System and Interface Framework for SCAPE as a Collaborative Infrastructure

Hong Hua[1], Leonard D. Brown[2], Chunyu Gao[2]
[1]*Department of Information and Computer Science, University of Hawaii at Manoa, Honolulu, HI 96822*
[2]*Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801*
Email: Hhua@hawaii.edu lbrown2|cgao@uiuc.edu

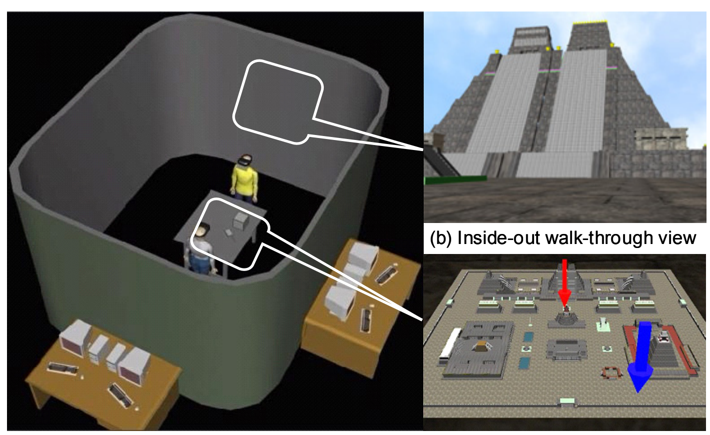
Abstract

We have developed a multi-user collaborative infrastructure, SCAPE (an acronym for Stereoscopic Collaboration in Augmented and Projective Environments), which is based on recent advancement in head-mounted projective display (HMPD) technology. SCAPE combines the functionalities of an interactive workbench and a room-sized immersive display to concurrently create both exocentric and egocentric perspectives. SCAPE intuitively provides a shared space in which multiple users can simultaneously interact with a 3D synthetic environment from their individual viewpoints, and each user has concurrent access to the environment from multiple perspectives at multiple scales. SCAPE also creates a platform to merge the traditionally separate paradigms of virtual and augmented realities. In this paper, we discuss the design principles we have followed to conceptualize the SCAPE system and briefly summarize SCAPE's hardware implementation. Furthermore, we discuss in detail the high-level design and implementation of the SCAPE architecture, and present a set of unique widget interfaces currently available in our implementation that enable and facilitate interaction and cooperation. Finally, we demonstrate SCAPE's unique visualization and interface capabilities via a testbed application—Aztec Explorer.

Keywords: Human computer interaction (HCI), Virtual reality (VR), Augmented reality (AR), Head-mounted display (HMD), Head-mounted projective display (HMPD), and Computer supported collaborative work (CSCW)

1. Introduction

There exists a large body of research efforts in the area of computer-supported collaborative work (CSCW) as well as work in tele-collaboration infrastructures and applications to facilitate collaborative interfaces [5, 18, 20, 21, 27]. Hollan and Stornetta [12] suggest that successful collaborative interfaces should enable users to go "beyond being there" and enhance the collaborative experience, instead of imitating face-to-face collaboration. Recent efforts have been made to develop tools and infrastructures to support collaboration in 3D virtual and augmented environments [2, 3, 24, 26].



(a) SCAPE display (c) Outside-in workbench view

Fig.1 SCAPE: stereoscopic collaboration in an augmented projective environment: (a) Concept illustration; (b) Simulation of inside-out egocentric walk-through view; (c) Simulation of outside-in exocentric workbench view.

We have developed a multi-user collaborative infrastructure, SCAPE (an acronym for Stereoscopic Collaboration in an Augmented and Projective Environment)

(Fig. 1) [17]. SCAPE, which is based on the recent development in head-mounted projective display (HMPD) technology, mainly consists of a workbench and a room-sized walk-through display, multiple head-tracked HMPDs, multi-modality interface devices, and an application-programming interface (API) designed to coordinate the components. It is capable of: (a) providing a non-distorted shared space in which multiple users can concurrently interact with a 3D synthetic environment from their individual viewpoints; (b) allowing each user to have concurrent access to the environment from multiple perspectives (both an egocentric inside-out view and an exocentric outside-in view) at multiple scales; (c) creating a platform to merge the traditionally separate paradigms of virtual and augmented realities in a single system; (d) enabling tangible interaction with a 3D environment and intuitive collaboration among a group of users.

The focus of this paper is to present the system and interface framework that enables SCAPE as an effective collaborative infrastructure. More specifically, we will discuss in detail the high-level design principles and guidelines that we have practiced to conceptualize and implement the SCAPE core architecture.

The rest of this paper is organized as follows: We will briefly review recent advances in 3D collaborative interfaces and recent development in HMPD technology in Section 2, describe SCAPE's conceptual design guidelines in Section 3, briefly summarize SCAPE's hardware implementation in Section 4, discuss in detail a set of design principles and implementation related to the system core architecture in Section 5, and present a set of unique interface modalities to enhance interaction and cooperation in Section 6. Finally, we will demonstrate SCAPE's key visualization and interface capabilities through a testbed application in Section 7.

2. Related Work

2.1. 3D Collaborative Interfaces

There are several different approaches to facilitating 3D collaborative work. An attractive and yet expensive solution is to use projection-based spatially immersive displays such as CAVE-like systems [6, 7, 8, 26] or the responsive workbench [25], which allows a number of users to concurrently view stereoscopic images by wearing LCD-shutter glasses. With these displays, users can see each other and therefore preserve face-to-face communication. However, the images can be rendered from only a single user's viewpoint, and therefore the stereo images are perspective-correct only for the tracked leader. The other non-tracked users will notice both perspective distortion and motion distortion.

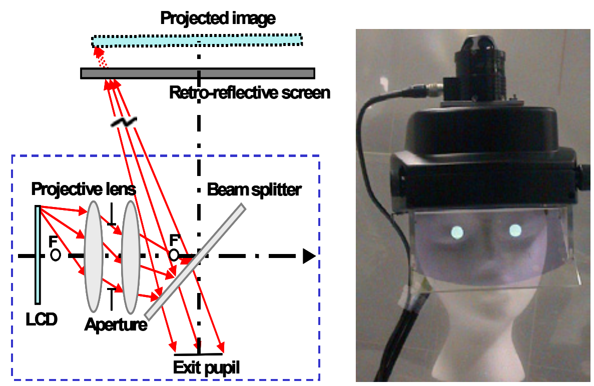
Several efforts have been made to overcome this limitation. Agrawala [1] proposed the two-user responsive workbench that allows two people to simultaneously view individual stereoscopic image pairs from their own viewpoints by using four different frame buffers. Two pairs of stereoscopic images are rendered sequentially at ¼ the display frame rate. The system cuts the display frame rate in half for each user compared to the single viewer approach, which leads to noticeable flicker and cross talk (e.g. 30Hz for each eye with ordinary display hardware having 120Hz maximum frame rate). Kitamura [24] proposed an alternative solution, namely IllusionHole, which allows three or more people to simultaneously observe individual image pairs from independent viewpoints without sacrificing frame rate. The IllusionHole display consists of a normal bench display and a display mask, which makes each user's drawing area invisible to others. However, the maximum number of users is limited, each user has a limited movement space, and the viewing area for each user is small.

Of great interest are systems that extend the VR-based paradigm by integrating physical objects into the physical workspace. Thus such augmented or mixed reality (AR-MR) interfaces facilitate the development of collaborative interfaces that go "beyond being there," while they also support seamless interaction with the real world, reducing functional and cognitive seams [2]. For example, Rekimoto [29] used tracked hand-held LCD displays in a multi-user environment and miniature cameras attached to the LCD panels to allow virtual objects to be superimposed on video images of the real world. Billinghurst [2] and Szalavari [31] proposed using see-through HMDs with head and body tracking in a collaborative interface, which allows multiple local or remote users to work in an augmented world. Bimber and his colleagues alternatively demonstrated the Virtual Showcase, which allows two or four tracked users to interact with the virtual content of the showcase while maintaining the augmentation of the virtual contents with real artifacts [3].

2.2. Head-Mounted Projective Display (HMPD)

Both the VR- and AR-based interfaces reviewed above typically address visualization from a perspective that is exclusively egocentric or exocentric. Such immersive displays as CAVEs and HMDs belong to the first category, and such semi-immersive displays as workbenches are of the second category.

Head-mounted projective display (HMPD), pioneered by Fisher [9] and Kijima & Ojika [23], is an emerging technology that can be thought to lie on the boundary between conventional HMDs and projective displays such as the CAVE systems [6]. An HMPD consists of a pair of miniature projection lenses, beam splitters, and displays mounted on the head and a supple retro-reflective screen placed strategically in the environment. Its monocular configuration is conceptually illustrated in Fig. 2-a. Unlike a conventional optical see-through HMD, an HMPD replaces eyepiece-type optics with a projective lens. Unlike a conventional projection display, an HMPD replaces a diffusing screen with a retro-reflective screen. An image on the miniature display is projected through the lens and retro-reflected back to the exit pupil, where the eye can observe the projected image. The uniqueness of a retro-reflective screen from a diffusing or specular surface lies in the fact that a ray hitting the surface at an angle is reflected back on itself, in the opposite direction. Due to the essence of retro-reflection, the location and size of the perceived images projected from the HMPD are theoretically independent of the location and shape of a retro-reflective screen. Furthermore, the projected images are only visible from the optical pupil of the display. This property enables a shared workspace in which each user views a synthetic environment from his or her own unique perspective. More in-depth discussion of HMPD technology compared with traditional HMDs can be found in [13].



(a) (b)

Fig. 2 Head-mounted projective display (HMPD): (a) Conceptual illustration; (b) HMPD prototype.

The HMPD concept has been recently demonstrated to yield 3D visualization capabilities with a large-FOV, lightweight and low distortion optics, and correct occlusion of virtual objects by real objects [22, 19, 13]. It has been recognized as an alternative solution for a wide-range of augmented applications [28, 22, 19, 17]. A custom-designed ultra-light compact prototype was developed in [14, 15]. The prototype achieves 52 degrees FOV and weighs about 750 grams, with a 640x480 VGA color resolution. Figure 2-b shows the front view of the prototype with a Hiball 3000 sensor attached.

3. SCAPE: A Collaborative Infrastructure

The HMPD technology intrinsically enables the capability of creating an arbitrary number of individual viewpoints in a shared workspace. In such a shared workspace, each user views a synthetic dataset from his or her non-distorted perspective without crosstalk with other users, while basic face-to-face communications with other local users are also retained. The single-user HMPD technology can be readily extended to a collaborative infrastructure by deliberately applying retro-reflective surfaces in the workspace and integrating multiple head-tracked HMPDs and interaction devices. This section will describe the conceptual design of the SCAPE to enable multi-

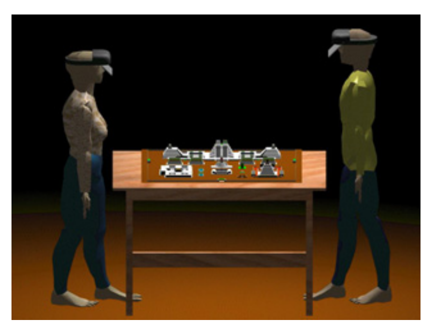


Fig.3 Illustration of an interactive workbench for collaboration.

scale collaborative visualization tasks.

A shared workspace based on the HMPD technology can potentially take many forms. One example is a multi-user interactive workbench environment (Fig. 3), whose surface is coated with retro-reflective film. Through the workbench display, each participant, wearing a head-tracked HMPD, is able to view and manipulate a 3D dataset from an individualized perspective. The workbench provides an outside-in perspective of a 3D dataset, in which the users can only explore the dataset from an exocentric viewpoint. Using the HMPD technology, it is also possible to create a CAVE-like room-sized workspace when egocentric perspectives, such as an immersive walk-through, are preferred. One or multiple walls are coated with retro-reflective film to create a shared workspace.

The first difference of the HMPD-based shared workspaces from the traditional CAVE and its kin is its capability of supporting an arbitrary number of non-distorted unique perspectives, which shares similarity with such systems as the two-user responsive workbench [1], IllusionHole [24], or Virtual Showcase [3]. If shared registration is properly achieved, when two users point to the same part of a dataset, their fingers shall touch. Furthermore, the ability to display multiple independent views offers the intriguing possibility of presenting different aspects or levels-of-detail (LOD) of a shared environment in each view. The second difference is that the combination of projection and retro-reflection in HMPDs intrinsically provides correct one-way occlusion cues: (1) computer-generated virtual objects is naturally occluded by real objects that are not coated with retro-reflective film; or (2) a user can see through real objects that are coated with retro-reflective film (Figures 11-b, 13-b, 14-a). Therefore, such HMPD-based shared workspaces allow augmenting a 3D synthetic dataset with physical objects or props which may be deliberately coated with retro-reflective material. This capability differentiates HMPD-based collaborative interfaces from those used traditional HMDs. In the meanwhile, it is worth to mention that one limitation of this one-way occlusion is that virtual objects will erroneously disappear if a virtual object is intentionally floating in front of a non-reflective real object.

Overall, either the workbench or the multi-wall display alone can only create a single perspective—an omni-present outside-in view for the workbench or an immersive inside-out view for the wall display—and a single scale of visualization (e.g. 1:1, minified, or magnified) with which to interact. This limitation of single perspective and single scale prevents a user from appreciating the larger context of the entire virtual environment. Stoakley et al addressed this concern in an HMD-based virtual reality system through a "World in Miniature (WIM)" metaphor [30]. Through a hand-held miniature WIM representation of a life-size virtual world, a user can interact with the environment by direct manipulation through both the WIM and the life-size world. Simultaneously the WIM representation also presents a second perspective of the virtual world. Their informal user studies show that an alternative view and scale of the visualization context can help users to establish spatial orientation in a virtual environment. In the WIM metaphor, however, the WIM perspective plays a supportive role to facilitate interaction with the immersive virtual world, which is the dominant context. Furthermore, the WIM metaphor is a single-user interface and does not emphasize collaborative aspects among a group of users in a shared space. An attempt to multi-scale collaborative AR interface by Billinghurst and colleagues in the MagicBook project explored the possibility of blending a user's experiences between reality and virtual reality by using a physical book as the main interface [Billinghurst]. While a user can read the book as normal, he or she can also see 3D virtual models appearing out of the pages through an HMD. The user can switch his or her viewing mode to fly into an immersive virtual environment to experience the story. HMD-based interface also allows multiple users to share the same MagicBook interface from individual viewpoint.

The conceptual design of SCAPE combines an interactive workbench with a room-sized display environment to create exocentric and egocentric perspectives simultaneously (Fig. 1-a). First of all, SCAPE intuitively provides a shared space in which multiple users can concurrently observe and interact with a 3D synthetic environment from their individual viewpoints.

Secondly, each user can have concurrent access to the synthetic environment from two different perspectives at two different scales such as an exocentric miniature view through the workbench (Fig. 1-c), and an egocentric life-size view through the room (Fig. 1-b). For convenience, we hereafter refer to the workbench view as the Micro-scene, and the immersive walk-through view as the Macro-scene. For example, the Macro-scene may be an expansive city with life-size buildings, and the Micro-scene can be a minified 3D map of the city (See testbed example in Section 7). Obviously the map can assist a user in exploring the city in many different ways such as navigation, path planning, distance estimation, and task coordination with collaborators. Conversely, the workbench may represent one-to-one scale and the room a magnified world. For example, consider an anatomy visualization task. On the workbench is projected a life-size human body and visualized through the immersive display is a greatly magnified view of the human vascular system; using the "molecular" scale of the immersive display, the user can thus travel within the pathways of individual blood vessels, while an indicator on the workbench shows relative anatomical location within the body. Moreover, not only in different scales and perspectives, the Micro-scene may also represent a different level of detail from that of the Macro-scene. Indeed, both the Micro- and Macro-scenes play equally important roles and they should seamlessly coordinate with each other.

Finally, SCAPE creates a platform to merge the traditionally separate paradigms of virtual and augmented realities. The workbench provides a means of performing augmentation tasks in which a Micro-scene may be registered with the physical workbench and objects placed on the bench, while the room provides a container for an expansive virtual environment which may be many times larger than the physical expansion of the room. Rather than switching from one to the other as in the MagicBook interface [ ], we attempt to seamlessly blend the multi-scale virtual and augmented interfaces to which a user can have concurrent access.

4. SCAPE Hardware Implementation

The SCAPE implementation is mainly affected by the characteristics of available retro-reflective materials suitable for screens. Practically, a retro-reflective material can only work well for limited angles. Imperfect reflective properties have direct or indirect impact on imaging characteristics and quality, and thus affect various aspects of the SCAPE design such as screen shape, screen distance and room-display size, field-of-view of the HMPDs, and environmental lighting. In-depth discussions on how the artifacts affect actual design were reported in [17].

The preliminary implementation of the SCAPE display environment currently consists of a 3'x5' workbench and a 12'x12'x9' four-wall arched cage made from retro-reflective film, multiple head-tracked HMPDs, multi-modality interface devices, computing facilities, and networking.

The shape of the cage is specified in Fig. 4-a; it is composed of four 6-foot flat walls and four arch corners with 3-foot radii. The height of the walls is 9 feet. The round corners, rather than squared corners as in CAVE systems, are designed purposely to minimize the gradual drop in luminance [17]. The walls and corners are all coated with the reflective film, and one of the corners is designed as a revolvable door. The enclosure allows full control of the environmental lighting. Naturally, a 6-wall display is possible if both the floor and the ceiling are coated with the film. Hiball3000 sensors by 3rdTech [www.3rdtech.com] are used for head tracking purposes, so our ceiling is installed with the 14'x14' array of LED

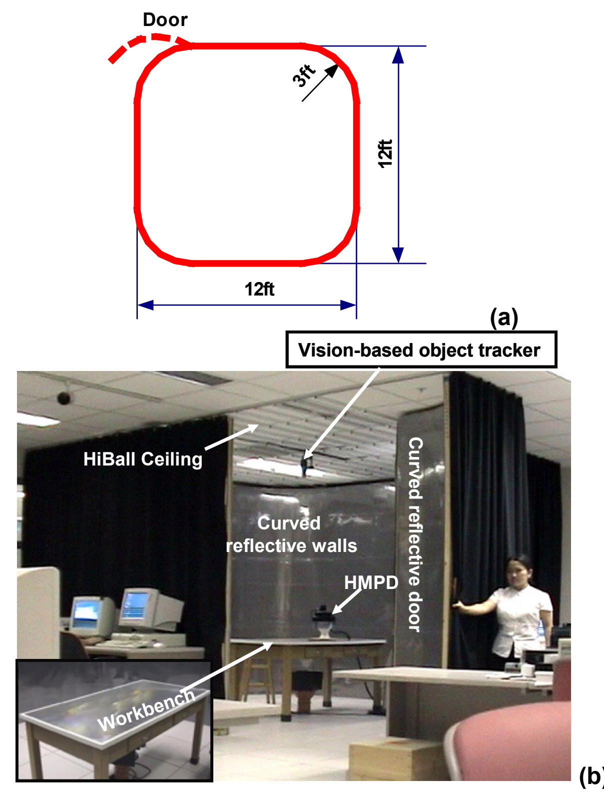


Fig.4 SCAPE implementation: (a) Shape and size specification of the room; (b) Experimental setup.

strips. Because of the minimal requirements on wall alignment and the low cost of the film, the expense in building the reflective cage is much less than that of building a CAVE. Figure 4-b shows the SCAPE setup.

Two HMPDs are driven by Dell Precision Workstations with P4 Dual Processors (Intel Xeon 2.4GHz) using NVIDIA Quadro4 900 XGL graphics cards. The head position and orientation of each user is detected by the Hiball3000 optical tracker. The stereoscopic image pairs are generated without distortion for each user according to their individual viewpoints.

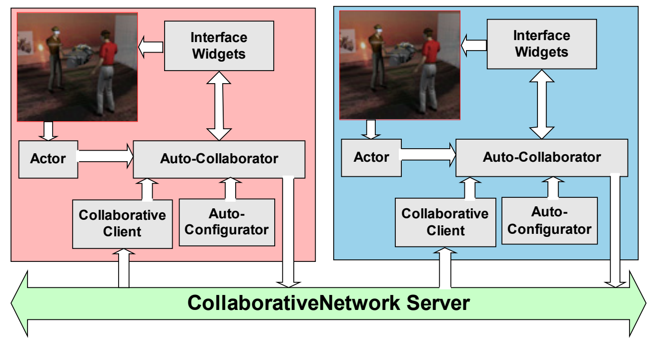


Fig. 5 Diagram of the SCAPE core architecture

In terms of interfaces, SCAPE employs a set of generic devices to manipulate and interact with virtual environments. An Ascension Flock-of-Birds (FOB) magnetic tracker is used to track moving objects such as hands or interface widgets. A tracked 5DT Data Glove [www.5dt.com] is used to manipulate 3D virtual objects on the bench and to navigate the walk-through immersive environment in the room (see the Aztec application example in Section 7). Besides these generic interface modalities, we have developed a set of unique augmented widgets to facilitate interaction and collaboration. These widget interface modalities will be described in Section 6.

5. SCAPE Core Architecture

SCAPE opens up new possibilities as well as challenges for design approaches to system architecture and user interfaces over traditional collaborative infrastructures reviewed in Section 2. For example, how can we maintain seamless integration between the Micro- and Macro-scene views for each individual user? Switching from one view to the other by some physical push-button will certainly jeopardize both functional and cognitive integration. In a collaborative application, should we grant each user an equal accessibility to the entire environment or should we grant one user higher priority than the others? Given that a large community of users and extremely intricate system configuration may be involved in a complex networked application, it becomes essential to deal with such issues as user management and system calibration.

Enabling SCAPE as an effective collaborative infrastructure is a custom-designed application-programming interface (API) referred to as the SCAPE Toolkit. The Toolkit is a cross-platform, modular, and extensible core framework providing high and medium-level programming control over the SCAPE workspace to enable augmented collaboration (Fig. 5). This core framework manages various aspects of the SCAPE workspace from networking, users, and interfaces to collaboration. We concentrate this discussion on four higher-level controls that facilitate interaction and collaboration: a transformation hierarchy that enables seamless integration of multi-scale multi-perspective visualization, collaboration modes that control various aspects of collaboration, an Actor interface that manages users and their priority, and an Auto-configurator module that calibrates and configures an application.

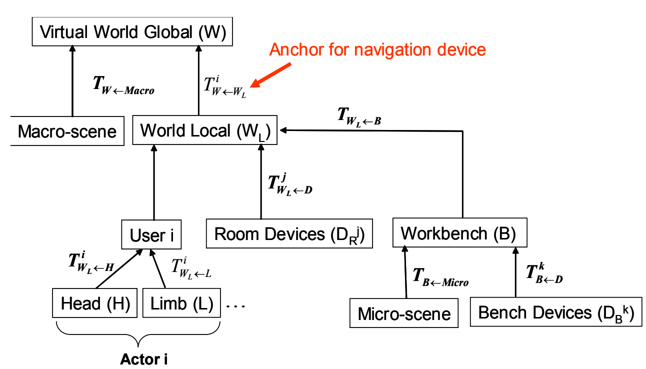


Fig. 6 SCAPE transformation hierarchy

5.1 Transformation Hierarchy

As we discussed in Section 3, SCAPE combines two scales of visualization from two perspectives, namely the Micro-scene from an exocentric workbench view and the large-scale Macro-scene from an egocentric immersive view. The Micro-scene can further be considered as an augmented view superimposed upon the physical workbench and objects placed on the bench, and the Macro-scene can be significantly larger than the physical extents of the room. Therefore, absolute sensor measurements of a user's head and hand as well as objects are required in a one-to-one scale to render the Micro-scene view, while relative or scaledsensor measurements are necessary to render the Macro-scene view beyond the room. It is essential to have an intuitive "Transport mechanism" to coordinate the two different scales of visualization.

SCAPE's transformation hierarchy (Fig. 6) is such a transport mechanism to maintain concurrent seamless integration between the different views and scales, without the necessity of switching from one to the other. At the root of the hierarchy is the virtual world global coordinate system (W), which is the universal reference governing the rest of the components in the environment. The scale of the global world should be determined by application contexts. The Macro-scene, residing in the W_{Macro} reference, is defined as an entity in the global reference with a composite transform $T_{W \leftarrow Macro}$. Within the global world context, we define a World Local (W_L) reference corresponding to the physical extents of the SCAPE room display. This local reference serves as a container to encapsulate the physical-related entities such as workbench, users, interface devices, and micro-scene. Within this local world context is the physical reality. The spatial relationships of all the physical-related entities are measured in one-to-one physical scale. Some of the physical entities such as users, workbenches, and room-related interface devices are defined relative to the world-local reference through their corresponding transformations. In a multi-user environment, this arrangement makes the device transformations independent of user association and allows flexibility of reconfiguring the overall system. Users attached with head trackers may walk physically with the extents of the SCAPE room to explore his or her world-local context. Other entities such as Micro-scene and bench-related devices are defined relative to the workbench reference W_B.

The local world reference may be anchored or transformed arbitrarily within the higher-level global-world context by manipulating a transform $T_{W \leftarrow W_L}$. This arrangement is analogous to driving a "car" in a virtual world. Inside the vehicle is the physical reality, looking through the vehicle's window is a virtual world, and driving the vehicle transports users in the virtual world.

The transport analogy described above can be achieved with a typical interface device used for traveling in large-volume virtual environments such as a wand, Data Glove, or six degree-of-freedom mouse. In our implementation, we combine two means of "Travel" to drive the vehicle. A user wearing a 5DT Data Glove [www.5dt.com] can nudge his or her position continuously forward and backward with simple hand gestures. Alternatively, we have designed a vision-based object tracking method that is capable of recognizing and tracking simple objects such as a number-coded ID marker placed on the workbench. A user can manipulate his or her world local reference in the global world by simply moving his or her physical "ID" on the workbench (Fig. 10). While a user can invoke the two means at will, the Data Glove interface enables fine-grained navigation and the ID marker enables rapid maneuvering to a largely different region.

5.2 Collaboration Modes

The transformation hierarchy described in the last sub-section is appropriate for one single user. In a multi-user collaborative environment, a fundamental viewpoint management question has to be addressed. Innately, SCAPE provides the capability of allowing each user to have equal access to a simulation. However, should we grant each

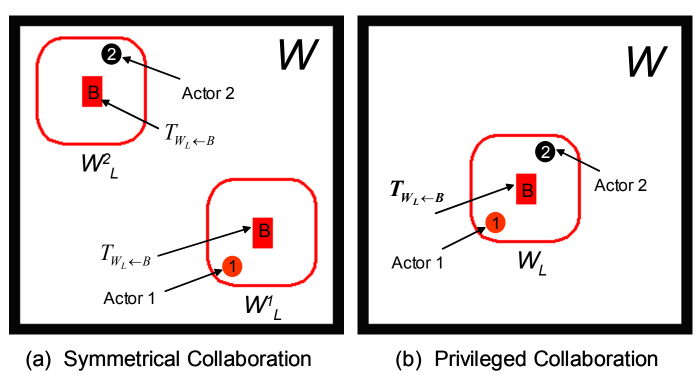


Fig. 7 Collaboration modes: Symmetrical vs. privileged.

local user an equal accessibility to the entire environment or should we grant one user higher priority than the others? In other words, a choice has to be made between symmetrical collaboration and privileged leader-mode collaboration. In symmetrical collaboration (Fig. 7-a), each user (i) has an individual anchor $T^i{}_{W \leftarrow W_L}$ to control his or her world-local location in the Macro-scene as well as his or her viewpoint in the world-local environment. In a privileged mode (Fig. 7-b), there is only one world-local anchor and a leader of the group controls the anchor. Reflecting the "car" analogy, the symmetrical mode is analogous to the case in which each user drives his or her vehicle individually, and the privileged mode is analogous to the case in which all users carpool and only one driver controls the vehicle. Different from the leader-mode in a traditional CAVE-like environment, each user has individual control of his or her viewpoint in the world-local environment. In both modes, we can apply filters to partition information into different layers so that users can actually access different layers or combinations of visualization. There are pros and cons for these two different modes.

Symmetrical collaboration provides each user equal control and accessibility, and consequently has more flexibility and self-control from a user's point of view. Therefore multiple tasks can be performed by individuals in parallel. Users can start their journey from different regions and they can "jump" from one area to the other. This parallel maneuvering capability is particularly important for mission-oriented applications, for example, searching for military targets in a large area. In another case, participants may have different expertise and different assignment, thus they are not necessarily interested in the same area in terms of spatial partition.

One of the disadvantages is that the symmetrical mode requires more interface resources. Each user needs to own "Travel" gadgets and controls his or her own vehicle. Another potential issue in symmetrical mode is perceptual contradiction. Potentially, there exist two types of contradiction. In one scenario (Fig. 8-a), users 1 and 2 are facing each other in the physical world, but they are looking in opposite directions in the Macro-scene. In another scenario (Fig. 8-b), they look away from each other in the physical world but are facing each other in the virtual world. These contradictory visual cues could cause spatial disorientation and other perceptual problems.

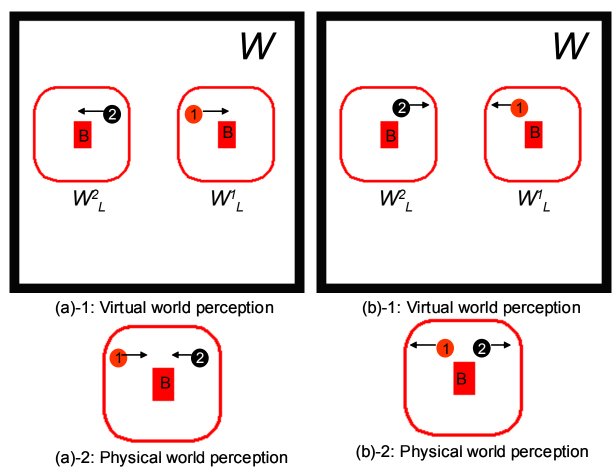


Fig. 8 Perceptual contradiction in symmetrical collaboration mode

On the other hand, in applications that have natural leadership or supervision requirements, the privileged mode has advantages over the symmetrical mode. For security reasons, a leader can supervise the pace of a process and controls accessibility to some sensitive resources and regions. For example, in a training program, the instructor may have access to more detailed information than students. Locking the group attention to the same context may also encourage more convenient group discussions and collaboration.

The SCAPE toolkit includes an Auto-Collaborator module to encapsulate the constructs above for a multi-user collaborative application. In the current Toolkit, we have only implemented the modes of symmetrical collaboration. The Auto-Collaborator will provide default support for both modes of onsite collaboration. Automation and packaging are still in a preliminary state of implementation. Indeed, we can possibly implement other collaboration modalities and allow users to configure an appropriate mode based on application needs. Users may also switch among the modes during a collaboration session.

5.3 User Management via Actor

Complex interactive, collaborative environments require a cohesive structure for maintaining devices and information specific to each user. The SCAPE Toolkit employs a high-level object associated with each user, called an Actor, to encapsulate all the real and virtual components of that user. Each Actor maintains its corresponding user's viewpoints into the multiple scales of visualization, interface devices, coordinate systems and transformations, as well as other user-related public and private data. For reasons of security, ethics, or convenience, we do not presume the symmetric access of all users to all data. Hence we limit the accessibility of certain data and devices by constraining their ownership. The Actors may be classified into three categories: guest Actor, power Actor, and super Actor.

Fundamentally, a guest Actor only inherits the basic accessibility to user-specific devices, private data, and components of the public scenegraph. Except his or her user-related status, a guest Actor may not be allowed to manipulate virtual objects, or modify any system-related status. For example, the guest Actor category is appropriate for a collaborator who only passively observes visualization, or for a user who has minimal accessibility and control of the visualization. Besides the basic accessibility, a power Actor has a wide range of ownership, accessibility, and interface options. For example, a power Actor is able to manipulate public virtual objects, possess certain interface widgets, and access certain privileged data. We can also group power Actors such that a subset of Actors can confer on specific privileged data as a group, independent of the larger Actor community, by allowing multiple ownership of the same privileged data. A super Actor, like a system administrator possessing "root" privileges in UNIX, has access and control of all levels of data and can override the actions of other Actors. For example, a super Actor can assign or suspend ownership to widget interfaces, control system status, and switch collaboration modes. There is only one super Actor present in an application, but a sub-group may have a group super Actor.

A hierarchical organization of the Actor community is illustrated in Fig. 9. Indeed, it demonstrates more intricate user relationships beyond the collaboration modes discussed in the previous section. In the actor community, we have a unique super Actor who supervises the community. Other Actors can be members of a group (i.e. as children of a group node) or can behave individually (i.e. as children of the root node). Within a group, the actors can collaborate symmetrically or otherwise asymmetrically. In the case of symmetric collaboration within a group, the actors equally control the group behavior. In the case of asymmetrical collaboration within a group, a leader naturally becomes the group super Actor.

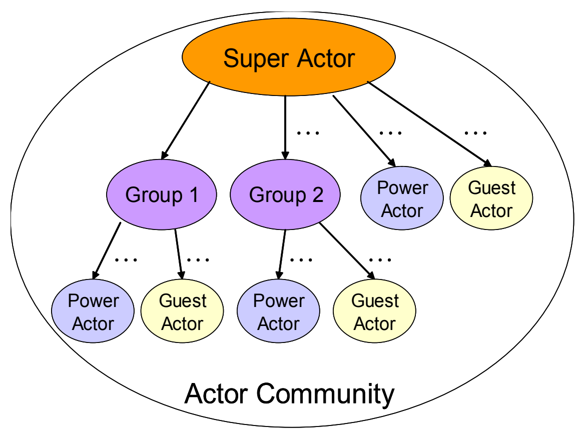


Fig. 9 Hierarchical organization of Actor community.

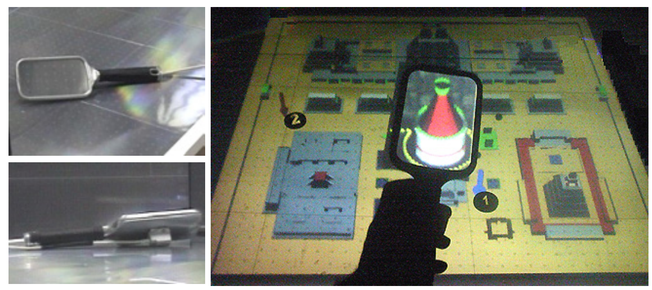
In this methodology, the states of certain scenegraph components are maintained and updated within specific Actors via user-defined behaviors. The private data are then loaded onto the scenegraph exclusively for the rendering of the particular owners' views, and remain unloaded otherwise. In the case of augmented widgets, Actors not owning a widget will see no virtual component when they manipulate the widget's physical device; for them, the widget is essentially "turned off." The ownership requirement also suggests that a widget may identify and interface intelligently with each Actor, restoring unique saved state or preferences from previous encounters.

5.4 System Calibration and Auto-Configurator

In SCAPE, each user is provided individual views into a shared multi-scale visualization. In order to maintain a shared synthetic environment with which to interact, proper calibration of the hardware is required so that the synthetic representations are consistent and continuous for all users from arbitrary perspectives. This requires the coordinate systems in the SCAPE transformation hierarchy to be properly aligned. This is referred to as the registration process.

The registration process takes three major steps: (1) Determining transformations that define the spatial relationships of all physical objects including workbench and all the tracking devices relative to the world-local reference; (2) Determining intrinsic and extrinsic parameters of each HMPD's viewing optics; and (3) Obtaining the viewing orientation and projection transformations for each user, based on viewing optics parameters, to generate view-dependent stereoscopic image pairs and to align the references. We have been using different types of trackers in our experiments. The first step involves individual calibration of each tracking system relative to the world-local reference. For the less accurate magnetic trackers, look-up-table calibration methods [10] may be used to compensate for the large magnetic distortion. The second and third steps involve a complex procedure to individually calibrate each HMPD system and a process to match the extrinsic and intrinsic viewing parameters of the virtual cameras in graphics generator for each viewer with those of his or her viewing device. We have developed systematic calibration methods to perform HMPD display calibration and a computational model for applying the estimated display parameters to viewing and projection transformations. Details about the calibration methods and procedures can be found in [16, 11].

In the SCAPE Toolkit, we have implemented methods for establishing an accurate computational model from the intrinsic and extrinsic parameters and for customizing the viewing and projection transformations for each user to generate their corresponding image pairs. The SCAPE ToolKit implements an Auto-Configurator class that enables stock program configuration options, including system configurations, networking, display parameters obtained through the calibration process, Actor interface options, interface and widget options, and collaboration modes. Currently, the calibration methods are implemented separately in Matlab code. In future work, we anticipate integrating the calibration functions into the Auto-Configurator module and automating the procedure.



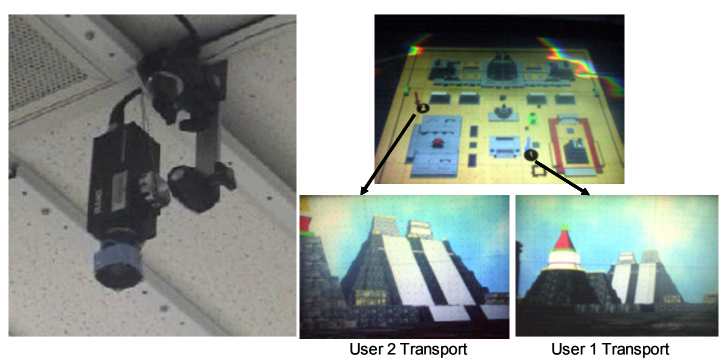
(a) **(b)**

Fig.11 Magnifier widget: (a) Implementation of 'Magnifier' device; (b) "Magnifier" at work.

6. SCAPE Interface Framework

In a collaborative context, interface designs are required to facilitate collaborative needs and to enhance collaborative experiences. For example, there is such a scenario during symmetrical mode collaboration when participants are virtually far apart but physically in a reachable distance. How do they effectively share data and views without changing their virtual locations? Besides the Micro- and Macro-scenes, we should also consider intermediary representations to facilitate user interaction with 3D contents that are either at low levels-of-detail, too large to manipulate, or far from reach. These open issues and challenges inflence the design principles we have kept in mind and practiced in the SCAPE implementation.

Besides a set of generic interface devices such as head tracker, hand tracker, and Data Glove, in SCAPE, we have developed a set of unique augmented devices or widgets to facilitate interaction and collaboration. They currently include a vision-based object tracker, Magnifier, CoCylinder, and CoCube. The object tracker interface allows augmentation and navigation in the immersive Macro-scene, while the rest of the widgets are designed to support intermediate levels of visualization between the Macro-scene and Micro-scene and to facilitate cooperative interfaces. The following paragraphs



(a) **(b)**

Fig.10 Vision-based object tracker: (a) Experimental setup; (b) Tracking user IDs in Aztec Explorer.

summarize the implementation and functionality of the widgets, and an in-depth discussion and implementation details can be found in [4].

Vision-based Object Tracker: To support augmentation of virtual objects with physical ones and to enable tangible interaction with the virtual world, we have developed a 2D vision-based object tracking method to recognize and track physical objects placed on the workbench. An infrared camera with infrared lamps mounted on the ceiling continuously captures the image of the objects placed on the bench (Fig. 10-a). Segmentation algorithms are applied to group and recognize the objects and to determine their 2D position and orientation. Under different application contexts, this tracking method with minor modification can be used to track multiple physical objects in augmented environments, recognize simple hand gestures to interact with virtual environments without special attachments or hand markers, and develop widgets to facilitate cooperation among multiple users. In particular, by identifying and tracking a number-coded user ID marker registered with the Micro-scene on the workbench, the tracking methods enable a user to control his or her anchor in the global world and to navigate "himself" or "herself" through the Macro-scene (Fig. 10-b). In a multi-user environment, each user owns an ID marker and the tracking method is capable of recognizing them in real-time. We anticipate extending this tracking method to support 3D tracking and more complicated objects by integrating multiple cameras.

Magnifier Widget: Given the context that the workbench presents a miniature visualization of a 3D dataset at a low level-of-detail, we have developed a "Magnifier" widget allowing a user to examine detailed views of the virtual data on the workbench via the lens inset without the need to directly retrieve the corresponding Macro-scene. The Magnifier is a hand-held device coated with retro-reflective film, with a motion tracker attached (Fig. 11-a). A virtual magnifier camera is associated with the Macro-scene, which is at a higher level-of-detail than the bench view. While moving the magnifier around above the bench, a user perceives a magnified view superimposed on the bench view corresponding to the image captured by the magnifier's virtual camera (Fig. 11-b). Thus, the magnifier metaphor naturally creates a through-the-window visualization at a medium level of detail that lies between the immersive Macro-scene and semi-immersive Micro-scene.

CoCylinder widget: As an alternative means of visualizing life-size artifacts, we have constructed a large cylindrical device whose surface is coated with retro-reflective film into which a life-size object is projected (Fig. 12). The cylindrical display measures 48 inches tall with a diameter of 15 inches. The display is installed on a rotation stage with an Ascension FOB sensor to measure the display's azimuth rotation. Therefore, this device intuitively allows collaborators encircling the display to concurrently view and manipulate the virtual object by physically walking around the device. This device also enables tangible interaction with the virtual object itself by physically rotating the display. Within the SCAPE context, collaborators can "capture" a virtual object from either the Micro-scene or Macro-scene and fit it into the cylindrical volume for convenient interaction and cooperation.

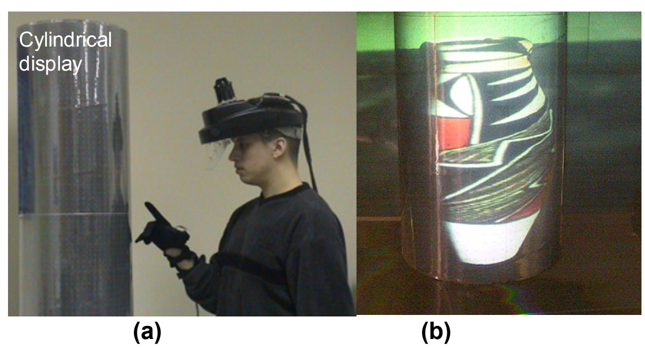
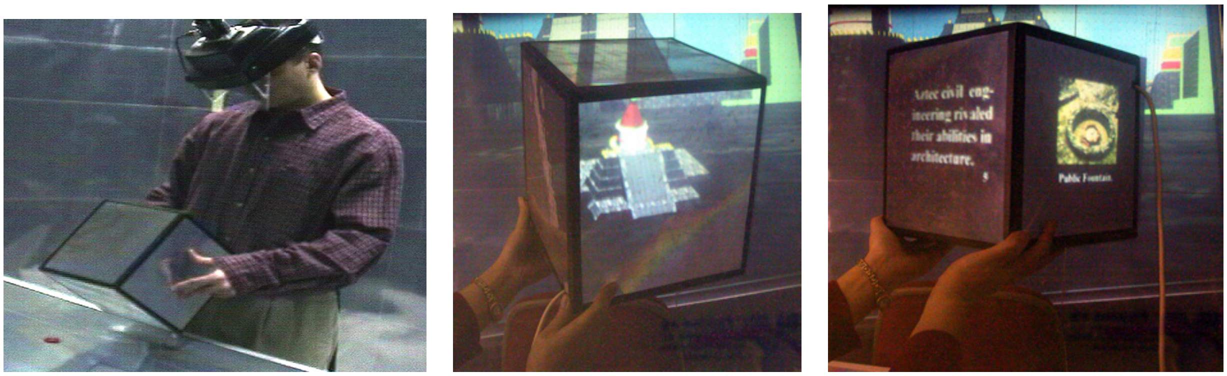


Fig.12 CoCylinder widget: (a) Device implementation; (b) 'CoCylinder' at work.

CoCube widget: To facilitate cooperative interaction in SCAPE environments, we have constructed a CoCube widget. This widget's hardware consists of a handheld 10-inch cube coated in retro-reflective film, with framed, diffusing edges attenuating the reflective viewing surfaces (Fig. 13-a). Attached to the inside of the Cube is a FOB sensor. The Co-Cube has implemented two distinct modes of interaction: selection and inspection. In the selection mode, the device allows a user to "capture" a large or distant virtual object from his or her surrounding Macro-scene through a ray-casting analogy (Fig. 13-b). The selected object is minified to fit within the cube volume, and thus the user can inspect the object from an exocentric viewing perspective. Similar to the CoCylinder, the CoCube widget allows a user to frame a virtual object from the Macro-scene into the cube device and share it with other collaborators via their unique perspectives. The virtual workspace of multiple users does not necessarily overlap as they may be exploring different regions of the Macro-scene or accessing different layers of information. Therefore, the CoCube device can be used as a tool to relay information from one user's workspace to the others, and thus grounds their cooperative activities.



(a) (b) (c)

Fig. 13 CoCube widget: (a) Implementation of "CoCube" device; (b) Object captured from Macro-scene; (c) Retrieval of documentary information from the selected object.

7. Testbed Application: Aztec Explorer

In this section, we present a testbed example—Aztec Explorer—to demonstrate some of the SCAPE characteristics, the API framework, and some aspects of the interface and cooperation features we have implemented.The testbed features a scale model of Tenochtitlan, an ancient Aztec city. The 3D scenegraph is modified from a freeware mesh obtained from 3DCAFE [www.3dcafe.com] and we have enhanced it with texturing mapping and created multiple levels-of-detail. Visualized through the workbench is a low LOD Micro-scene rendered only with Gouraud shading (Fig 14-a), and visualized through the SCAPE room display is a high LOD Macro-scene rendered with texturing mapping at one-to-one physical scale (Fig 14-b). Two individual viewpoints (capable of unlimited users if resources are available) are currently rendered for two head-tracked users. Users can either discuss the Aztec city planning with the other participants through the workbench view, or explore its architectural style via the walk-through. The two users collaborate in the symmetrical mode.

During bench-view collaboration, the users share exactly the same Micro-scene but from individual perspectives, and therefore collaboration takes place in an intuitive face-to-face manner. They can simply point to a temple via hand to direct the group's focus of attention. The "Magnifier" widget (Fig. 11) is shared among users with ownership and allows a user to closely examine the magnified view of particular temples (Fig. 11-b, 14-a). The Macro-scene is a fully immersive life-size environment (Fig. 14-b), which measures two kilometers across. There are three distinct but seamlessly combined methods to navigate the expansive virtual world as discussed in Section 5.1. A user may walk around physically within the extents of the SCAPE room to explore his or her world-local context and his or her views are updated accordingly based on absolute measures of the head-tracker. The user wearing a 5DT Data Glove may also manipulate his or her world-local context W_L^i relative to the world-global reference by making pre-coded hand gestures. For example, a user can transport his/her world-local reference by making an "index finger point" gesture for "forward" or "thumb up" gesture for "backward", rather than physically walking "forward" or "backward", which overcomes the physical constraints on mobility. Each user is also assigned a unique physical ID, for instance, a numbered checker piece in our experiment (Fig. 14-c). The user can place his or her ID on the bench area which is registered with the Micro-scene. The vision-based object tracker described in Section 6 is capable of simultaneously recognizing multiple IDs and determining their 2D locations in the Micro-scene. Each user's ID location in the Micro-scene corresponds to a unique location in the Macro-scene (Fig. 10-b). Therefore, by manipulating his or her physical ID on the workbench, the user can instantly transport his or her world-local context W_L^i. While the head-tracker and Data Glove enable fine-grained navigation in the Macro-scene, the tangible ID metaphor is a transport mechanism to facilitate rapid navigation in the sufficiently large Macro-scene.

To provide the user and companions an awareness of his or her location, a virtual avatar (e.g. simply a color-coded arrow in Fig. 10-b) is created for each user in the Micro-scene and is visible in the bench view to all participants. Each avatar represents the current location of its associated user in the Macro-scene and is updated accordingly as he or she walks through the scene. The virtual avatars are registered properly with the ID checkers (not necessarily overlapped), and the bench view thus can be thought of as a shared map to explore the expansive city.

When multiple users need to confer with each other on a virtual structure such as a temple in the Macro-scene, they can use the CoCube widget to capture the temple from the Macro-scene. They can inspect and share the framed object by manipulating the physical cube (Fig. 13-b), and they can also optionally toggle to a documentary mode to read about the structure's history (Fig. 13-c). We have specified 14 buildings that can be individually captured via the CoCube widget.

Overall, Aztec explorer demonstrates SCAPE's unique visualization and interface capability: intuitively creating a perspective-correct shared workspace for multiple users; seamlessly integrating egocentric and exocentric perspectives at multiple scales; merging the traditionally separate paradigms of virtual and augmented realities; and interacting and collaborating with the synthetic environments through tangible widgets.

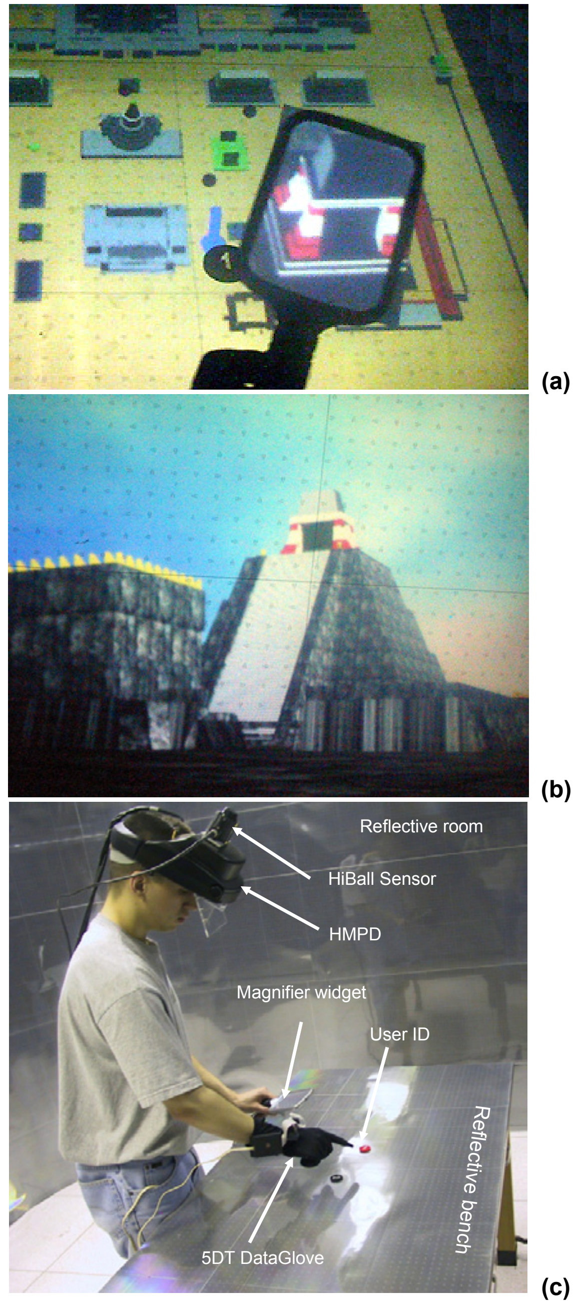


Fig. 14 Aztec Explorer: (a) Exocentric workbench view; (b) Egocentric walk-through view; (c) Experimental setup.

8. Conclusions and Future Work

We have developed a multi-user collaborative infrastructure, SCAPE, based on the head-mounted projective display (HMPD) technology. This article discussed the motivations and design principles we have followed to conceptualize the SCAPE system, described the current implementation of the SCAPE hardware, discussed the high-level design principles of the SCAPE framework, and summarized the unique widget interface modalities currently available in our implementation.

In the future, more efforts will be made to tackle some of the fundamental challenges in the SCAPE system. For example, as a collaborative AR interface, a particular challenge is to achieve shared registration. We will put more emphasis upon developing collaboration methods and interaction techniques that facilitate comfortable and intuitive cooperation among multiple users. We will also perform a formal user study to investigate the cooperative effectiveness and ease of use of the SCAPE's multi-modal interface and widgets. Finally, we will explore the possibility of extending SCAPE to remote collaboration scenarios by integrating visual and audio acquisition facilities, and evaluate the system as a tool for remote collaborative applications with our collaborative laboratories over high-speed networks.

Acknowledgements

We specially thank the following people for their stimulating discussion and collaboration on display development: Jannick P. Rolland of the ODA Lab at the University of Central Florida, Frank Biocca of the M.I.N.D lab at the Michigan State University, and Narendra Ahuja of the Computer Vision and Robotics Lab at the University of Illinois at Urbana-Champaign. We also specially acknowledge 3M Inc. for generously supplying the retro-reflective film. This paper is based on work supported by the National Science Foundation Grants IIS ITR 00-83037 and ITR 03-01874.

References

1. Agrawala, M., Beers, A. C., Fröhlich, B., Hanrahan, P., McDowall, I., Bolas, M., 1997, "The two-user responsive workbench: support for collaboration through individual views of a shared space," In *Proceedings of SIGGRAPH 1997,* ACM Press/ACM SIGGRAPH, New York, USA. 327-32.
2. Billinghurst, M., and Kato, H., 1999, "Collaborative mixed reality," in *Mixed reality: merging real and virtual worlds*, Y. Ohata and H. Tamura (Edts), Co-published by Ohmsha and Springer-Verlag, 261-284.
3. Bimber, O., Fröhlich, B., Schmalstieg, D., and Encarnacão, L. Miguel, 2001, "The virtual showcase," *IEEE Computer Graphics & Applications*, Vol. 21, No. 6, pp.48-55.
4. Brown, L., Hua, H., Gao, C., 2003, "A widget framework for augmented interaction in SCAPE," to appear in ACM 2003 Symposium on User Interface Software and Technology.
5. Buxton, W., Sellen, A., and Sheasby, M., 1997, "Interfaces for multiparty videoconferencing," In K. Finn, A. Sellen, and S. Wilber (Eds.), Video Mediated Communication. Hillsdale, New Jersey: Erlbaum, 385-400.
6. Cruz-Neira, C., Sandin, D. J., and DeFanti, T. A., 1993, "Surround-screen projection-based virtual reality: the design and implementation of the CAVE", *Proceedings of SIGGRAPH 1993,* ACM Press/ACM SIGGRAPH, New York, NY, USA, 135-142.
7. Czernuszenko, M., Pape, D., Sandin, D., DeFanti, T. A., Dawe, G. L., and Brown, M. D., 1997, "The ImmersaDesk and infinity wall projection-based virtual reality displays," *Computer Graphics*, Vol.31(2), 46-49.
8. Dias, J. M. S., Galli, R., Almeida, A. C., Belo, C. A. C., and Rebordao, J. M., 1997, "mWorld: A multiuser 3D virtual environment," *IEEE Computer Graphics & Applications*, vol.17, no.2, pp.55-65.
9. Fisher, R., 1996, "Head-mounted projection display system featuring beam splitter and method of making same," US Patent 5,572,229, November 1996.
10. Ghadisaedy, M., Adamczyk, D., Sandin, D., Kenyon, R. V., and DeFanti, T. A., 1995, "Ultrasonic calibration of a magnetic tracker in a virtual reality space," In *Proceedings of the IEEE VR 1995*, 179-88.
11. Gao, C., Hua, H., and Ahuja, N., "Easy calibration of a head-mounted projective display for augmented reality systems," Proceedings of *IEEE VR 2003*, March 22-26th, 2003, 53-60.
12. Hollan, J., and Stornetta, S., 1992, "Beyond being there," *Proc. CHI'92*, New York, ACM Press, 119-125.
13. Hua, H., Girardot, A., Gao, C., Rolland, J. P., 2000, "Engineering of head-mounted projective displays". *Applied Optics*, 39 (22), 3814-3824.

14. Hua, H., Gao, C., Biocca, F., and Rolland, J. P., 2001, "An Ultra-light and Compact Design and Implementation of Head-Mounted Projective Displays," *Proceedings of IEEE VR 2001*, March 12-17, 2001, Yokohama, Japan, 175-182.
15. Hua, H., Ha, Y., Rolland, J., 2003, "Design and assessment of ultra-light and compact projection lens," Applied Optics, 42(1), 97-107.
16. Hua, H., Gao, C., and Ahuja, N., 2003, "Calibration of a head-mounted projective display for augmented reality systems," in *Proceedings of IEEE and ACM International Symposium on Mixed and Augmented Reality 2002*, October, 2002, 176-185.
17. Hua, H., Brown, L., Gao, C., and Ahuja, N., 2003, "A new collaborative infrastructure: SCAPE," *Proceedings of IEEE VR 2003*, March 22-26th, 2003, 171-179.
18. Ichikawa, Y., Okada, K., Jeong, G., Tanaka S. and Matsushita Y., 1995, "MAJIC videoconferencing system: experiments, evaluation, and improvement," In Proceedings of ECSCW'95, 279-292.
19. Inami, M., Kawakami, N., Sekiguchi, D., Yanagida, Y., Maeda, T., and Tachi, S., 2000, "Visuo-haptic display using head-mounted projector", *Proceedings IEEE VR 2000*, Los Alamitos, CA, USA, 233-40.
20. Ishii, H., Kobayashi, M., and Grudin, J., 1992, "Integration of inter-personal space and shared workspace: Clearboard design and experiments," *Proc. CSCW'92*, 33-42.
21. Ishii, H., Kobayashi, M., and Arita, K., 1994, "Iterative design of seamless collaboration media," *Communications of the ACM,* vol 37 no. 8, p 83-97.
22. Kawakami, N., Inami, M., Sekiguchi, D., Yangagida, Y., Maeda, T., and Tachi, S., 1999, "Object-oriented displays: a new type of display systems—from immersive display to object-oriented displays", *IEEE SMC'99 Conference Proceedings*, 1999 IEEE International Conference on Systems, Man, and Cybernetics, Vol.5. Piscataway, NJ, USA. 1066-9.
23. Kijima, R., and Ojika, T., 1997, "Transition between virtual environment and workstation environment with projective head-mounted display", *Proceedings of IEEE VR 1997,* Los Alamitos, CA, USA, 130-137.
24. Kitamura, Y., Konishi, T., Yamamoto, S., Kishino, F., 2001, "Interactive stereoscopic display for three or more users," In *Proceedings of SIGGRAPH 2001,* ACM Press/ACM SIGGRAPH, New York, USA. 231-240.
25. Kröger, W., and Fröhlich, B., 1994, "The responsive workbench," *IEEE Computer Graphics and Applications*, 12-15.
26. Lehner, V. D. and DeFanti T. A., 1997, "Distributed virtual reality: supporting remote collaboration in vehicle design," *IEEE Computer Graphics & Applications,* vol.17, no.2, 13-17.
27. Macedonia, M. R., and Noll, S., 1997, " A transatlantic research and development environment," *IEEE Computer Graphics and Applications*, Vol. 17, No. 2, 76-82.
28. Parsons, J., and Rolland, J. P., 1998, "A non-intrusive display technique for providing real-time data within a surgeons critical area of interest," *Proceedings of Medicine Meets Virtual Reality98*, 246-251.
29. Rekimoto, J., 1996, "Transvision: a hand-held augmented reality system for collaborative design," In Proceedings of Virtual Systems and Multimedia (VSMM'96), Gifu, Japan, 18-20.
30. Stoakley, R., Conway, M. J., and Paucsh, R, 1995. Virtual Reality on a WIM: Interactive Worlds in Miniature. *Proc. of Human Factors in Computing Systems (CHI)*, ACM, Denver, CO, USA, 265-272.
31. Szalavari, Z., Schmalstieg, D., Fuhrmann, A., and Gervautz, M., 1998, "Studierstube: an environment for collaboration in augmented reality," *Virtual Reality, vol.3, no.1, 37-48.*