representations reduce demand for object manipulation.	Mutual interaction support with location references.	Perspective sharing reveals invisible fragments of the shared environment.	Individual manipulation of viewing parameters in privately used Photoportals.	Group transit through portals reduces demand for steering-based navigation.	Rapid creation and management of shared WIMs and previews of target locations.

Table 1. Several major challenges of collaborative 3D environments (in columns) have been addressed in prior work. Our Photoportals add a range of alternative strategies that exploit the concept of shareable references for collaborative interaction.

The design of collaborative 3D user interfaces involves further challenges that relate to 3D viewing perspectives and group navigation. A surrounding 3D environment cannot be overseen completely and the appearance of 3D objects depends on the viewing perspective. As a result, each user only sees fragments of a shared 3D scene at a time. Hindmarsh et al. addressed this fragmented visibility problem for the case of individually navigating users in a distributed virtual reality system. They suggested extended peripheral vision and exaggerated representations of other users' actions to improve the perception of remote participants [20]. Another type of fragmented visibility is interpersonal occlusion. Geometric features may be hidden behind other geometry from several perspectives. Argelaguet et al. suggested tackling this problem with cut-away visualizations that let indicated features show through occluding parts of the shared scene [2]. Note that both techniques imply a distorted perception of the scene which can be highly detrimental, e.g. if the evaluation of geometric properties is purposed.

We suggest an alternative approach to resolve fragmented visibility that preserves the appearance of the scene. Objects and locations that cannot directly be seen by everybody can be captured in a Photoportal and shown to everybody else. Geometry clipping at the surface of the virtual photos implicitly enables cut-away views towards occluded geometry.

Exploring a 3D virtual world requires navigation during which the visual and the vestibular perceptual systems perceive contrary motion signals: visual motion flow without physical acceleration. This perceptual conflict may induce cybersickness, even more so if the virtual motion is not expected. Input for group navigation should thus be observable by all involved users and they should be allowed to interfere easily if they do not agree. Kulik et al. suggested a large stationary device to increase mutual awareness of group-related input. They also developed a set of augmented group navigation techniques that avoid collisions of individual users with the virtual environment – at the price of temporarily giving up the notion of a shared 3D space [26].

By using Photoportals, the demand for group locomotion can be drastically reduced. Individual users can prepare a location reference to a new travel target and then invite the whole group to enter this place directly, using the photo as a portal.

Our design process was inspired by the tangible interaction framework of Hornecker and Buur [21]. Following their guidelines, our interaction techniques build on direct haptic manipulation with isomorphic mappings and enable lightweight interaction in experimental steps (Tangible Manipulation). As demanded by their concept of Spatial Manipulation, we ensure a consistent space among users and interactive objects in which the movement of objects and the users' bodies has a comprehensible meaning. While the effects of individual perspectives in 3D environments imply the problem of fragmented visibility, our Photoportals provide simple ways to solve it. The operation of our interface elements generally involves the whole body in action. With a set of complementary interfaces that include one or many Photoportals, we provide multiple access points to engage users in the collaboration (Embodied Facilitation). In the sense of preparing and exchanging scene references for further interaction, Photoportals also support the externalization of ideas and provide unambiguous references for the communication and collaboration (Expressive Representation).

Reference-based interaction techniques have rarely been applied to solve design challenges of collaborative systems. Beck et al. suggested a shared World in Miniature for collaborative tour planning, but did not consider further usage scenarios [5]. Spindler et al. suggested Tangible Windows as private views towards a shared 3D scene [39]. Their design builds on the concept of handheld magic lenses [7] and extends it with techniques for object manipulation and navigation. They also mentioned the advantage of private views for parallel interaction of multiple users, but no particular workflow of collaborative interaction was outlined in their paper. We show how the concept of secondary scene representations can minimize common collaboration obstacles including fragmented visibility and the coordination of manipulation and navigation.

REFERENCE-BASED 3D USER INTERFACES

Photoportals are visual references to captured objects and locations in a shared 3D environment. While this concept is new in the realm of collaboration support, similar reference-based interaction techniques are successfully used for manipulating virtual objects that are out of reach: A reference to a distant part of the scene held in one hand provides a reference for input from the other hand.

Such bi-manual interaction also leverages the benefits of proprioceptive cues [17]. Voodoo Dolls, for example, provide the user with an additional miniature representation of selected objects [32]; a World in Miniature (WIM) can represent the whole virtual environment on a smaller scale [40]. WIMs are often used to visually support wayfinding. Several extensions have been proposed to the technique, including scaling and scrolling [44] and a navigation technique that flies the user into selected locations of the handheld miniature [31].

Hinckley et al. suggested physically graspable props as reference objects [19]. If the prop's physical shape fits into the human hand, users can take advantage of their manual dexterity. The necessary dislocation of visual space and motor space, however, may be detrimental. We aimed for a novel interaction technique that combines the advantages of physical grasping and a congruent input-output space.

Stoev and Schmalstieg suggested building on the paradigm of magic lenses for reference-based interaction [41]. They proposed virtual lenses that show parts of the scene from a different viewpoint and defined a set of corresponding "through-the-lens" interaction techniques. Several of them, including pick ray manipulation at the remote location and virtual travel "through the lens," were similar to corresponding techniques of our Photoportals. Our work builds on this example and extends it with additional viewing modes and a refined user interface for collaborative settings.

The concept of interconnecting different virtual locations is also known as "locales" that can be assembled to virtual environments of potentially unlimited size [4]. Greenhalgh et al. extended this concept with temporal links that capture sequences of user activity [13, 12]. They explored various application scenarios for this functionality, including the intuitive authoring of dynamic 3D content and the post-hoc evaluation of user interaction. They also proposed several manifestations of temporal links mostly corresponding to the requirements of digital storytelling. Most comparable to our sequence recordings with Photoportals is Holovid, a WIM-like scene representation that includes captured motion sequences. Other than temporal links in the distributed system MASSIVE-3 [14], our Photoportals support the recording of temporal sequences that include life-sized 3D video avatars of local and remote participants. Our tangible camera interface furthermore enables inexperienced users to operate the recording and manage several of them in a virtual gallery.

Technically, our Photoportals are most closely related to the pioneering work of Stoev and Schmalstieg [41], but our implementation in the context of collaborative virtual reality goes beyond their original proposal. We created a novel tangible interface for the immediate creation and management of scene references that corresponds to the familiar functionality of digital cameras. The resulting ease of use enables the application of Photoportals in collaborative settings with naive users. Furthermore, our Photoportals incorporate other successful concepts from prior research into a unifying interface. This includes all of the above mentioned extensions to WIM-like scene representations.

The immediate creation of secondary scene representations can be compared to that of Voodoo Dolls, but instead of a grasping gesture in screen space, our interface exploits the metaphor of digital photography. This novel approach directly supports adaptations of the capturing process to create specialized representations for different use cases, as well as the management of multiple scene representations in a virtual gallery. Our tangible camera interface facilitates the handling and the exchange of Photoportals.

MULTI-USER INTERACTION IN VIRTUAL ENVIRONMENTS

We developed and tested our interaction techniques with a state-of-the-art immersive telecollaboration system that connects two multi-user 3D power walls. The participants in front of each display are captured using color-and-depth cameras. The resulting color- and depth-image streams are transferred to the other location where real-time 3D video representations of remote users are shown [5]. For a group of collocated users, each user can independently walk in front of the screen, but navigation through the environment is always performed for the entire group [26]. Otherwise, the consistency of the shared virtual environment cannot be preserved. We coupled virtual locomotion for the whole group and limited its control to input from a single group navigation device (Spheron) that is purposefully large in order to increase awareness and its accessibility for everybody in the group (Figure 1 right). The six degree-of-freedom device consists of a large 3D trackball and a separate joystick-type handle for 3D translational input. Its operation requires comparably large movements that are visible to all group members.

Both groups use a multi-user 3D power wall display with a resolution of 1920 by 1200 pixels. The larger display measures 4.3 by 2.7 meters screen size and provides stereoscopic image pairs for up to six users [26]. The smaller one supports only two users with a screen size of 3 by 2 meters.

The setup generally provides high workspace awareness – at least as long as the remote group is visible within the boundaries of the screen and the participants' actions correspond to the real-world, as in direct walking and bare-handed gesturing. As soon as augmented, non-isomorphic interaction techniques are involved, it becomes much more difficult to maintain awareness of each other's actions. However, interaction with virtual objects requires a certain level of abstraction, as the data cannot be grasped directly. Therefore, we focused our efforts on the design of interfaces and interaction techniques that support the mutual understanding of "magic" techniques like teleporting, remote manipulation or scaling.

The pick ray is a common solution for interaction with virtual objects. We also provide this basic technique, but not to everybody simultaneously. We experienced conflicts of users that try to operate on the same object. In a first attempt to increase the group's coordination, we limited direct manipulation to the selection of a virtual reference point on a geometry. The whole scene could then be rotated around this point using the 3D trackball of the Spheron device as a physical reference. The manipulation of individual objects required indirect input that was operated with motion input relative to the 3D trackball. The technique prevented input conflicts, but only at the price of inhibiting simultaneous input. Moreover, the dislocation of physical input and visual output was confusing for many users.

One of the main motivations for object manipulation is the desire for their examination from various perspectives. An unintentional side effect often is the change of the object placement relative to the surrounding scene. The demand for object manipulation can thus be reduced if users can easily view objects from every angle without actually moving them. The scene rotation around a selected reference point, as described above, provides such functionality, but only for the whole group. If someone wants to explore a particular detail, not everybody else will share this interest. Collaborating users often prefer to divide their focus temporarily in order to explore various aspects of a shared context and then exchange their findings. Consequently, we aimed to create an interaction technique that allows individual examination of different scene details and promotes immediate sharing of gathered perspectives without disturbing anybody or compromising the coherence of the shared virtual environment.

Secondary scene representations such as a WIM or Voodoo Dolls allow users to virtually get hold of some part of the virtual environment and observe it from all sides while the original scene remains unchanged. Providing users with such references can thus be very useful – if their creation and handling can be done without hassle. Successful collaboration support with a shared WIM has been shown earlier [5]; the rapid creation of a WIM with dedicated parameters, has not been considered so far. Voodoo Dolls can be created rapidly with a grasping gesture in screen space but, as a result of this gesture-based creation, the doll is attached to the user's pinched fingers and vanishes if the fingers are released. This complicates their handling; in particular, rotation beyond the range of the wrist and sharing between multiple people.

We propose the metaphor of digital photography for the immediate creation of secondary scene representations to which we refer as virtual photos or Photoportals. In correspondence to actual photography, snapshots can be made instantly and the capturing device provides a physical handle to the resulting virtual photo. Moreover, as known from photographic cameras, various controls enable the parameterization of the capturing process for different use cases and facilitate the management of multiple virtual photos in a gallery.

PHOTOPORTALS

Photoportals combine the capabilities of reference-based interaction techniques in a consistent user interface that builds on the metaphor of photography. It consists of a tracked tangible camera prop with a set of physical buttons and a virtual display through which the captured scene can be perceived. Users can take virtual photos or record a video of anything they find interesting in the scene. These virtual photos or videos are not just images like their real-world counterparts, but they can be augmented with powerful features and functionalities. We exploit the following four elements of photos for the interaction with virtual environments: 1. a frame that contains the captured scene, 2. a focused object or location of interest, 3. the camera viewpoint, and 4. recording time.

Once a view has been recorded, the respective scene features can be individually examined without necessarily involving the whole group. If interesting aspects shall be discussed, they can easily be shown to others. A collection of successively taken photos shows parts of the virtual environment that are currently not visible. During an architectural walkthrough, for example, users can take photos of details and perspectives they want to discuss with others – now or later. This approach corresponds to groups exploring real environments and it helps with structuring the collaborative review process. It is even more attractive in virtual environments. Our virtual photos are essentially portals and thus provide shortcuts to the captured locations. Consequently, we dubbed our mixed-reality interface Photoportals.

The most fundamental implementation of Photoportals is a monoscopic projection of the scene on the virtual screen of our device. This 2D viewing mode is most effective if a particular vista towards the scene shall be shown to others. It corresponds exactly to real-world photography with the live-view function of digital cameras. Alternatively, our 3D viewing mode shows the captured scene in stereo and continuously adapts the view according to the users' head positions in relation to the camera's virtual screen. In this case, the Photoportal can be considered a window to a remote place. This mode is most appropriate for accessing remote locations.

Furthermore, the visual appearance of our Photoportals can be switched from a single window to a box shape, which is a composite of six individual portal windows that are used in conjunction to show the same scene location from six orthogonal directions. The portal box is well-suited for the examination and manipulation of remote objects. In combination with the 3D viewing mode, it resembles a fish tank containing objects of interest. Protruding parts of the objects are automatically clipped. In combination with the 2D viewing mode, the display box shows six aligned orthographic projections of the captured part of the 3D scene. This mode is particularly useful for the accurate examination of geometric relations, as well as for constrained object manipulation.

We also experimented with other combinations of viewing and display parameters, but we found that the combination of two viewing modes (2D and 3D) with two display setups (Window and Box) provides the most sensible manifestation of Photoportals. The four resulting Photoportal modes are:

- 2D-Window (Figure 2 green Photoportal to the right)
- 3D-Window (Figure 2 blue Photoportal to the left)
- 2D-Box (Figure 3 green Photoportal to the right)
- 3D-Box (Figure 3 blue Photoportal to the left)

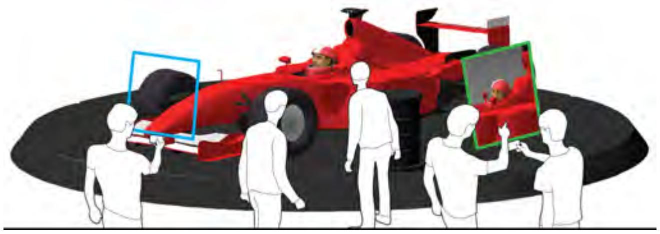

Figure 2. Capturing scene references in both window display setups. The blue Photoportal to the left illustrates the 3D-Window mode. The live preview in the blue frame integrates seamlessly with the scene behind. The green Photoportal to the right illustrates the 2D-Window mode. The captured 2D image corresponds to a real photo. It appears identical from all perspectives.

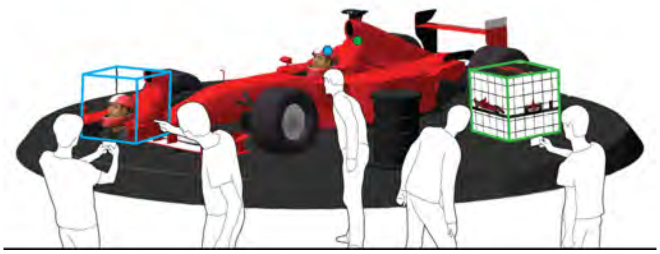

Figure 3. Captured scene references in both box display setups. The blue Photoportal to the left illustrates the 3D-Box mode. The captured scene is shown within the blue box as a miniature. The green Photoportal to the right illustrates the 2D-Box mode. The captured car model appears aligned within the green box with orthographic projections shown on all of its sides. The blue and green intersection points on the surface of the car model in the background define the selected geometry.

REFERENCE-BASED INTERACTION WITH PHOTOPORTALS

Photoportals offer a unifying interface for reference-based interaction techniques, some of which are novel, others building on prior work. The following subsections, organized by interaction goals, explain these techniques in detail. At the end of each subsection, we summarize the implemented functionalities with references to related work where applicable.

Taking Photos and Videos

Taking a virtual photo or video is an almost equivalent operation to its real-world counterpart. The user orients the camera towards the target object or location and then triggers the capturing with a button on the device (see Figures 2 and 4). We use a two-state button as known from autofocus cameras. In our case, the intermediate state enables a live preview of the Photoportal which helps to optimize the captured perspective. Fully depressing the trigger button captures the current perspective as a photo or video.

The Photoportals in our system do not only capture virtual content. The photos and videos also record the users as 3D video avatar representations (Figure 1 left). This feature can be used to capture user postures and movement sequences, e.g. for studies of ergonomics. During virtual videotaping, the current color- and depth-image streams of the captured local and remote users are stored for later playback.

In both box modes (2D-Box, 3D-Box), the capturing process is optimized for the creation of an object representation that may later serve as a reference for remote manipulation (Figure 3). If the user encloses 3D objects with the portal box, these objects become captured as they are, but with all protruding parts cut away. If, instead, the box is empty during the capturing process, the camera tries to fetch distant objects along its capturing direction. We search for candidate objects by means of an intersection ray that extends from the Photoportal camera into the scene along the capturing direction. The first intersected object will be represented in the box. The original object placement in the scene remains unaffected, but equivalent to the creation of a Voodoo Doll, [32] the view is parameterized such that the geometry will appear as a miniature inside the handheld Photoportal Box (with its size remaining equal in 2D image space). In 2D-Box mode, the algorithm sets the viewing parameters such that the bounding box of the selected object fits inside the Photoportal box and is aligned to the coordinate system of the parent node in the scene graph. The live preview in the box modes shows only the intersection point of the capturing ray with the scene.

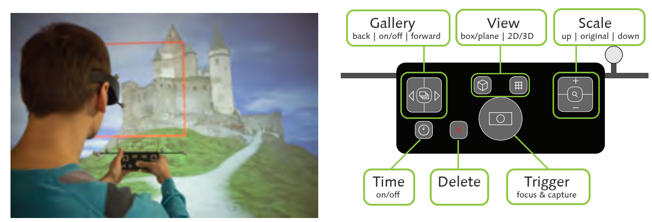

Figure 4. Left: The Photoportal in 3D-Window mode. The live-preview integrates seamlessly with the surrounding scene. Right: The assignment of buttons on the tangible camera device.

All virtual photos are automatically saved in a gallery that is associated with each Photoportal camera device. They can be reviewed, adapted and deleted at any time. After the initial capturing, the viewpoint can still be adjusted by moving the camera device while pressing the trigger button. The viewing and display modes of a virtual photo can be modified at any moment, as well. In analogy to the zoom control of real-world cameras, our Photoportal input device offers a two-way button scaling of the displayed scene representation.

Photoportals do not capture static images, but a certain viewpoint towards a 3D scene. If changes to that same scene are applied after a photo has been taken, they affect the situation that is visible on the photo. Some users expected that a scene that had been captured on a photo could not be changed anymore. Correspondingly, we extended the capturing functionality of Photoportals to record versions of the scene if the recording of temporal aspects is enabled via the *time* button on the camera device (see Figure 4 right). This allows users to go back to earlier versions of the scene if they are unsatisfied with their manipulations. In collaborative work, it also enables users to apply manipulations to a separate version and show their results to everybody else. Then the group can accept the suggested changes by entering this version of the scene. We currently record the full scene graph for a version and only support a full exchange of scenes, but we aim to include a versioning system that supports branching and merging of content.

Recording temporal aspects naturally supports the capturing of action sequences similar to the creation of temporal links in the work of Greenhalgh et al. [13]. Our virtual video sequences also include the captured performances of the life-sized 3D video avatars of local and remote users.

In summary, our capturing techniques offer the following features:

- Immediate creation of scene references
- Two display setups:
  - *Photoportal Window* for location references
  - *Photoportal Box* for object references and WIM
- Two viewing modes:
  - *2D View* for aligned perspectives
  - *3D View* for portal functionalities
- Recording of versions and animations [13]
- Capturing of local and remote users as 3D video avatars
- Storage and management of scene references in a gallery

Sharing Perspectives

Our 3D display system enables the perception of a shared 3D environment among multiple users. However, sharing the same 3D environment does not guarantee that everybody sees the same (fragmented visibility). Our Photoportals provide three techniques to increase the visibility of relevant scene details for all involved users: perspectives can be shared, occluding objects can be cut away and additional object representations can be acquired within arm's reach.

If users want to share a particular perspective of the scene, e.g., to show the alignment of several objects, they can take a photo of this view and immediately show it to others on the virtual display of the camera device (Figure 5). The 2D-Window mode is often preferable in this case. It ensures that all users can see exactly the same perspective on the Photoportal display. However, the lack of stereo and motion parallax also hampers depth perception. In cluttered environments, such as a car's engine compartment, it can become impossible to distinguish between objects that are located at different depth levels along the perspective. In these occasions, the 3D-Window mode can be the better choice.

Each group of users typically has a single camera device, but multiple devices per group are also supported. Each device maintains its own gallery but their content can be shared too. Taking a photo of the virtual screen of another camera device copies the shown photo (Figure 5 right).

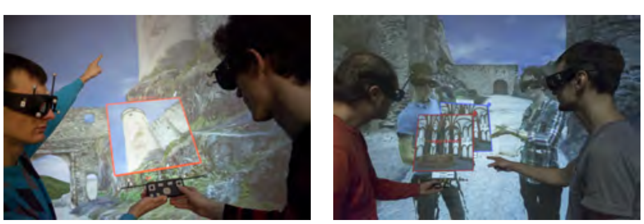

Figure 5. Left: A user shares an interesting perspective with a local collaborator. Right: The exchange of captured photos is supported with a copy functionality between Photoportal cameras.

If the groups' attention shall be drawn to a particular object that cannot easily be seen behind surrounding geometry, they can make use of the Photoportal as a cutting plane. The near clipping plane of our Photoportals is aligned to the virtual display; hence the tool facilitates the exposure of hidden objects through clipping of occluding geometry. Such cutaway views can be captured, readjusted and scaled as any other virtual photo in our system. This facilitates the group discussion of zones that cannot be reached comfortably (Figure 6). Using the display box for cutaway views provides additional information on adjacent geometry. The box's shape can even be adjusted to the object of interest.

Figure 6. Two users examining interior parts of a building model with a Photoportal used as a cutting plane (left) or a 3D cutout (right).

In summary, our interface supports communication and mutual exchange with the following features:

- Rapid capturing of individual perspectives
- Shareable handheld view representations
- Cut-away views to reveal hidden geometry

Navigation Techniques

Photoportals provide references to remote locations and it seems natural to use these for navigation too. Earlier frameworks for reference-based interaction suggested navigation support in terms of camera-manipulations in a WIM [40, 41], but the resulting viewpoint motion can be uncomfortable from an egocentric perspective. Our Photoportals provide the means for the ad-hoc creation of a WIM. We simply need to scale down the scene representation in the 3D-Box mode (Figure 7). Miniaturized 3D avatar representations of local and remote users are also visible inside the WIM.

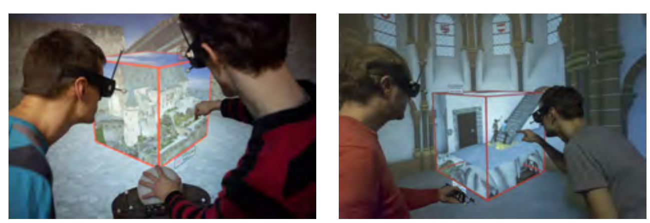

Figure 7. Using Photoportals as a WIM. The image on the right side illustrates the benefits of implicit geometry clipping for wayfinding inside a virtual building. Note the 3D video avatars of remote users in the miniature view.

For viewpoint navigation, however, we do not manipulate the position of these avatars in the WIM. Instead we switch to 3D-Window mode, adjust the scaling and enter the visible locations through the Photoportal. Each user can individually enter a remote scene directly by putting his head into a handheld Photoportal in 3D viewing mode. The user who does this experiences a seamless transition from the main scene that is shared with everybody else to the place displayed by the Photoportal. If the distance between the portal window and the user's head is less than 15 cm, the main scene becomes fully replaced by the portal scene - but only for that user. The others remain completely unaffected by this individual transition. They only see the respective user putting the head into the portal (Figure 8 left).

Group transit to another location requires more steps to ensure everybody's awareness and agreement. Our group interaction policy requires that all group-related input must be operated via a highly-visible central control station, which is generally the aforementioned Spheron in our setup. Placing the Photoportal device on the board around the 3D trackball triggers a coupling between the two input devices. In 2D- or 3D-Window mode, this coupling activates a public gallery with the device's collection of virtual photos on our large physical screen (Figure 8 right). The near clipping planes of all user views become aligned with the gallery in order to avoid occluding scene geometry in front of it. With the buttons of the camera device, the users can browse through the Photoportal collection and select a location they want to enter. Pressing the capturing button grows the selected Photoportal to full screen coverage without affecting the scale of the displayed scene. This entering process deactivates the public gallery and resets the near clipping planes of all user views to the default distance of 15 cm in front of the eyes.

The scene shown in a Photoportal can be scaled up or down before entering. Therefore, navigation in terms of entering Photoportals implicitly supports multi-scale navigation similar to that suggested by [25].

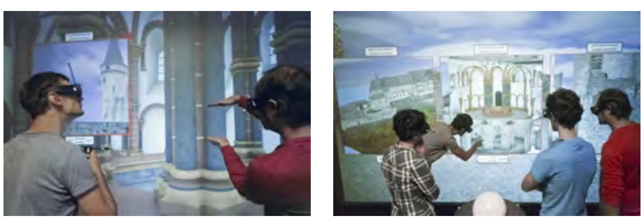

Figure 8. Two ways of entering a Photoportal. Individually, users can put their head into the portal (left). Group transitions require the selection of the target view from a public gallery (right).

In summary, Photoportals support navigation tasks with the following features:

- Shared scene overview for collaborative tour planning [5]
- Preparation of 3D location references
- Multi-scale navigation [25]
- Individual viewpoint transition to remote locations [41]
- Group transit to remote locations

Manipulation Techniques

Similar to other reference-based interaction techniques, our Photoportals enable the remote manipulation of distant objects. The basic principle for all these techniques is the transformation of pick ray input towards the Photoportal display into the shown scene. This has been described previously by Stoev and Schmalstieg [41]. Our implementation of pick ray manipulation in the 3D-Window mode corresponds to their implementation, including the suggested drag-and-drop operation to extract objects from the Photoportal into the main scene. We extended the concept with different manipulation techniques that correspond to the characteristics of our Photoportal modes and added further functionality.

We found that the single-display Photoportal does not support the manipulation of distant objects very well. Without laborious viewpoint adaptations, the distance to the captured object remains the same and it can only be accessed from one direction. The drag-and-drop operation for object extraction, instead, works well (Figure 9 left). We extended the technique with functionalities for object scaling, duplication and animation. When dragging an object from a scaled portal scene into the main scene, the object retains the scaling of the scene from which it is extracted. Inserting objects into a Photoportal works vice versa. A duplicate of an object is created if the user holds the trigger button of the camera device during the object extraction. Extracting a moving object from a video sequence retains its motion path and automatically loops this motion at the new location in the main scene. This feature facilitates the creation of simple animation sequences.

Remote manipulation with the Photoportals in 3D-Box mode corresponds largely to the Voodoo Dolls technique [32]. The corresponding capturing process fits the focused object or location into the box; hence it provides a secondary representation within arm's reach. In our implementation, the remote scene reference appears more like a WIM and is attached to the physical camera device which facilitates its handling among multiple users.

The 3D-Box mode does not support object extraction. Instead, we increased the manipulation range. Once an object inside the Photoportal has been selected with the pick ray, it can be moved beyond the borders of the box. The object does not get lost when passing the border and it is generally still visible in the main scene. The viewport of the Photoportal can be moved thereafter such that the object of interest becomes enclosed by the box again for further manipulation.

The 2D-Box mode additionally offers constrained manipulation for the precise manual alignment of remote objects. Pick ray interaction on one of the six orthogonal projection views is limited to 2D translations along the respective display plane and 1D rotation about its normal (Figure 9 right).

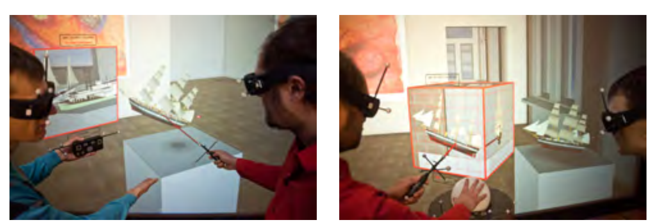

Figure 9. Left: A 3D object is extracted from a Photoportal in 3D-Window mode. Right: The Photoportal in 2D-Box mode supports constrained manipulation. For more relaxed operation, the Photoportal box can be coupled with the Spheron device.

Depending on the task, object manipulations may take time. Continuously holding the Photoportal device to provide a reference for manipulation input can become exhausting. For these situations, we couple the Photoportal interface in box display mode with the Spheron device. The Photoportal box becomes detached from the camera device and is then displayed hovering above the Spheron (Figure 9 right). Rotation input to the trackball is now directly applied to the orientation of the Photoportal box.

In summary, Photoportals support manipulation tasks with the following features:

- Manipulation of remote objects [41, 32, 40]
- Mutual provision of manipulation references
- Drag-and-drop transitions between 3D locations [41]
- Drag-and-drop copy operations with implicit scaling
- Constrained object manipulation

USER EXPERIENCE WORKSHOP

Photoportals offer novel functionalities that can improve collaborative interaction in 3D virtual environments – but only if they can be operated without hassle. Therefore, the general usability of the interface was a primary concern during development. For the evaluation of our novel interface, we invited design students from a partner university to realize their own ideas of interactive virtual environments. In preparation for this workshop, we visited the students at their university and introduced them to the interactive possibilities of our system using photos and videos to illustrate various functionalities. Groups of two to four students were then asked to develop application scenarios.

The six resulting scenarios ranged from 3D design reviews over a futuristic 3D file browser to more gaming-oriented applications. Most fascinating was the possibility of creating experiences that would not be possible in the real world, e.g. traveling through hyperlinks, walking upside-down or meeting oneself in recorded scenarios. After five weeks of preparation, eight of the students came to our lab to realize their ideas. We focused on four scenarios which had to be managed by groups of two students.

Procedure

Before we started to work on the students' scenarios, we familiarized them with the multi-user 3D displays and the 3D capturing and reconstruction technology for telepresence applications. The participants were also thoroughly introduced to the operation of the Photoportals in groups of two. In an existing virtual environment, they were shown each functionality and then asked to operate it themselves. By following a predefined script, we made sure that each user operated every function alone, but also in collaboration with the other. This took between 60 to 90 minutes per group. Then each user was asked to complete a questionnaire on the usability of the overall system and the Photoportal interface. It consisted of a system usability scale (SUS) [9] and more specific items on 7-point Likert scales. In particular, we were interested in ratings of the different Photoportal functionalities and the usefulness of the system to promote collaboration. We worked on the realization of the students' ideas during the following two days. This gave us the opportunity to observe users during the unsupervised use of the interface and to discuss their opinions in brainstorming sessions.

Results

The overall feedback was very positive. The students were excited about the system and also rated the general usability of the Photoportals positively. On a scale between 1 (worst) and 7 (best) a mean score of 5.25 (SD=1,04) was given. The mean score on the system usability scale was 71.65 (SD=10.43), which is above average [27]. One of the students added the comment that using the Photoportals primarily required learning the button mappings of modes and functions, which he considered to be normal for any new electronic device. During their independent work with the system, we observed that it took everybody a while to remember the mapping.

The ease of using the individual functionalities was also rated on the same scale (1-7). Capturing, deleting and exchanging photos, as well as entering the portals individually or as a group, were rated to be most easy with an average score of 6.17 (SD=0.86). The later adjustment of perspectives, including the scaling of the scene, was rated slightly lower (M=5.31, SD=1.10). This corresponds to our observations that users had difficulties estimating the scale of distant objects and that relocating a Photoportal capturing position in a scene can be cumbersome if the scene is shown at a large scale. The ease of using the Photoportals to create cutaway views was rated on a similar level with an average score of 5.43 (SD=1.28). In practice, it requires the same operations to create a cutaway view as any other adjustment of the perspective. When asked to decide between cutaway views in plane or box mode, all but one voted for the box view. The ease of creating object copies was rated with an average score of 5.25 (SD=1.39). We observed some users having difficulties remembering that the trigger button had to be pressed in order to create a copy rather than removing the object from the portal.

The two remote manipulation techniques were rated differently. The average score for remote manipulation of objects in the 3D box mode was 5.5 (SD=0.93). The constrained remote manipulation in 2D box mode was still rated positively, but worst among all of our functionalities (M=4.63, SD=0.92). During the discussions, one participant mentioned that, for accurate placement, the input should be further constrained (e.g. only 2D translation) and that the orthogonal viewing in 2D box mode can be hard to understand in complex scenes. We observed that our participants tended to use the portals to navigate to the optimal position for manipulation rather than using it for remote manipulation.

All but one participant agreed that using Photoportals in a multi-user environment like ours can improve the collaboration of participants. One participant was neutral on that question; the average score was 6.13 (SD=1.36). Some participants argued that the overall system clearly promotes collaboration, but that it takes more than a few days' experience to evaluate the impact of individual aspects like the Photoportals on the collaborative use of the system.

We observed that two aspects of Photoportals for collaboration were immediately used: perspective sharing and territoriality. Our test users frequently assumed Photoportals as temporary private spaces (territoriality) and used them to prepare perspectives towards interesting parts, sometimes while somebody else was steering the rest of the group through the environment. These scene references were then shown to others (perspective sharing) – in most cases they referred to detailed scene features or suggested a new target location. More closely-coupled collaboration, like preparing a reference view for another user, occurred only rarely. Instead of asking each other for such support, most users wanted to operate the device themselves – despite being otherwise busy.

DISCUSSION

Our collaborative virtual reality system supports collocated interaction in a shared 3D environment and the inclusion of remote participants through 3D video avatars. We experienced that these characteristics provide an overall high level of workspace awareness, but that it can still be hard to understand and follow each other's actions in the virtual world if these go beyond real-world behavior.

This deficit and the desire to better support parallel activities led us to the design of a novel interface metaphor: Photoportals provide powerful reference-based interaction techniques for collaborative settings. Primarily, they support the communication about details of a shared virtual environment. Users can immediately capture objects, locations and views they want to discuss. Motion sequences can also be recorded. The resulting virtual photos and videos can then be shown to others directly or stored for later use. The physical camera device further promotes exchange between users. The interface also supports post-hoc adjustment of the view towards the captured scenes, including scaling, which helps to emphasize relevant aspects.

The novel interface metaphor also supports the navigation requirements of groups in virtual environments. It only takes several seconds to create an appropriately scaled and oriented WIM that can be used for wayfinding. The possibility of entering locations of interest directly through the virtual photo reduces the need for traveling long distances, which can be particularly cumbersome for groups. Furthermore, users can enter saved versions of the scene which can also include the captured movements of objects, themselves and other users. This allows them to travel into the past of a design process, create alternative versions or observe an interaction sequence including all the movements of the users. A further important aspect in that regard is the possibility of using the interface as a private display or territory. If a user wants to inspect a particular location within the scene that might not be relevant for everybody else, a Photoportal can be used to provide an individual view of this place.

Photoportals can also be used for object manipulation. In this context, they can be considered as an amalgamation of Voodoo Dolls and tangible props. Like Voodoo Dolls, our Photoportals clearly show what they are referring to, while grasping the physical capturing device facilitates the handling of these references. Our technique contributes further advantages for collaborative scene editing. The capturing of a remote object or location for manipulation does not yet change anything in the shared scene. It only manifests the intention and provides a reference to the relevant part of the scene. The collection of photos thereby provides a storage space for planned activities. The actual editing can be postponed to a moment that fits into the collaborative workflow and allows participants to discuss the intended changes. Moreover, if the group ultimately does not agree on the applied changes, they can use Photoportals to return to an earlier captured version.

Our interfaces and interaction techniques have been designed to encourage the engagement of each individual through multiple access points and different levels of involvement. At the lowest level, each user can explore the shared scene from various angles by walking around. The consistency of the shared environment facilitates gestural communication about the application content. As a next step, one can take a camera device for capturing photos of interesting features and show them to the others. Alternatively, the pick ray can be used to select and manipulate objects in the scene. The group can also be navigated through the virtual environment. Such group-related operations are controlled through the centrally located Spheron device to increase awareness and minimize conflicts.

This selection of tools provides the users with complementary functionalities and thus invites cooperation. For example, the Photoportals provide references for manipulation with the pick ray and the Spheron can be coupled with the Photoportals. We observed that users easily distribute fully orthogonal tasks like navigation and manipulation amongst each other, but that more closely intertwined collaboration only occurs among users that are proficient with the involved technologies and know each other well.

CONCLUSION AND FUTURE WORK

We presented Photoportals as a consistent interface for reference-based 3D interaction techniques in distributed multi-user virtual reality systems. Photoportals allow the quick creation of references and provide easy access to stored ones. Our user experience workshop revealed their usability and showed indications of their benefits for collocated and remote group interaction. The development of our collaborative interaction techniques followed well-established guidelines from the field of computer-supported collaborative work. In particular, the tangible interaction framework of Hornecker and Burr guided our developments [21]. The consistency of the shared 3D space inhabited by locally and remotely collaborating users leverages workspace awareness and promotes full-body interaction.

Our Photoportals allow users to take photos and videos of the surrounding scenery while they are collaboratively exploring a virtual environment. Recordings can be shared among collocated and remote users and serve as portals into the captured 3D scenes. Photoportals provide additional access points for interaction, minimize the issue of fragmented visibility and provide unambiguous references for communication, manipulation and navigation. They also provide users with a private interaction space where they can prepare views that refer to individual areas of interest. Photoportals can also refer to alternate versions of the scene in which users can explore their ideas for modifications. In that sense, Photoportals support the externalization of ideas that can be collaboratively discussed before they are applied to the shared virtual environment.

Given the development in 3D scanning, light field acquisition and display technology, we envision that, in the future, static real-world photos will only be placeholders for interactive portals to the captured scenes in a simulated mirror world and suggest virtual reality as a development platform for appropriate user interfaces.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the EC FP7 projects 3D-PITOTI (ICT-2011-600545) and VR-HYPERSPACE (AAT-2011-RTD-1-285681). Our work was also supported by the German Federal Ministry of Education and Research (BMBF) under grant 03IP704 (project Intelligentes Lernen). The 3D model of Castle Vianden (Luxembourg) that can be seen in Figures 4-8 is courtesy of ArcTron 3D GmbH (`http://www.arctron.com`). We thank the participants of our studies and the members and students of the Virtual Reality Systems group at Bauhaus-Universität Weimar (`http://www.uni-weimar.de/medien/vr`) for their support. Special thanks to Stefanie Wetzel, who realized the first software prototype of Photoportals.

REFERENCES

1. Agrawala, M., Beers, A. C., McDowall, I., Fröhlich, B., Bolas, M., and Hanrahan, P. The two-user responsive workbench: support for collaboration through individual views of a shared space. In *Proc. of Siggraph 1997*, SIGGRAPH '97, ACM Press/Addison-Wesley Publishing Co. (1997), 327–332.
2. Argelaguet, F., Kulik, A., Kunert, A., Andujar, C., and Froehlich, B. See-through techniques for referential awareness in collaborative virtual reality. *Int. J. Hum.-Comput. Stud. 69*, 6 (June 2011), 387–400.
3. 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. In *Proceedings of IPT '98, Workshop CD-ROM* (1998).
4. Barrus, J. W., Waters, R. C., and Anderson, D. B. Locales: Supporting large multiuser virtual environments. *IEEE Comput. Graph. Appl. 16*, 6 (Nov. 1996), 50–57.
5. Beck, S., Kunert, A., Kulik, A., and Froehlich, B. Immersive group-to-group telepresence. *IEEE Transactions on Visualization and Computer Graphics 19*, 4 (Apr. 2013), 616–625.
6. Benford, S., Bederson, B. B., Åkesson, K.-P., Bayon, V., Druin, A., Hansson, P., Hourcade, J. P., Ingram, R., Neale, H., O'Malley, C., Simsarian, K. T., Stanton, D., Sundblad, Y., and Taxén, G. Designing storytelling technologies to encouraging collaboration between young children. In *Proc. of CHI 2000*, ACM Press (2000), 556–563.
7. Bier, E. A., Stone, M. C., Pier, K., Buxton, W., and DeRose, T. D. Toolglass and magic lenses: the see-through interface. In *Proc. of SIGGRAPH 1993*, ACM Press (1993), 73–80.
8. Bricker, L. J., Baker, M. J., and Tanimoto, S. L. Support for cooperatively controlled objects in multimedia applications. In *CHI '97 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '97, ACM Press (1997), 313–314.
9. Brooke, J. SUS: A quick and dirty usability scale. In *Usability evaluation in industry*, P. W. Jordan, B. Weerdmeester, A. Thomas, and I. L. Mclelland, Eds., Taylor and Francis (London, 1996).
10. García, A. S., Molina, J. P., González, P., Martínez, D., and Martínez, J. An experimental study of collaborative interaction tasks supported by awareness and multimodal feedback. In *Proc. of VRCAI 2009*, ACM Press (2009), 77–82.
11. Greenberg, S., and Marwood, D. Real time groupware as a distributed system: concurrency control and its effect on the interface. In *Proc. of CSCW 1994*, CSCW '94, ACM Press (1994), 207–217.
12. Greenhalgh, C., Flintham, M., Purbrick, J., and Benford, S. Applications of temporal links: Recording and replaying virtual environments. In *Proc. of IEEE VR 2002*, VR '02, IEEE Computer Society (Washington, DC, USA, 2002), 101–.
13. Greenhalgh, C., Purbrick, J., Benford, S., Craven, M., Drozd, A., and Taylor, I. Temporal links: recording and replaying virtual environments. In *Proc. of ACM Multimedia 2000*, MULTIMEDIA '00, ACM Press (New York, NY, USA, 2000), 67–74.
14. Greenhalgh, C., Purbrick, J., and Snowdon, D. Inside massive-3: flexible support for data consistency and world structuring. In *Proc. of CVE 2000*, CVE '00, ACM Press (New York, NY, USA, 2000), 119–127.
15. Gross, M., Würmlin, S., Naef, M., Lamboray, E., Spagno, C., Kunz, A., Koller-Meier, E., Svoboda, T., Van Gool, L., Lang, S., Strehlke, K., Moere, A. V., and Staadt, O. blue-c: a spatially immersive display and 3d video portal for telepresence. *ACM Trans. Graph. 22*, 3 (July 2003), 819–827.
16. Grossman, T., and Balakrishnan, R. Collaborative interaction with volumetric displays. In *Proc. of CHI 2008*, CHI '08, ACM (2008), 383–392.
17. Guiard, Y. Asymmetric Division of Labor in Human Skilled Bimanual Action: The Kinematic Chain as a Model. In *Journal of Motor Behavior 19* (1987), 486–517.
18. Gutwin, C., and Greenberg, S. A descriptive framework of workspace awareness for real-time groupware. *Comput. Supported Coop. Work 11*, 3 (Nov. 2002), 411–446.

19. Hinckley, K., Pausch, R., Goble, J. C., and Kassell, N. F. Passive real-world interface props for neurosurgical visualization. In *CHI '94: Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (1994), 452–458.

20. Hindmarsh, J., Fraser, M., Heath, C., Benford, S., and Greenhalgh, C. Object-focused interaction in collaborative virtual environments. *ACM Trans. Comput.-Hum. Interact. 7*, 4 (Dec. 2000), 477–509.

21. Hornecker, E., and Buur, J. Getting a grip on tangible interaction: a framework on physical space and social interaction. In *Proc. of CHI 2006*, ACM Press (2006), 437–446.

22. Hornecker, E., Marshall, P., Dalton, N. S., and Rogers, Y. Collaboration and interference: awareness with mice or touch input. In *Proc of CSCW 2008*, ACM Press (2008), 167–176.

23. Jones, A., McDowall, I., Yamada, H., Bolas, M., and Debevec, P. Rendering for an interactive 360 degree light field display. *ACM Trans. Graph. 26* (July 2007).

24. Klinkhammer, D., Nitsche, M., Specht, M., and Reiterer, H. Adaptive personal territories for co-located tabletop interaction in a museum setting. In *Proc. of ITS 2011*, ITS '11, ACM Press (2011), 107–110.

25. Kopper, R., Ni, T., Bowman, D. A., and Pinho, M. Design and evaluation of navigation techniques for multiscale virtual environments. In *VR '06: Proceedings of the IEEE conference on Virtual Reality*, IEEE Computer Society (2006), 175–182.

26. Kulik, A., Kunert, A., Beck, S., Reichel, R., Blach, R., Zink, A., and Froehlich, B. C1x6: a stereoscopic six-user display for co-located collaboration in shared virtual environments. *ACM Trans. Graph. 30*, 6 (Dec. 2011), 188:1–188:12.

27. Lewis, J. R., and Sauro, J. The factor structure of the system usability scale. In *Proc. of HCI International 2009*, HCD 09, Springer-Verlag (2009), 94–103.

28. Maimone, A., and Fuchs, H. Encumbrance-free telepresence system with real-time 3d capture and display using commodity depth cameras. In *Proc. of ISMAR 2011*, IEEE Computer Society (2011), 137–146.

29. McGrath, W., Bowman, B., McCallum, D., Hincapié-Ramos, J. D., Elmqvist, N., and Irani, P. Branch-explore-merge: facilitating real-time revision control in collaborative visual exploration. In *Proc. of ITS 2012*, ITS '12, ACM (2012), 235–244.

30. Morris, M. R., Ryall, K., Shen, C., Forlines, C., and Vernier, F. Beyond "social protocols": multi-user coordination policies for co-located groupware. In *Proc. of CSCW 2004*, ACM Press (2004), 262–265.

31. Pausch, R., Burnette, T., Brockway, D., and Weiblen, M. E. Navigation and locomotion in virtual worlds via flight into hand-held miniatures. In *Proc. of SIGGRAPH 1995*, SIGGRAPH '95, ACM Press (1995), 399–400.

32. Pierce, J. S., Stearns, B. C., and Pausch, R. Voodoo dolls: seamless interaction at multiple scales in virtual environments. In *Proc of I3D 1999*, ACM Press (1999), 141–145.

33. Pinelle, D., Barjawi, M., Nacenta, M., and Mandryk, R. An evaluation of coordination techniques for protecting objects and territories in tabletop groupware. In *Proc. of CHI 2009*, CHI '09, ACM (2009), 2129–2138.

34. Pinho, M. S., Bowman, D. A., and Freitas, C. M. Cooperative object manipulation in immersive virtual environments: framework and techniques. In *Proc. of VRST 2002*, VRST '02, ACM Press (2002), 171–178.

35. Riege, K., Holtkamper, T., Wesche, G., and Frohlich, B. The bent pick ray: An extended pointing technique for multi-user interaction. In *Proc. of 3D UI 2006*, 3DUI '06, IEEE Computer Society (2006), 62–65.

36. Ruddle, R. A., Savage, J. C. D., and Jones, D. M. Symmetric and asymmetric action integration during cooperative object manipulation in virtual environments. *ACM Trans. Comput.-Hum. Interact. 9*, 4 (Dec. 2002), 285–308.

37. Schmalstieg, D., Fuhrmann, A., Hesina, G., Szalavári, Z., Encarnação, L. M., Gervautz, M., and Purgathofer, W. The studierstube augmented reality project. *Presence: Teleoper. Virtual Environ. 11* (February 2002), 33–54.

38. Scott, S. D., Sheelagh, M., Carpendale, T., and Inkpen, K. M. Territoriality in collaborative tabletop workspaces. In *Proc. of CSCW 2004*, CSCW '04, ACM Press (2004), 294–303.

39. Spindler, M., Büschel, W., and Dachselt, R. Use your head: tangible windows for 3d information spaces in a tabletop environment. In *Proc. of ITS 2012*, ACM Press (2012), 245–254.

40. Stoakley, R., Conway, M. J., and Pausch, R. Virtual reality on a wim: interactive worlds in miniature. In *Proc. of CHI 1995*, ACM Press/Addison-Wesley Publishing Co. (1995), 265–272.

41. Stoev, S. L., and Schmalstieg, D. Application and taxonomy of through-the-lens techniques. In *Proc. of VRST 2002*, ACM Press (2002), 57–64.

42. Tani, M., Horita, M., Yamaashi, K., Tanikoshi, K., and Futakawa, M. Courtyard: integrating shared overview on a large screen and per-user detail on individual screens. In *Proc. of Chi 1994*, CHI '94, ACM Press (1994), 202–.

43. Tse, E., Histon, J., Scott, S. D., and Greenberg, S. Avoiding interference: how people use spatial separation and partitioning in sdg workspaces. In *Proc. of CSCW 2004*, ACM Press (2004), 252–261.

44. Wingrave, C. A., Haciahmetoglu, Y., and Bowman, D. A. Overcoming world in miniature limitations by a scaled and scrolling wim. In *Proc. of 3D UI 2006*, 3DUI '06, IEEE Computer Society (Washington, DC, USA, 2006), 11–16.