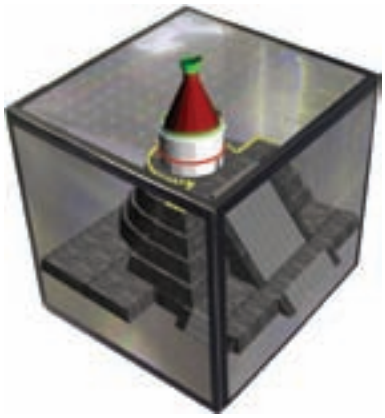


Merging the Worlds of Atoms and Bits:

Augmented Virtual Environments

Hong Hua

Interactive 3D visualization has found myriad applications in flight simulation, scientific visualization, training and education. This article reviews recent advancements in creating interactive augmented virtual environments that integrate digital information into the fabric of the physical world and enhance our sensory perceptions.

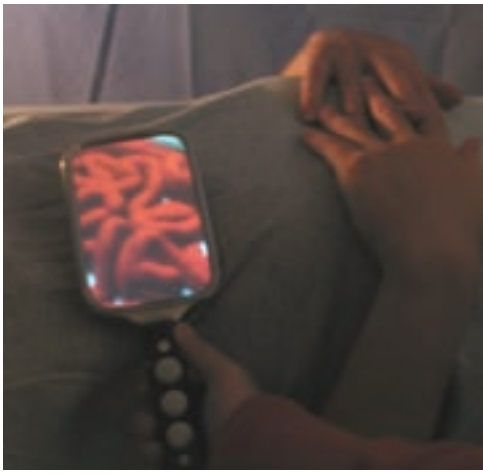
A physician is evaluating a patient lying on an exam table. To aid in her diagnosis, the doctor superimposes a 3D image of the patient's stomach and intestines in life-sized scale onto his ventral side. The doctor, wearing a goggle-like 3D display, may view the dataset from different perspectives by changing her own viewpoint, or may instruct the patient to move his body to gain a better perspective.

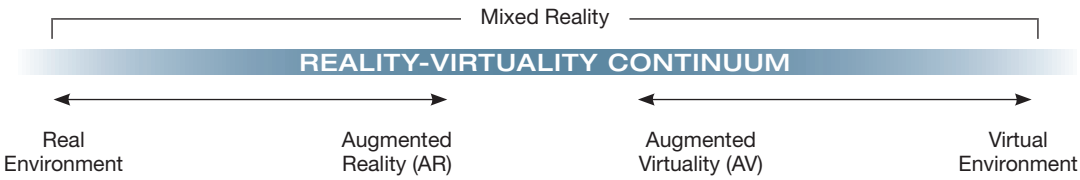
If the physician wishes to analyze a feature of the patient's internal physiology in more detail, she can grab and magnify the 3D image of this feature for viewing through an immersive room display, and even "fly" into the specified region by walking about the room. The medical images and other related digital information become an integrative part of the physical environment, and the doctor is able to freely navigate through the merged worlds of atoms and bits.

1047-6938/06/10/0026/6-$15.00 ©OSA

A proof-of-concept demonstration for medical visualization.

Images courtesy of Hong Hua





Although this scenario may seem farfetched, human-computer interfaces such as these—known as augmented virtual environments—will have a real and profound effect on the way in which we perceive and interact with digital information.

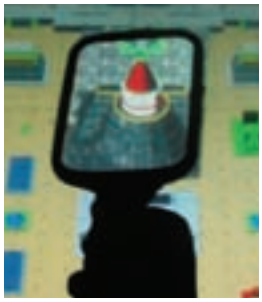
Introduction to augmented virtual environments

Researchers have been developing technologies to support interactive 3D visualization for more than three decades. The first graphics-driven 3D display prototype was demonstrated in the 1960s by Ivan Sutherland, a computer graphics pioneer. The term "virtual reality" was first used in the early 1980s by computer scientist Jaron Lanier. It is broadly defined as a high-end human-computer interface that presents real-time 3D simulation and interaction through multiple sensorial modalities, giving users the feeling of being immersed in a computer-generated world.

In the early 1990s, the term "augmented reality" was coined by researchers at Boeing who were developing an experimental system to help workers assemble wiring harnesses. It differs from virtual reality in that it seeks to supplement, rather than replace, users' perceptions of their real environment with computer-generated simulations.

Over the past decade, research and development has bloomed in the areas of virtual and augmented reality. Many enabling technologies such as 3D displays, motion trackers, 3D graphics hardware and interaction techniques have emerged, as have many branches of related research. The relationships of these evolving areas can be elucidated by assessing the reality-virtuality (RV) continuum, which was proposed by Milgram and Kishino to characterize various visual displays. As illustrated in the figure above, the "real" world and entirely computer-generated virtual environments represent opposite poles of the spectrum. Excluding the poles, the segment toward the center is known as mixed reality.

A mixed reality system requires that the overlaid virtual information be aligned with the user's real-world senses—a process known as registration.



Augmented reality differs from virtual reality in that it seeks to supplement, rather than replace, users' perceptions of their real environment with computer-generated simulations.

The left portion of the mixed reality segment represents augmented reality, where digital information is overlaid on a real-world scene, and the right is referred to as augmented virtuality, in which a virtual environment is selectively enhanced with real-world data.

The major difference between mixed and virtual reality is that the former uses the physical world as its reference and scale to render virtual information correlated with its real counterpart, while the latter may have an arbitrary reference point and scale without correlation with the physical surroundings. Thus, a mixed reality system requires that the overlaid virtual information be aligned with the user's real-world senses—a process known as registration.

This classical view of the RV continuum emphasizes the visual aspect of the transition between the physical and digital worlds. For instance, one of the very first visions for mixed reality interfaces is aimed at data visualization, where the goal is to correctly register virtual information relative to the real-world objects so that the viewer perceives a synthetic view beyond the physical reality. The concept of the RV continuum can be readily extended to other sensory forms such as sound or touch.

Interactivity is also an essential component of a 3D environment; it allows users to manipulate virtual objects, perform system control or add textual annotations. Many 3D interface techniques have been explored. Motion trackers,

joysticks, 3D mice and datagloves have been the dominant input devices.

The interactive techniques using these devices differ quite a bit from the way people naturally interact with the real world. For example, in reality, a user directly manipulates physical objects with his or her hands. However, in a virtual environment, selecting an object is usually more complicated: It requires a means to indicate an object for selection, a way to confirm it and a mechanism through which to provide feedback during the interaction.

Thus, when users deal in the physical and virtual worlds, they are forced to switch between operation modes, resulting in interaction discontinuity. To address this, researchers hope to develop tangible interaction techniques that allow seamless integration of both realms—for example, a method that allows a user to directly select, examine and share a virtual object with bare hands and familiar gestures.

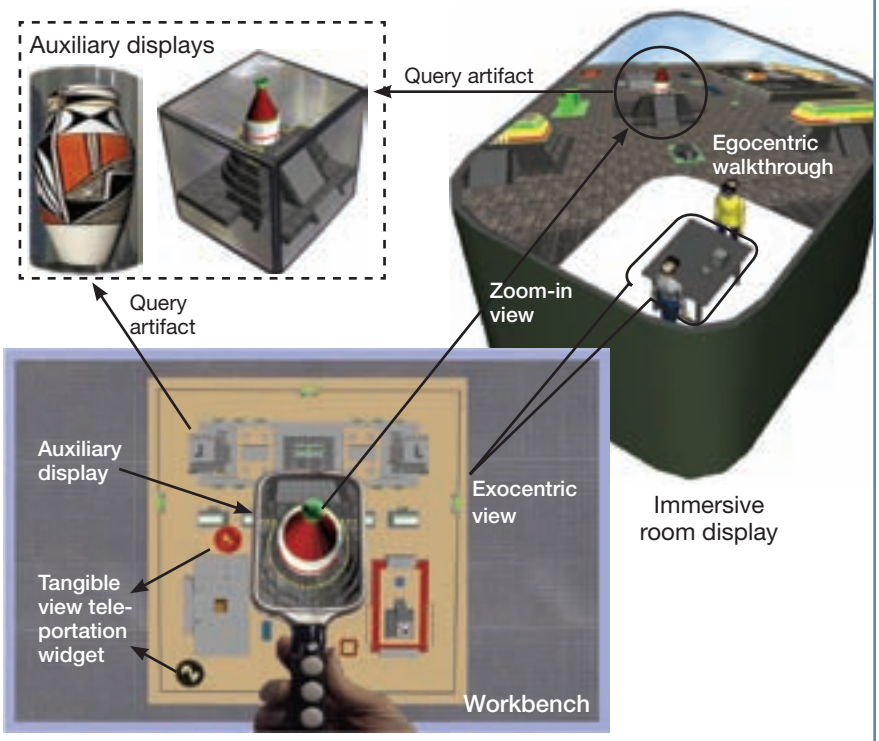
An augmented virtual environment is an interface that aims to integrate digital information with the physical world, from the perspectives of not only how we perceive but also how we interact with the merged worlds of atoms and bits. From the display perspective, an augmented virtual environment is one in which users experience a smooth transition between the physical and virtual worlds and have the capability of traversing arbitrary levels of immersion into the digital realm. From the interaction perspective, such an environment emphasizes interaction with both worlds in intuitive, unified manners such as direct manipulation and physics-based interaction metaphors.

3D display technology bridges the paradigms

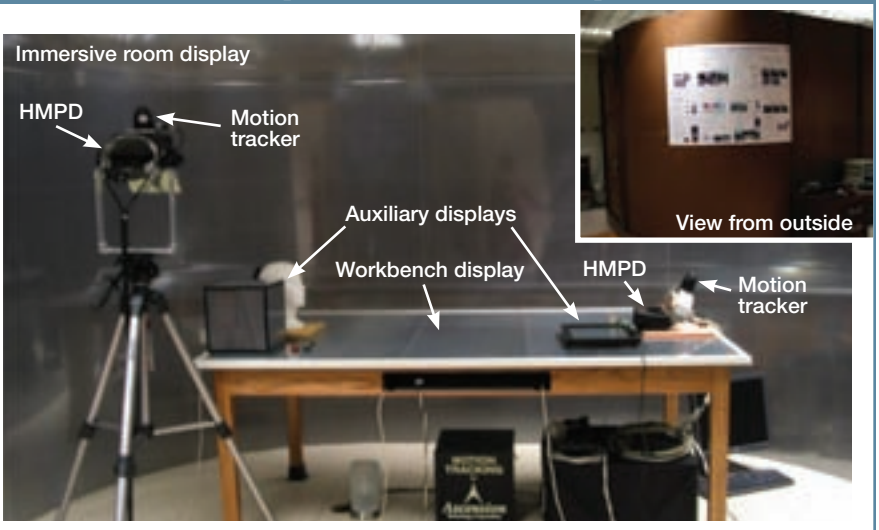
Three-dimensional displays allow observers to perceive depth effects. They have been one of the enabling technologies for creating interactive 3D environments. (An excellent review of the principles and history of 3D display techniques can be found in an article by Keigo Iizuka in the July/August OPN.) Applying the

Stereoscopic Collaboration in Augmented & Projective Environments (SCAPE)

[Conceptual simulation]



[Prototype implementation]



Immersive walkthrough: sample view through the surrounding wall display.

Exocentric overview: sample view through the workbench.

principles of 3D imaging, scientists have developed many 3D display systems. For instance, a well-known but expensive approach is the projection-based spatially immersive displays such as CAVE Automated Virtual Environments (CAVEs). These spatially immersive displays can accommodate a number of users, who wear compact liquid crystal shutter glasses or polarizers to view stereoscopic images in a shared space.

In these displays, however, it is usually difficult to present a combination of virtual and real information, which is called creating "see-through" capability, to maintain a consistent registration between the virtual and physical environments from multiple perspectives, and to support the use of a natural interaction mechanism. For instance, when two users attempt to point to the same virtual object in their views, their fingers likely do not meet in physical space.

See-through head-mounted displays (HMDs) have been the dominant devices used for augmented reality applications. These displays fall into two categories: video see-through and optical see-though. In a video see-through HMD, the display blocks the direct view of the physical world and one or two miniature video cameras mounted on the top of the headgear are used to capture the real-world view—which is then electronically fused with the computer-generated virtual environment. In an optical see-through HMD, the direct view of the real world is maintained and the computer-generated virtual scene is optically superimposed onto the real scene via a beamsplitter.

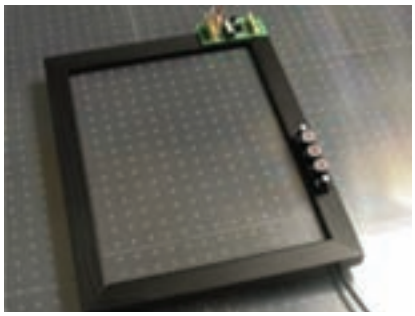
The optical see-through approach allows a user to see the real world with full resolution and is less intrusive into the user's view of the real scene than the video see-through approach. Therefore, it is the preferred method for tasks where hand-eye coordination or non-blocked real-world view is critical. Both types of displays have been applied in various augmented applications, from medical training to entertainment.

Head-mounted projective display (HMPD) technology, pioneered by Fisher (U.S. Patent 5,572,229), has recently matured as an alternative approach to 3D visualization systems. The technology deviates from the conventional approaches to HMD designs by replacing eyepiece- or microscope-type optics in a typical HMD design with projection optics, which are then combined with a retroreflective screen, as opposed to the diffusing screen in a conventional projection system. The unique combination of projection and retroreflection enables stereoscopic capability.

In the late 1990s, several groups of researchers in Japan and the United States conducted preliminary research exploring the HMPD concept for various visualization applications. In 1999, when I spent 10 months in Dr. Rolland's lab, she and I studied the imaging properties of retroreflective materials and their effects on image quality. We investigated the engineering challenges of designing miniature projection optics. These efforts have led to the success of custom-designed compact prototypes that have been used in both of our labs for developing augmented applications.

While investigators have made much progress in 3D display technology, it is still a challenge for them to design wide field-of-view, compact and non-intrusive optical see-through HMDs. The hope is that new display concepts will emerge that can fully enable the vision captured by an augmented virtual environment and bridge the paradigms along the RV continuum.

SCAPE
[Physical props and interfaces]



Prototypes of devices that integrate interaction functions: (top) a framed window attached to a workbench display; (center) a hand-held device with a mirror-like prop; (bottom) a hand-held device with a volume prop.

Multi-scale collaborative augmented virtual environments

Most research into immersive virtual environments has focused on the development of visualization tools and interaction techniques for single-scale, single-user-interaction environments. With the ever-increasing scope of data available for exploration, however, new work may aim to simultaneously present a complex dataset in several levels of scale from different view perspectives. There has been great interest in developing multi-scale virtual environments and techniques

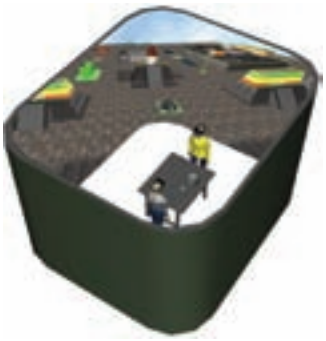
to navigate such environments. For instance, the "Worlds in Miniature" (WIM) metaphor by Stoakley and his colleagues is a well-received technique for creating a dual-scale, dual-perspective virtual environment, where the WIM is essentially a mini 3D map providing an overview of an expansive virtual world.

Developing multi-scale augmented virtual environments is challenging due to its requirement for maintaining seamless integration and correct registration between the different views and scales referenced to the physical world and achieving smooth navigation through the environments. Using traditional see-through HMDs, several groups of researchers have explored the WIM metaphor in the design of multi-scale augmented reality interfaces, such as the Magicbook project by Mark Billinghurst and his colleagues. By exploring several unique properties of HMPD technology, my group has developed a multi-scale collaborative infrastructure, referred to as SCAPE (Stereoscopic Collaboration in Augmented and Projective Environments).

A schematic simulation of the SCAPE conceptual design and a prototype implementation are shown in the images on p. 29. The retroreflective screens that HMPDs rely on to view projected images are not necessarily in flat shapes or placed at fixed distances from the projection optics, as required by the conventional projection-based displays such as CAVEs. Thus, theoretically they can be tailored into arbitrary shapes and strategic placements without warping or blurring perceived images. In practice, however, artifacts of retroreflective materials can cause image quality degradation.

The existing SCAPE environment consists of two primary display surfaces: a workbench and a curved immersive room display, both of which are coated with retroreflective materials. A computer-generated low-detailed microscene is registered with the workbench and physical objects placed on it, while a correlated high-detailed, life-size, immersive walk-through, or macroscene, is visualized through the surrounding room display.

With the ever-increasing scope of data available for exploration, new work may aim to simultaneously present a complex dataset in several levels of scale from different view perspectives.



A computer-generated low-detailed microscene is registered with the workbench and physical objects placed on it, while a correlated high-detailed, life-size, immersive walk-through, or macroscene, is visualized through the surrounding room display.

Thus, the primary displays visualize a dataset simultaneously in dual scales and perspectives in a hybrid form of augmented and virtual reality—an exocentric overview at a low level of detail on the workbench display in augmented reality mode and a life-sized egocentric walkthrough view in high detail on the room display in a virtual environment.

Photographs of several testbed applications captured in the SCAPE environment are shown on p. 27 and p. 29. Such dual-mode visualization can be further extended into multiple hierarchical levels of realization, which can take into account specific application requirements. For example, in a planetary science application, we have been exploring the design of primary displays that allow a scientist to simultaneously view the entire planet with only recognizable landmarks, a regional map of a selected area with moderate terrain detail, or a zoomed-in walkthrough of a locale.

Often, the discrete views provided by the primary displays alone are insufficient—for example, in the immersive view, the dataset appears overwhelmingly detailed or zoomed in, whereas on the workbench, the same visualization may appear at inadequately low resolution or look too zoomed-out to distinguish notable characteristics.

In such instances, more continuous scale control or intermediary levels of detail and scale are needed to fine-tune the visualization for a given task. The challenge is creating display methods for intermediary scales and effectively managing the complexity resulting from the display varieties.

In the SCAPE environment, we attempt to introduce various auxiliary displays that can provide alternative perspectives and zoomable scale controls of a particular point of interest on the primary displays. The auxiliary displays can function as dual-purpose devices: display and interaction. A device not only provides a surface on which to view information, but also a tangible artifact

with which users can interact and share. The tangibility of these devices makes them particularly attractive in designing a rich range of user interaction techniques.

With the increasing complexity of visualization tasks, many applications require the coordination of a team of experts. Researchers are working to support both individual and group interaction with a 3D augmented virtual environment for multiple co-located users.

Group work in the real world often involves the interleaving of individual and group efforts, frequent information exchange and a considerable amount of both explicit and tacit communication between collaborators. The transition between shared and individual activities is smooth and often invisible. In a 3D visualization system, however, the support of multiple-user interactions and the maintenance of smooth transition and communication impose great challenges in both display and user interaction techniques.

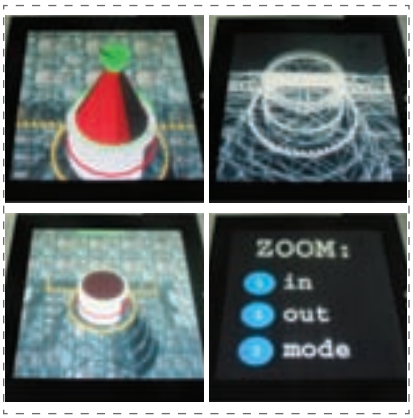
For instance, it is necessary to develop techniques to support different styles of collaborative work, such as the symmetrical, leader-privileged and slave modes. In the symmetrical mode, each user has individual control of his or her view across the entire space and the group works together in parallel. In the leader-privileged mode, a group leader controls a global anchor to select a region of interest, but all users can explore the region freely from their individual perspectives. The slave mode is similar to the collaboration method of most CAVE-like environments, where a group leader not only has global control privilege but also forces other users to stand in his or her shoes.

It is also necessary to develop techniques that support simultaneous user interactions, facilitate activity awareness and provide a convenient means for sharing views and information. In SCAPE, we implemented a flexible transformation hierarchy that is potentially capable of supporting several styles of collaboration and a variety of interaction techniques.

[Exploring tangible interaction techniques in SCAPE]

Augmented views of planar and cube-shape zoom lenses in a 3D augmented virtual environment.

Lens operators



Sample views of an embedded planar lens with multi-functions (i.e., zoom, clip and rendering-mode switch).

Interacting with augmented virtual environments

Everyone is familiar with the interaction techniques that are typically used with desktop computers. For example, the mouse and keyboard are recognized as default input devices. Moreover, in the 2D user interface, pull-down menus and drag-and-drop interaction techniques are universally accepted. However, these familiar interface components are often not appropriate for an interactive 3D environment.

The concept of a tangible user interface (TUI), also known as tangible computing, is a promising interaction technique for a 3D environment. In this approach, a user interacts with digital information using graspable physical handles called phicons (physical icons).

In parallel with Milgram's notion of the RV continuum, TUI emphasizes the vision of seamlessly blending the worlds of atoms and bits. Perhaps one of the most appealing properties of TUI is that a user can interact with both the virtual and real worlds by direct manipulation—which is something people can accomplish intuitively and with ease. For instance, 2D or 3D virtual objects can be accurately registered with their physical counterparts to create integrated tools. Users can thus select, examine and share a virtual object by directly manipulating its physical handler.

The tangibility of 3D interaction can be promoted in subtle manners. For instance, the physical shape of a device can be selected to suggest its intended functions, or its setting can be made to indicate how to use it. The direct manipulation of the physical device can be mapped to meaningful operation on the virtual dataset. This area of research has not been well explored, but is now drawing more attention in the human-computer interaction community.

In the SCAPE environment, several auxiliary display widgets have been explored to facilitate the development of tangible interaction methods. These displays are treated as dual-purpose widgets, which not only expand the

1047-6938/06/10/0032/6-$15.00 ©OSA

visualization capability of the primary displays but act as tangible interfaces for manipulating digital information. One example of the widgets is a device that integrates multiple interaction functions into a physical prop, referred to as a tangible magic lens (TML).

A TML-based device employs a lens analogy. It consists of a method of selecting a point of interest in a context display (e.g., the workbench or wall displays), an inset window or a volume for displaying an alternative view of the point of interest, a stack of semantic lens operators that defines how the point of interest is visualized, and an interface technique that allows a user to interact with, and toggle among, the lens functions.

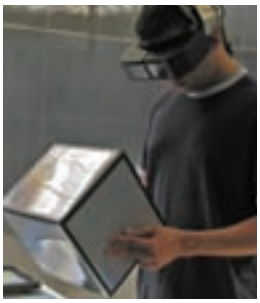
Examples of the operators include a "zooming" lens to control the scale of the 3D contents displayed in the inset window, a rendering mode operator that cycles through different drawing methods (e.g., textured vs. wireframe) or a filtering function that tunes the level of information complexity (e.g., Boolean addition and subtraction). These operators can be "stacked" together to create accumulating effects from a set of basic operators.

Three prototypes of TML-enabled devices are presented on p. 30. The top device is a framed window attached to a workbench display. The display uses a decoupled method for selecting a point of interest, allowing a large inset window without ergonomic concerns. The center device is a hand-held device with a mirror-like prop, which allows a user to specify a point of interest on the fly by holding the lens directly over a desired area on the workbench.

The bottom image is a hand-held device with a volume prop, which allows a user to select a 3D object from the immersive display environment, scale it down to fit into the volume, and examine the object in the same way a user would do with a physical artifact.

The images on the facing page demonstrate a set of sample views of these widgets with differing functions and combinations of lens operators.

In a 3D visualization system, the support of multiple-user interactions and the maintenance of smooth transition and communication impose great challenges in both display and user interaction techniques.



The concept of a tangible user interface is a promising interaction technique for a 3D environment. In this approach, a user interacts with digital information using graspable physical handles called phicons (physical icons).

The widgets essentially support tangible navigation, scale manipulation, object examination and object sharing in a 3D augmented virtual environment.

Conclusions and future directions

Many technical challenges persist in creating an augmented virtual environment that offers users a rich range of visualization and interaction capabilities for addressing complex application tasks. These challenges not only stem from the lack of non-intrusive 3D display technologies, but also a dearth of adequate methods for interacting with the virtual world in a tangible manner. Future efforts should be made to develop less-intrusive display technologies and novel visualization and interaction methods to support applications that involve heterogeneous data sets.

[**Hong Hua** (hhua@optics.arizona.edu) is with the College of Optical Sciences, The University of Arizona, Tucson, Ariz.]

OSA Member

[References and Resources]

>> C. Cruz-Neira et al. "Surround-screen projection-based virtual reality: the design and implementation of the CAVE," Proc. of SIGGRAPH 1993, 135-42 (1993).

>> P. Milgram and F. Kishino. "A taxonomy of mixed reality visual display," IEICE Transactions on Information Systems, special issue on "Networked Reality," E77-D, **12** (1994).

>> H. Ishii and B. Ullmer. "Tangible bits: towards seamless interfaces between people, bits and atoms," ACM Proc. Human Factors in Computing (CHI'97), 234-41 (1997).

>> H. Hua et al. "Engineering of head-mounted projective displays," Appl. Opt., **39**(22), 3814-24 (2000).

>> R. Azuma et al. "Recent advances in augmented reality," IEEE Computer Graphics and Applications, **21**(6), 34-47 (2001).

>> L. Davis et al. "Alice's adventures in Wonderland: A unique technology enabling a continuum of virtual environment experiences," Computer Graphics and Applications, **23**(2), 10-12 (2003).

>> H. Hua et al. "System and interface framework for SCAPE as a collaborative infrastructure," Presence: Teleoperators and Virtual Environments, **13**(2), 234-50 (2004).

>> L. Brown and H. Hua. "Magic lenses for augmented virtual environments," IEEE Computer Graphics and Applications, **26**(4), 64-73 (2006).