Understanding and Constructing Shared Spaces with Mixed-Reality Boundaries

STEVE BENFORD, CHRIS GREENHALGH, GAIL REYNARD, CHRIS BROWN, and BORIANA KOLEVA
The University of Nottingham

We propose an approach to creating shared mixed realities based on the construction of transparent boundaries between real and virtual spaces. First, we introduce a taxonomy that classifies current approaches to shared spaces according to the three dimensions of transportation, artificiality, and spatiality. Second, we discuss our experience of staging a poetry performance simultaneously within real and virtual theaters. This demonstrates the complexities involved in establishing social interaction between real and virtual spaces and motivates the development of a systematic approach to mixing realities. Third, we introduce and demonstrate the technique of mixed-reality boundaries as a way of joining real and virtual spaces together in order to address some of these problems.

Categories and Subject Descriptors: H.4.3 [**Information Systems Applications**]: Communications Applications—*computer conferencing and teleconferencing*; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems—*artificial realities*; H.5.3 [**Information Interfaces and Presentation**]: Group and Organization Interfaces—*theory and models*

General Terms: Human Factors, Theory

Additional Key Words and Phrases: Augmented reality, collaborative virtual environments, CSCW, media-spaces, mixed reality, shared spaces, telepresence, video, virtual reality

1. INTRODUCTION—SHARED-SPACE TECHNOLOGIES

Shared-space technologies aim to create distributed electronic environments where participants can exploit spatial properties such as containment and movement in order to manage their communication. This approach in its various forms has been motivated by a range of issues that are seen as being important to cooperative work. These include

—creating a persistent context for on-going activity;

Authors' address: Department of Computer Science, The University of Nottingham, Nottingham, NG7 2RD, UK; email: {sdb; cmg; gtr; ccb; bnk}@cs.nott.ac.uk.

ACM Transactions on Computer-Human Interaction, Vol. 5, No. 3, September 1998, Pages 185–223.

—enabling peripheral as well as focused awareness of the activities of others, a critical issue noted by Heath and Luff [1991] and Hughes et al. [1992] in their studies of real-world cooperative work;

—facilitating chance encounters, a goal of early media-spaces such as *Cruiser* [Root 1988]; and

—promoting usability through the use of a spatial metaphor (e.g., the frequent use of the virtual office metaphor as in Cook et al. [1991]).

In a more general sense, spatial approaches to Computer Supported Cooperative Work (CSCW) might be seen as a shift of focus toward supporting the context within which work takes place, rather than the process of the work itself. Thus, the spatial approach contrasts with other more process-oriented approaches to CSCW, such as that of workflow systems.

A review of the literature suggests to us that spatial approaches to CSCW can be grouped into five general categories: media-spaces, spatial video-conferencing, collaborative virtual environments, telepresence systems, and collaborative augmented environments.

1.1 Media-Spaces

Media-spaces involve the enhancement of existing workspaces (typically offices) with integrated audio/video communication facilities as a basis for providing a range of general communication services [Bly 1993]. Among the best-known examples of media-spaces are the *Cruiser* [Root 1988] and *RAVE* [Gaver et al. 1992] systems. Media-spaces focus on support for social browsing, peripheral awareness, and the establishment and maintenance of long-term working relationships among physically separated people. Typical services include the ability to glance into other people's offices, to establish longer-term office share relationships via open connections, or to gain a general sense of people's likely availability through a *Portholes* style interface [Dourish and Bly 1992]. The *Multiple Target Video* (*MTV*) media-space explored the use of multiple cameras in a physical location to provide different views of the activities of others. These included face-to-face views, in-context views of the participant in relation to their workspace, and desk views for the sharing of documents [Gaver et al. 1993].

Several authors have conducted theoretical or experimental evaluations of media-space technology, focusing on problems with video communications such as limited field of view and the lack of dynamic navigation [Gaver 1992]. Discussions of the MTV system identified problems in working with fragmented views of a remote physical space. These included difficulties with switching between different camera views and making reference to objects within a remote physical environment [Gaver et al. 1993].

1.2 Spatial Video-Conferencing

Video-conferencing involves the use of combined video and audio communications to support focused meetings. This can be contrasted to the media-space objectives of supporting peripheral awareness and long-term working relationships. Video-conferencing may be experienced in both desktop mode (i.e., using multimedia workstations) and through dedicated links between public meeting rooms. Video-conferencing systems may include additional communication tools such as shared-document editors.

A problem with traditional video-conferencing systems is that they do not easily support forms of spatial referencing, such as gaze direction, whereby participants can infer who is attending to whom at any moment in time from their representations. Gaze direction has been identified as a key element of conversation management [Sacks et al. 1994], and other research has indicated the general importance of understanding the viewpoints of others when engaged in collaborative work (e.g., the experiments of Shu and Flowers on viewpoint representation in collaborative 3D design tasks [Shu and Flowers 1992]). However, if one looks at the camera in a video conference, one appears to gaze at all of the participants simultaneously; there is no way for other participants to distinguish at whom one is gazing.

Several researchers have recognized this problem and have attempted to introduce support for gaze direction into video-conferencing. This results in what we term "spatial video-conferencing" systems. For example, the *Hydra* system utilized an arrangement of miniature televisions on a table top to give each participant a consistent representation of gaze [Sellen and Buxton 1992], and the *MAJIC* system used video projection techniques to achieve the same effect for three participant meetings [Ichikawa et al. 1995]. As a slightly different example, the *Clearboard* system demonstrated the integration of two video streams and a shared drawing surface into a spatially consistent environment in order to support two-person design meetings [Ishii and Kobayashi 1992]. However, it should be noted that a generalized solution which works well for larger numbers of participants or meetings where participants dynamically join and leave has yet to emerge; the above systems would require considerable physical hardware reconfiguration if a new participant were to suddenly arrive.

1.3 Collaborative Virtual Environments (CVEs)

CVEs involve the use of networked virtual reality systems to support group work [Benford et al. 1994]. The key concept behind CVEs is that of shared virtual worlds: computer-generated spaces whose occupants are represented to one another in graphical form. Each occupant can control his or her own viewpoint and can interact with others and with representations of data and computer programs within the virtual space. In a CVE the shared space defines a consistent and common spatial frame of reference. In other words, there is a mutually available coordinate system through which the relative positions and orientations of different objects can be understood.

This is combined with support for independent viewpoints that are represented through avatars with the intention that it becomes possible to infer where someone is attending and what that person is seeing from his or her representation. Note that making such an inference is not the same as actually seeing what they are seeing. CVEs also aim to provide an integrated, explicit, and persistent context for cooperation that combines both the participants and their information into a common display space. This is in contrast to multimedia systems that typically display communication and data in separate windows. Furthermore, the possibility of including a wide variety of data representations creates the potential to support a broad range of cooperative applications such as training, visualization, simulation, design, and entertainment.

Representative examples of CVEs include the DIVE system from the Swedish Institute of Computer Science [Fahlén et al. 1993], *Collaborative Workspace* from NTT [Takemura and Kishino 1992], our own *MASSIVE* system [Greenhalgh and Benford 1995], and large-scale military simulation systems such as *NPSNET* [Zyda et al. 1993]. Recently, a number of commercial services have emerged that deliver social virtual worlds over the Internet (e.g., *Alphaworlds* and *On-Live Traveller*).

A commonly held assumption behind CVEs has been that participants somehow leave the physical world behind and enter into a virtual world in order to communicate with one another. In contrast, recent experiments with our own *MASSIVE* system [Greenhalgh and Benford 1995] suggest, that in order to make sense of a participant's actions in a virtual world, other participants require an understanding of actions and events within that participant's local physical environment. Bowers et al. present a conversation analysis of transcripts of meetings in MASSIVE in which they argue that the perceived trustworthiness of an embodiment can be influenced by real-world events such as users leaving their embodiments unoccupied when attending to real-world interactions or several users sharing a single embodiment [Bowers et al. 1996]. This issue of providing knowledge of the physical world within the virtual world becomes even more pertinent when the participants wish to actually manipulate and refer to physical objects such as paper documents as part of their cooperation.

1.4 Telepresence Systems

The concept of telepresence involves allowing participants to experience a remote physical space through computer and communications technologies. This may include the ability for the remote participant to view the space, to navigate the space, and even to interact with objects in the space. It should be noted that our use of the term telepresence to describe access to remote physical spaces is more restrictive than some uses of the term that refer to remote access to any kind of space, physical or virtual. Telepresence applications typically involve the creation of a physical proxy of the remote person in the form of a robot which has cameras attached to it and which may be able to move through the physical environment to varying degrees [Stone 1991]. In some cases the remote user may actually experience the physical space through the same kinds of immersive technology as may be used in collaborative virtual environments, except that in this case live video is displayed in the head-mounted display instead of 3D graphics. Telepresence is a field of research in its own right with applications focusing on areas such as control of remote robots in hazardous or inaccessible environments and navigation through remote regions using mobile robots. However, telepresence applications are now beginning to be discussed in the context of CSCW in systems such as the *GestureCam* [Kuzuoka et al. 1995] that explores the notion of remote surrogates in cooperative work.

1.5 Collaborative Augmented Environments

Recently, researchers have begun to explore the possibilities of shared augmented environments. The technology of augmented reality involves the overlaying of graphical objects onto a real word scene with some degree of dynamic registration between the two. This may be achieved through a variety of technologies, typical examples of which include overlaying graphics onto conventional video displays and the use of see-through head-mounted displays (see Milgram and Kishino [1994] for a comprehensive classification of approaches). Early experiments with collaborative augmented reality include the *shared-space system* [Billinghurst et al. 1996], in which users share virtual objects across a physical table top and *Studierstube* [Schmalstieg et al. 1996], in which virtual objects are also displayed in a physical space between multiple users. Both of these systems utilize see-through head-mounted displays coupled to Polhemus motion trackers.

An alternative approach to the augmentation of physical environments with digital information is given by Ishii and Ullmer's notion of tangible bits [Ishii and Ulmer 1997]. They explore how ambient display media such as sound, light, and airflow can provide peripheral awareness of background information and how this can be coupled with graspable objects and interactive surfaces in order to extend this to interaction with foreground information. The long-term goal of this approach is the natural integration of digital and physical information.

2. CLASSIFYING SHARED-SPACE TECHNOLOGIES ACCORDING TO TRANSPORTATION, ARTIFICIALITY, AND SPATIALITY

We now introduce the three dimensions of transportation, artificiality, and spatiality as a means of classifying shared-space technologies. There are three motivations behind this classification. First it allows us to explore the design trade-offs involved in applying these technologies to support different cooperative activities, e.g., in determining the costs and benefits of supporting different spatial properties such as containment, topology, movement, and shared coordinate systems. Second, such a classification may identify gaps in the technology where new approaches might be developed. In other words, it can provide the inspiration for new avenues of research. Third, producing a clean classification that is simple and inclusive and that clearly separates the different technologies suggests that we have managed to abstract out the key principles that define them and have understood the primary distinctions between them.

Later on, we shall reflect on the relationship between our classification and that of Milgram and Kishino in their original development of the mixed-reality approach [Milgram and Kishino 1994].

2.1 Transportation

The dimension of transportation concerns the extent to which a group of participants and objects leave behind their local space and enter into some new remote space in order to meet with others, versus the extent to which they remain in their local space and the remote participants and objects are brought to them. It therefore characterizes the essential difference between the concepts of local and remote.

One extreme of this dimension involves wholly thinking in terms of the participants' local physical environment. This would be the case in a physical face-to-face meeting and is the dominant tendency in augmented reality. At the other extreme is total involvement with a remote environment of some kind. Immersive virtual reality and immersive telepresence applications are exemplars of this extreme. At intermediate levels of transportation we find split levels of involvement, where participants attend to aspects of both their immediate physical environment and the remote environment. As one moves toward the transportationless extreme, the remote environment becomes less significant and impinges less on the immediate context. As one moves toward the totally transported extreme, the immediate environment becomes less significant; for example, it may seem that less of the local environment is being drawn into the remote environment.

The nature of the interface technology used has a significant effect upon transportation. For example, shared projected interfaces provide a high level of immersion. The most extreme examples are *CAVEs*, purpose-built rooms onto whose surfaces multiple synchronized views are projected from the outside [Cruz-Neira et al. 1992]. However, such displays may be less transporting than head-mounted displays because a number of physically co-located participants can share a single display and so remain aware of one another and of their shared physical environment. Desktop virtual reality interfaces typically expose the user to greater interference from local stimuli and distractions and thereby situate them more fully in their local surroundings. Indeed, early trials with the *MASSIVE* system identified the so-called "degree of presence" problem, where users who were embodied in the virtual world temporarily stepped out of their bodies due to some local distraction, causing confusion for other users who were trying to interact with them [Bowers et al. 1996; Greenhalgh and Benford 1995].

Our concept of transportation is similar to the virtual reality concept of immersion. Both are concerned with the extent to which an interface technology has been designed to introduce a participant into a new environment while at the same time excluding sensory stimuli from the local environment. However, transportation differs from immersion in two key respects. First, unlike immersion, transportation also includes the possibility of introducing remote participants and objects into the local environment that then becomes augmented rather than excluded. This is the dominant tendency in augmented reality and ubiquitous computing and may be an important starting point for designing technologies that need to be integrated with existing tools as part of the everyday working environment. Second, transportation considers how groups of participants and possibly other objects such as physical documents might be transported together. Immersion has typically focused on individual participants. Even where sharable interfaces such as projected displays have been used, the effect on and role of local objects has not been considered.

Like immersion, transportation is in principle a quantifiable property of a technology. Different technologies can be objectively located along this dimension according to the extent to which the display of the remote environment excludes the local environment and the amount of information that is projected from the local into the remote. This is not necessarily the same as the extent to which users perceive that (or behave as if) they are present in a new space. Just as virtual reality research makes a distinction between the technological concept of immersion and the psychological concept of presence [Sheridan 1992; Slater et al. 1994], so we also separate transportation from notions of presence or co-presence. Early experiments with virtual reality technology have suggested that while the degree of presence experienced may increase with the degree of immersion, other factors also make a profound contribution. These include whether users can see their own virtual body images [Slater et al. 1994] or the use of physical walking as a means of moving through a virtual environment [Slater et al. 1995]. The same distinction can be seen in shared-space technologies. The MAJIC system provided a highly transporting interface through a combination of large-screen displays and background substitution; however, users still remained aware that they were sitting in their own physical offices [Ichikawa et al. 1995].

2.2 Artificiality

The dimension of artificiality concerns the extent to which a space is either synthetic or is based on the physical world. This spans the extremes from wholly synthetic to wholly physical environments, i.e., between the total synthesis of an environment, independent of all external reality from nothing but bits in a computer to the electronically mediated delivery of a physical place, firmly based in everyday reality. A technology can be located along this dimension according to the ratio of physical to synthetic information present in its space. Video-conferencing and telepresence applications are typical of the physical extreme, as their information is drawn directly from the physical world. The use of CVEs for abstract data visualization or computer art exemplifies the synthetic extreme.

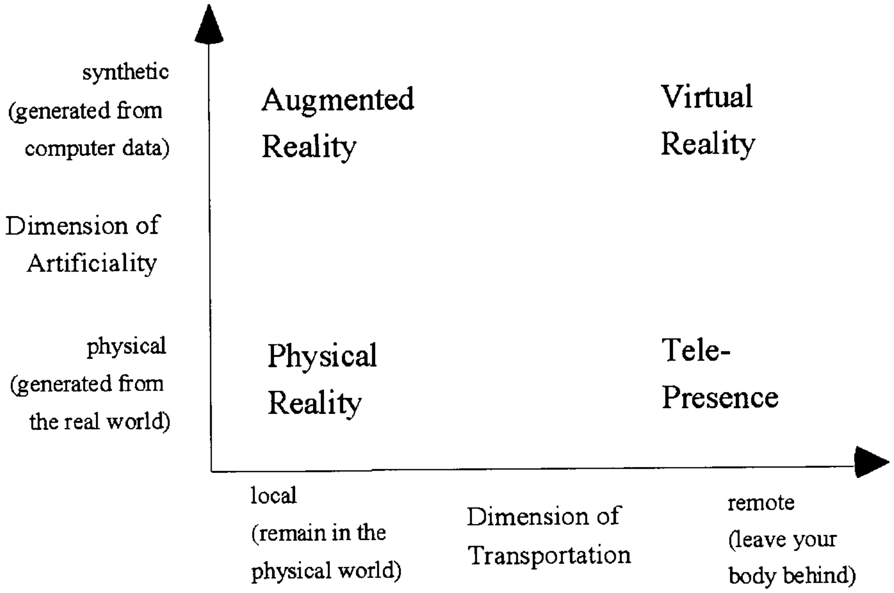

Fig.1. Broad classification of shared spaces according to transportation and artificiality.

Physical information may include images of people's faces and bodies, images of paper documents, views of physical spaces, and telemetry data drawn from sensors in the real world. Synthetic information may include 3D geometry, synthesized sounds, and external digital documents that may be brought into a synthetic space. For example, *Clearboard* [Ishii and Kobayashi 1992] combines digital documents with views of physical spaces, and our own *WWW-3D* Web visualization integrates HTML documents into a CVE [Benford et al. 1997c].

Associated with this dimension is a distinction in the underlying technologies. Video technology naturally lends itself to the capture and reproduction of physical scenes, while 3D graphics naturally lends itself to the synthesis of abstract scenes. However, within this general trend, more elaborate video manipulation technology can warp video signals away from the physical, while scene analysis techniques and real-world data capture can provide the basic data for generating 3D graphical scenes that correspond directly to physical reality.

2.3 Classification According to Transportation and Artificiality

Figure 1 provides a high-level overview of the classification of shared-space approaches according to the dimensions of transportation and artificiality.

It is possible to locate four major strands of technology on this classification. Familiar meeting technologies associated with physical reality (e.g., physical tables, whiteboards, etc.) combine both the local and physical. Virtual reality combines the remote and synthetic. Telepresence combines

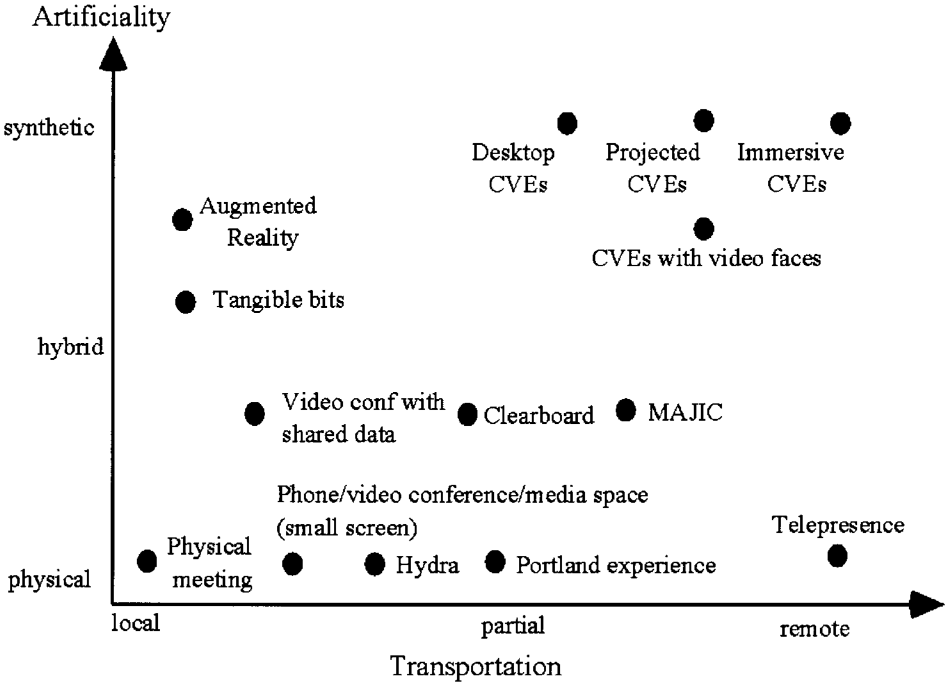

Fig. 2. Detailed classification of shared spaces according to the dimensions of transportation and artificiality.

the remote and physical. Augmented reality overlays the purely synthetic on the local environment.

Figure 2 offers a more detailed version of this classification that is populated with specific meeting technologies.

The users of a telephone conference essentially remain in the physical world, and all supplementary material is a direct rendering of remote physical events. Thus, telephone conferencing is fundamentally both local and physical. Similarly, in a traditional small-screen video conference or a media-space users remain in their local physical world and view remote video footage of physical events. The use of a small screen minimizes the display space afforded to the remote scene and thus assigns a dominant role to the local environment. Video-conferencing can be enhanced with shared editors so as to support discussion and joint editing of electronic documents and drawings [Sarin and Greif 1985]. This introduces a certain amount of synthetic information into the conference and hence involves greater artificiality than minimal video-conferencing.

Hydra [Sellen and Buxton 1992] introduces a consistent multiparty space in which gaze direction and the ordering of participants is meaningful (e.g., X will be to the left of Y for every participant). The sharing of a conference space weakens the role of the local space. However, all the information displayed is still a direct representation of the physical world. The use of large screens in video-conferencing between two sites (e.g., Olson and Bly

[1991]) may give users a stronger sense of involvement in the remote scene, implying a degree of transportation, but still retains its physical nature. The *MAJIC* system, especially with background substitution, excludes the local environment to an even greater extent, and the use of larger projected displays also strengthens the role of the remote scene when compared to the small displays used in *Hydra*. The use of synthetic backdrops introduces a degree of synthesis and therefore moves *MAJIC* away from the strictly physical end of the artificiality dimension. Clearboard [Ishii and Kobayashi 1992] creates a boundary between two physical spaces through a shared drawing surface. However, it also allows for digital information to be positioned on the surface between these spaces and so introduces a degree of artificiality.

Shared augmented reality systems such as *Shared Space* [Billinghurst et al. 1996] and *Studiestube* [Schmalstieg et al. 1996] supplement the user's immediate physical surroundings with additional synthetic information (e.g., annotations and projections), but the immediate surroundings remain the first consideration. They are essentially local in terms of transportation and primarily synthetic in terms of artificiality. The same argument applies to the use of ambient display media for introducing digital information into the periphery of a user's local physical environment as proposed by Ishii and Ullmer [1997].

CVEs are primarily synthetic, being composed of 3D geometry and, in the case of information visualizations, other digital objects (e.g., the WWW-3D browser that provides access to HTML Web pages through a 3D visualization of Web structure [Benford et al. 1997c]). However, CVEs may introduce some physicality through the use of texture-mapped video streams for conveying facial expressions or for providing views into remote physical environments [Reynard et al. 1998].

The role of different underlying technologies in locating systems along these dimensions should be noted. Choosing to implement a shared space using video or using interactive graphics will tend to locate it more toward the physical or synthetic ends of the artificiality dimension respectively. The choice of display technology (small screen, large screen, projected, or immersive) tends to locate the resulting system along the transportation dimension. Thus, to generalize, the medium affects artificiality, and the display affects transportation.

2.4 Mixed Realities

This classification highlights the close relationships between the various approaches and in turn raises the issue of how they might be integrated. This leads us to the idea of *mixed realities* as new forms of shared space that span these dimensions and that integrate the local and remote and the physical and synthetic.

However, before exploring the idea of mixed realities in more detail, we will first extend our classification to include a further dimension, that of spatiality. This will help us to reason about the way in which these various

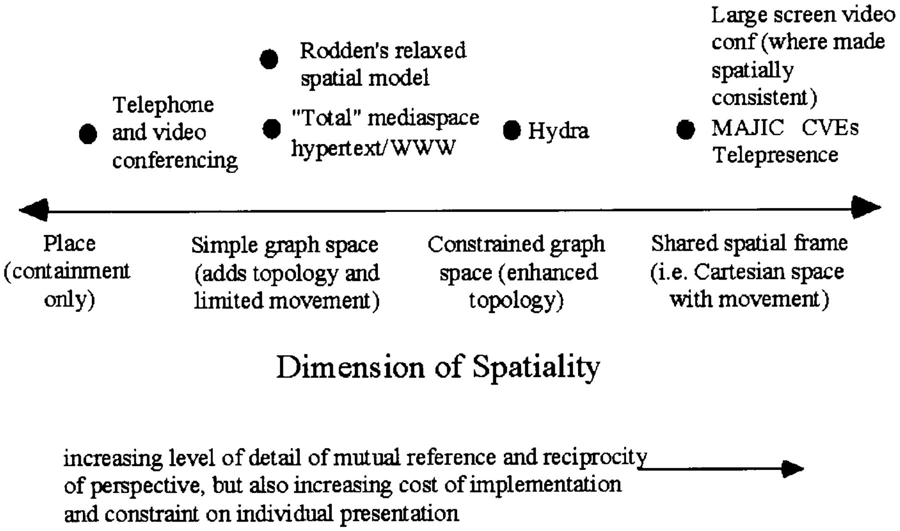

Fig. 3. Classification of shared spaces according to the dimension of spatiality.

approaches differ in their underlying interpretation of shared space and the possible consequences of these differences for collaborative activity. In Section 3 we will use this definition of spatiality to reason about the problems encountered in staging a mixed-reality poetry performance. This analysis will then motivate the specific approach of *mixed-reality boundaries* that is developed in Sections 4, 5, and 6.

2.5 Spatiality

A further dimension that may be used to characterize shared-space systems is their degree of spatiality. This concerns their level of support for fundamental physical spatial properties such as containment, topology, distance, orientation, and movement as shown in Figure 3. Its extremes are characterized by the notions of place, a containing context for participants; and space, a context that further provides a consistent, navigable, and shared spatial frame of reference (e.g., a Cartesian coordinate system).

Unlike the previous two dimensions, that might potentially be applied to CSCW systems in general, spatiality (obviously) applies specifically to the kinds of shared-space system discussed in this article.

We can populate the dimension of spatiality with specific examples of shared-space technologies. We begin with the basic notion of shared space. Logically, the shared space in which a cooperative activity occurs can be defined to be those aspects of the system which are independent of any single participant or group of participants and which are agreed on by the participants. This is expanded and illustrated in the following observations of current systems.

In minimal video-conferencing (single camera per group of participants, no shared electronic data space), the space which is independent of each participant, and on which all participants can agree, is the set of participants and their allocation to cameras. There is no higher-level relationship between the groups of participants. The use of a single camera per participant prevents participants from tailoring their visual communication actions toward any subset of recipients. Although one can be seen to move within the image provided by the camera, other participants cannot interpret this with respect to their own spatial frame of reference. In other words, the movement is only taking place within one participant's local space, not within a shared space. This type of system is characteristic of the placeful extreme on the spatiality scale. This is because there is a place—the conference—in which the participants agree they are, in some sense, present. However, there is no shared internal structure to this place, no dimensions or controls, and therefore little spatiality, other than containment.

By creating a shared drawing surface between two video views, the *Clearboard* system extends the spatiality of video-conferencing to allow some movements and gestures of the participants to be visible within a shared space. Not only could the participants see one another's faces, but they could also see their hands moving as they worked with objects that were placed on the boundary between them.

The management and control view of a media-space may include explicit notions of place, e.g., offices (or rather the camera views of those offices). Additionally, these places may be organized into a linked network. That network is itself a place (the media-space in its entirety) but is also a graph-space. Within the media-space as a whole, participants can move in well-defined ways between different subplaces. The set of subplaces and the ways in which they are related define the total media-space. This type of system therefore moves beyond basic containment to include some further spatiality in the form of a graph-like topology and limited movement.

The *Hydra* system is particularly interesting as it supports the property of orientation and so is more strongly spatial again. In *Hydra*, the space is the conference within which are the participants. Unlike a minimal video conference, the participants in a Hydra conference are in well-defined relationships to one another, and all participants agree about these relationships. The participants are arranged in a ring with exactly one neighbor on each side, and all participants agree about who is neighbor to whom. However the participants cannot be given consistent spatial locations because each participant "stretches" his or her part of the ring so that anyone can look at all the other participants' local representatives (camera-screen systems) simultaneously. This stretching or warping of the conference space preserves the relationship "is-looking-at," but distorts the angles and inferred positions in any postulated shared Cartesian space. Thus, Hydra's space is structured as a graph but with additional constraints which introduce a limited set of spatially related characteristics such as who is looking at whom.

Video-conferencing systems that use larger displays such as *Clearboard*, the *Portland Experience* [Olson and Bly 1991] (this used a combination of small displays in offices and larger displays in public areas), and *MAJIC* have a shared space that is very closely related to the real world but that does not exist in its entirety in any one physical location. For example, *MAJIC* creates a new meeting space that has mutually consistent extents, positions, etc., but which is composed from three disjoint portions of physical space. Its effect is to take the three 120 degree segments of space (each of the offices involved in the meeting) and join them together into a new physical space. Thus, the conference space is spatial in the same way that a normal meeting space is, and all participants can agree about the nature and arrangement of the conference space. This system therefore introduces the fundamental notion of a shared spatial frame of reference: a commonly defined Cartesian coordinate system through which relative positions and orientations can be measured. However, there is still only limited ability to move about within this spatial frame. For example, one cannot step into another participant's region of the shared space and when moving within one's own space, the views of the other spaces do not change at all.

Another class of fully spatial system is a 3D collaborative virtual environment, in which all participants observe the same (virtual) 3D space and see objects with the same extents and in the same relative positions and orientations. The same can be said of shared augmented-reality systems. Participants differ only in their personal viewpoint, and viewpoints are represented to other participants by bodies; from these each participant can infer what the other participants are seeing. A key aspect of such systems is that they support general movement within a shared spatial frame. Participants can explore a spatial environment, independently moving their own viewpoints while at the same time being aware of the viewpoints of others through their avatars.

2.6 How Much Spatiality Should There Be?

It is important to stress that we are not suggesting that more spatiality is always desirable. Instead, we argue that it is necessary to think in terms of the likely benefits and costs of increasing spatiality. Increasing spatiality implies the ability to establish a reciprocity of perspective at an ever finer level of detail. With containment, participants may only reason about each other's presence in a common place. With a shared topology, they may refine this to reason about mutual location and possibly distance (at least loosely) within a structured space. With a consistent spatial frame of reference, they may reason about mutual orientation and gaze direction and may be able to spatially reference shared objects (e.g., by pointing or by using spatial language such as "over there" or "to the left of").

Another important aspect of spatiality is movement. Beyond the general ability to explore and learn about large synthetic spaces, we suggest that movement may play an important role in dynamic group formation. Many real-time CSCW technologies focus on support for small groups and assume a relatively static model of group membership (e.g., the members of a video conference are often known in advance, and changes in membership are relatively infrequent). However, our experience of the real world suggests that there is no logical reason why real-time systems might not support hundreds or thousands of users (e.g., crush halls, open spaces in town centers, and stadia). Indeed, our own work in the area of entertainment applications of CVEs (so-called Inhabited TV) provides an example of larger-scale real-time CSCW applications [Benford et al. 1998]. In such situations, the ability to move through a crowded space in order to locate others becomes an important issue. Movement also plays a key role in joining and leaving groups. For example, Adam Kendon in his studies of group interaction describes how people use spatial movements to dynamically establish and maintain formations that frame their interaction with others and within which gaze usually operates [Kendon 1990, pp. 209–238].

However, these abilities to reason about mutual position and orientation and to move through a shared environment come at some cost. Considerable implementation and system effort may be needed in order to support an increasing level of spatiality (e.g., to maintain a common 3D coordinate system and to support real-time rendering with a moving viewpoint). In addition, increasing constraints are placed on the individual interface in terms of maintaining an objective world view (i.e., there is less freedom to tailor individual displays). Developers must therefore carefully consider the requirements of a given application. A requirement for knowledge of group membership may be met by containment. A requirement to manipulate a shared 3D artefact may require a shared spatial frame of reference. A requirement to support a large and dynamic population of users may require individually navigable viewpoints.

We can apply this line of argument to the "space versus place" debate that is on-going within the CSCW community and which ran throughout several papers at CSCW'96. For us, the essence of this debate concerns the degree of spatiality that is appropriate to a given application. Thus, on the one hand, several researchers have argued for a focus on "place" without the need for reproducing all of the properties of physical space. For example, the use of locales as developed in Fitzpatrick et al. [1996] stems from the desire to abstract away from physical space and to work with a

> *more encompassing meaning of space in the virtual, independent of graphical and VR depictions, which is driven by social world needs and needs of individuals participating in multiple social worlds* [Fitzpatrick et al. 1996, p. 341].

At the same conference, Harrison and Dourish argued for the defining role of place in framing social behavior [Harrison and Dourish 1996]. They cite examples of systems such as *USENET News* which, they claim, engender a notion of place, and therefore appropriate behaviors, without the need for reproducing all of the properties of physical space.

We have strong sympathies with such approaches which abstract away from the properties of physical space. Indeed, this has been a feature of our own work in developing the *MASSIVE* system. However, as discussed above, we argue that this debate needs to be informed by a detailed understanding of just what the underlying spatial properties are and what their likely benefits and costs might be. For example, both Rodden [1996] and Fitzpatrick et al. [1996] retain the spatial property of containment and, to some extent, topology, but relax the requirement for having a common spatial frame of reference. For us, this implies a choice about cost/benefit trade-offs such that general location is seen as being necessary, while orientation with respect to specific objects arranged within that location is not. This may be a sensible choice for USENET News, but not necessarily for collaborative 3D design (see Shu and Flowers [1992]). What is important, then, is to have a framework which allows one to reason about the degree of spatiality inherent in a given application and the likely costs and benefits that this will entail—hence our dimension of spatiality.

In summary, shared-space systems can be characterized according to their degree of spatiality. The least spatial systems support only the fundamental spatial property of containment. The subsequent introduction of other spatial properties such as topology, movement, and a shared spatial frame of reference results in increased spatiality. The benefits of increased spatiality may be the ability to establish a reciprocity of perspective at a finer level of detail or to dynamically form groups from among a larger population. However, the associated costs may be an increased implementation overhead and increasing constraints on the interface in terms of presenting a synchronized view of the space.

2.7 Comparison with the Taxonomy of Milgram and Kishino

To conclude our classification, we offer a brief reflection on the relationship between our proposed taxonomy and that of Milgram and Kishino which they developed as part of their discussion of mixed reality [Milgram and Kishino 1994]. Like ourselves, they introduced a taxonomy in order to highlight underlying differences between technologies and to motivate an exploration of mixed realities. Their departure point was a classification of different display technologies for augmented reality according to a virtuality continuum. This is a dimension of classification whose extremes were occupied by real and virtual environments and whose midpoints included various forms of mixed reality such as augmented reality and so-called augmented virtuality (the addition of real-world objects to virtual environments). They then refined this classification to include the further dimensions of

—*Extent of world knowledge*: the degree to which the computer system knows about which objects are in a given physical space and where they are (i.e., the extent to which the physical space can be modeled within the computer);

- —*Reproduction fidelity*: the level of image quality of either the real or virtual world; and
- —*Extent-of-presence metaphor*: the extent to which an observer is intended to feel present within the scene.

There are some broad parallels between our taxonomy and that of Milgram and Kishino. Like theirs, our taxonomy also motivates the consideration of mixed realities as forms of shared space that fall toward the midpoints of the dimensions (i.e., that combine notions of local and remote and physical and synthetic). *Extent of world knowledge* is related to *spatiality* in the sense that they both concern the degree to which the system has knowledge of the positions and locations of objects. However, the former focuses on the extent to which the computer system understands the structure and contents of a local physical space whereas the latter focuses on the level of spatiality that is defined across a combination of potentially many physical and virtual spaces that are joined together. The way in which *transportation* differs from the *extent-of-presence metaphor* was discussed in Section 2.1 above, both in terms of introducing remote objects to local spaces and by including transportation of groups of participants and local objects. *Artificiality* is perhaps most closely related to the *virtuality continuum* in that they share a concern for the balance between the physical and the virtual. It is notable that reproduction fidelity is not identified as a primary distinguishing feature in our work.

The differences of detail and interpretation between these classification schemes probably stem from an underlying difference in our starting points. Being derived from a review of CSCW technologies, our work explicitly considers non-Cartesian spaces arising from the interconnection of multiple distinct locations (e.g., the topological spaces created by media-spaces) as well as more familiar Cartesian spaces. In contrast, Milgram and Kishino's work is based on a classification of augmented-reality technologies. Consequently, they are more focused on how to overlay two Cartesian spaces at a fine level of detail (e.g., overlaying synthetic instructions, labels, or graphics on real-world objects) and with issues such as reproduction fidelity and world modelling. These distinctions have led us to a major difference in our approaches to the development of mixed realities as we shall see later in this article.

This concludes our classification of shared-space technologies and our introduction to the dimensions of transportation, artificiality, and spatiality. The following section begins our exploration of mixed reality by using these dimensions to analyze the problems that arose in staging a performance that spanned physical and synthetic theaters. In turn, this will lead us to the idea of mixed-reality boundaries as a way of creating new kinds of shared space that span the local and remote and physical and synthetic.

3. AN EXPERIENCE OF MIXING REALITIES

This section reflects on the practical experience of staging a mixed-reality poetry performance simultaneously within physical and synthetic environments. The aim of the performance was that a series of poets would simultaneously perform poetry to a traditional audience in a physical theater and to an on-line audience in a CVE. This experience demonstrated the complexity involved in the design of mixed-reality applications and provided input to the development of mixed-reality boundaries as described in later sections.

The poetry performance was staged in November 1996 as part of Nottingham's *NOWninety6* media arts festival. The performance involved four hip-hop poets performing in turn in a CVE that had been purposely designed by a graphic artist. The CVE was also populated by 10 virtual audience members at a time, where each audience member was an autonomous, mobile, and embodied participant who could explore the environment and could interact with the other participants using live audio links.

The event was housed in two quite different physical spaces, a theater and a nearby café. The theater was structured as a traditional performance space with a stage and seated audience and was intended to provide a well-managed and well-disciplined environment within which the poets could perform. The café contained the workstations that were used by the 10 virtual audience members and was intended to be a noisy and relatively unstructured environment. Projection was used to provide different views of the virtual world within the theater and café. Specifically they are as follows:

—a view of the virtual performance space was projected into the physical theater alongside the stage so that the physical audience could see the virtual poet embodiment moving alongside the physical poet and could also witness the behavior of the virtual audience members. This view was controlled by a dedicated camera person using a specially created interface that navigated according to a performer-centerd view (i.e., whose movements involved pans, tilts, and zooms with respect to the virtual poet's location). Audio from the virtual world (i.e., conversations between the on-line audience members) was also mixed into the house public address system.

—a view as seen by one of the virtual audience members was projected onto the wall of the café so that all of its occupants (up to 100 people) could keep track of events in the virtual world. The audio from the poets was also broadcast into this space.

The virtual environment itself was structured as a central stage area in which the virtual poet avatars appeared. Poets were assigned their own personalized avatar. Each poet avatar featured a moving head and hands driven by Polhemus motion trackers attached to the head and hands of the associated physical poet. The virtual audience members were assigned angel-shaped avatars and used a conventional workstation to access the virtual world.

Four so-called "outer worlds" were located around the virtual stage. From the outside, they appeared as simple colored cones. On the inside, they

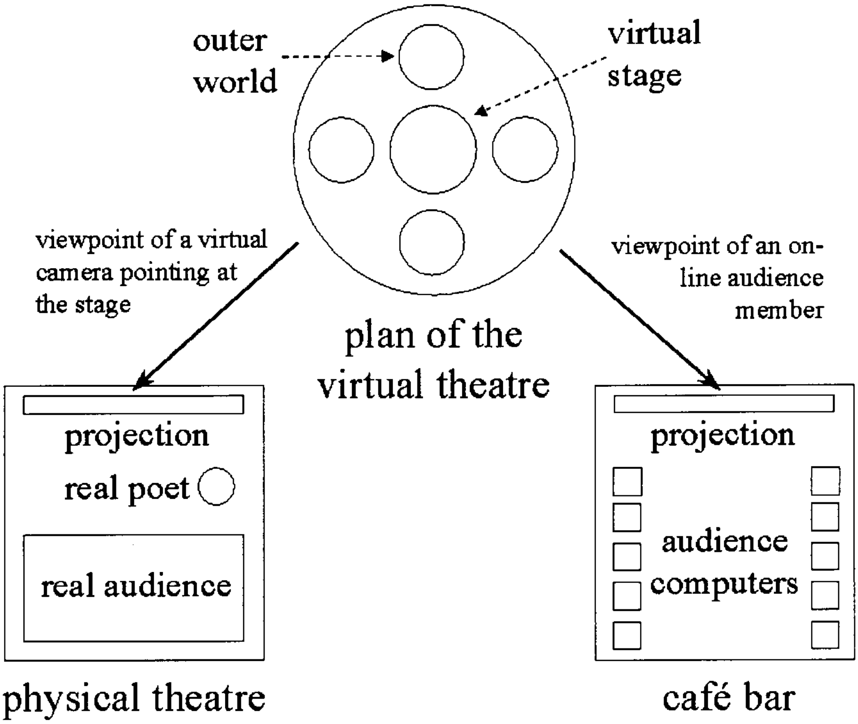

Fig. 4. Projecting views of the virtual theater into the physical theater and café.

contained additional graphics and fragments of text from the poems. These worlds were intended to encourage exploration of the environment and social mingling among the on-line audience members during the breaks between performances.

Figure 4 summarizes the relationship between the virtual theater and the physical theater and café. It shows both the structure of the virtual theater and the way in which this was projected into the two physical spaces.

The screenshot in Figure 5 is taken from near to the virtual stage in the virtual theater. It includes an example virtual poet embodiment as well as one of the angel embodiments used by the on-line audience members.

In terms of the dimensions of artificiality and transportation defined previously, our poetry performance combined synthetic and physical information and established different notions of local and remote for different participants. The on-line audience members were partially transported into a synthetic theater where they could see the virtual poet who, in turn, was driven by the movements of a physical poet in a remote physical space. The audience in the physical theater could see the local physical poets as well as their synthetic counterparts and the on-line audience members. The poetry performance therefore represents a form of mixed reality.

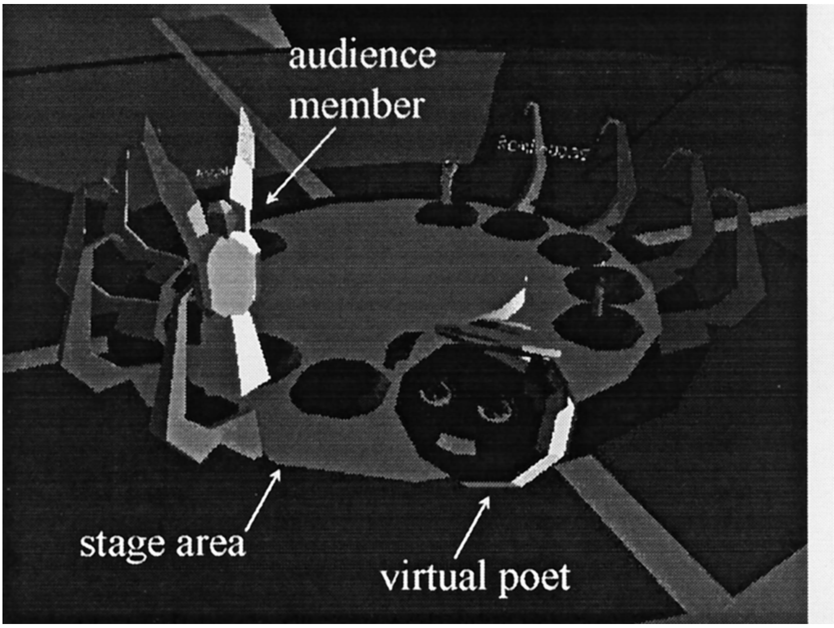

Fig. 5. A view of the stage area and example embodiments.

3.1 Experiences from the Poetry Performance

The following observations have been derived from a variety of sources including video recordings of the event, informal interviews with participants, a final public discussion session, and our own reflections. There were a number of problems with the poetry performance, especially with regard to the interaction between the poets and the angels (the virtual audience members). The angels mostly ignored the poets; they often moved away from their seats within seconds of a performance beginning; they constantly spoke to one another over the poetry; and they sometimes even wandered into the center of the stage during a performance. The poets equally ignored the angels; they did not appear to recognize or react to their apparently rude behavior and carried on the performance regardless. The whole event became separated into two parts: a poetry performance in a physical theater that was enhanced with some interesting graphics; and a social chat environment for members of the public.

There were probably several causes for these problems, spanning social and technical issues. For example, there was little opportunity for user acclimatization and training, and the novelty of the technology meant that there was a lack of established social norms. CVEs do not necessarily have to be a “walk up and use” success, and there may well be significant learning required to use them appropriately. Of course, the angels may simply have preferred to talk to one another rather than to listen to the poetry. However, they had paid money to attend a poetry recital, and their behavior was markedly different from that of their colleagues in the physical theater. We therefore believe that the design of the event itself was a major contributing factor to the above problems. In the following, we focus on two problems in particular, both concerned with the relationship between the synthetic and the physical spaces and the degree of spatiality that was provided.

3.1.1 *Lack of a Globally Integrated Spatial Frame.* The performance involved four distinct spaces, the physical theater, the virtual theater, the café and the tracking space of the Polhemus trackers attached to the poet. The latter can be considered as a separate space, defined by its own local coordinate system within which the poets' movements were recognized and reproduced. Applying the dimension of spatiality, these spaces acted as containers for the different participants and were connected into a distinct topology that defined limited possibilities for moving between them. Each individual space also defined a local shared spatial frame of reference (i.e., its inhabitants could move about within it and could reason about mutual positions and orientations). However, there was *no global, shared, spatial frame of reference* that spanned the combination of spaces. The use of a moving viewpoint for the projection of the virtual theater into the physical theater, while visually attractive, meant that the spatial relationships between the two spaces were constantly shifting. As a result, it was hard for a poet and the angels to establish any reciprocity of perspective. Specifically, a poet could not turn to face or point at an individual angel in order to reference them or to engage their attention. This problem was no doubt compounded by the positioning of the projection screen to the side of and behind the poet. In order to face the virtual world, they would have needed to turn their back on the physical audience, and even then they would have had a very limited view of the screen. Given a choice between the demands of a physically present audience and the demands of a tangential projected display with a shifting viewpoint, it is hardly surprising that the poets chose to perform to the physical audience and not to the virtual one.

3.1.2 *Asymmetrical Awareness between Physical and Virtual Theaters.* A second cause of our problems may have been that the angels in the virtual theater suffered from a lack of knowledge of the activities within the real theater. In contrast to a member of the physical audience who could see the physical poet, the virtual poet, the angels, and their neighbors, an angel could only see the virtual poet and other angels. They could not see the physical poet and could not see or hear the reactions of the physical audience. Had they been able to, they might have been less isolated from the performance and perhaps even felt some social pressure to regulate their behavior.

In summary, we propose that the lack of a globally integrated spatial frame of reference combined with asymmetrical awareness between the physical and virtual theaters played a major part in these two spaces becoming separated. Both spaces worked well as local social environments in their own right, but communication between them was problematic.

These observations suggest the need for a more systematic approach to joining together physical and synthetic, and local and remote spaces. In particular, greater attention must be paid to how different aspects of spatiality are supported across the boundaries between such spaces and the extent to which awareness between them is possible and is mutual. In a broader sense, the problems that we observed involved the regulation of social behavior within virtual spaces and are related to Harrison and Dourish's notion of place. We suggest that the lack of appropriate spatiality was a major contributing factor to the lack of the intended "placeness."

The remainder of this article develops a potential solution to these problems in the form of mixed-reality boundaries that join together different kinds of shared space.

4. MIXED-REALITY BOUNDARIES

In this section we introduce the concept of mixed-reality boundaries and explore their properties. This is motivated by the observations of the previous sections, namely, that

—the classification of shared-space technologies points toward the close relationships between current approaches and suggests their possible integration to create new forms of mixed reality and

—our early experience of social interaction between physical and synthetic spaces implies the need for a more systematic approach to joining them together.

Milgram and Kishino refer to mixed reality as "the merging of real and virtual worlds" such that "real world and virtual world objects are presented together within a single display" [Milgram and Kishino 1994]. We would broaden this to consider the joining together of whole environments. There might be many valid ways of constructing mixed realities based on different kinds of display technology. For example, Milgram and Kishino's own approach is based on the use of different augmented-reality displays to present the user with a synthesized image of a single physical space combined with a virtual one. This might be suited to a range of applications which involve fine-grained interaction with physical objects and their virtual analogues such as medical imaging, telesurgery, machine maintenance, and the control of robots.

Our approach to mixed reality is quite different. Driven by the concerns of supporting new forms of awareness and communication between the inhabitants of many distributed spaces, we develop the approach of creating transparent boundaries between physical and synthetic spaces. Thus, instead of being superimposed, two spaces are placed adjacent to one another and then stitched together by creating a "window" between them. One advantage of our approach is that we might use multiple boundaries to join many different spaces into a much larger superspace. In the long term, these two different approaches may prove to be complementary, with one focusing on the "microlevel" issues of merging two specific spaces and the

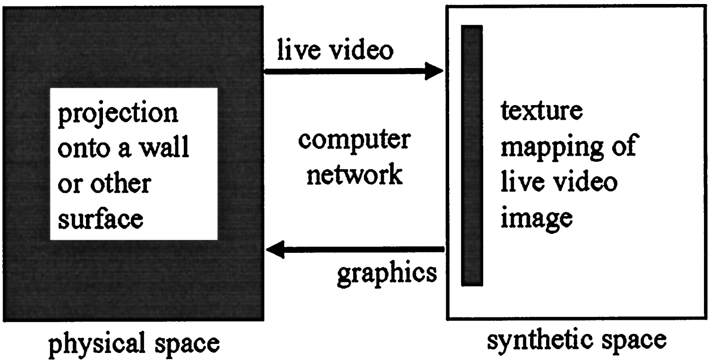

Fig. 6. Creating a simple mixed-reality boundary.

other focusing on addressing the "macrolevel" issues of building large mixed-reality structures from many individual spaces.

We begin by considering the general properties of a single mixed-reality boundary that might be used to join two spaces. We illustrate this idea with a simple example, although it should be noted that such boundaries might potentially take many different and possibly more sophisticated forms. Our example, illustrated in Figure 6, shows how a physical space might be linked to a synthetic space through the creation of a simple boundary. This is based on a combination of projecting graphics into the physical space and texturing video into the virtual space. In other words, the changing geometry of the synthetic space and the avatars within it would be transmitted across the network, rendered and then projected into the physical space. At the same time, a live video image of the physical space would be transmitted across the network and then displayed in a synthetic space through a process of dynamically texture mapping the incoming frames so that it appeared as an integrated part of the virtual environment. Consequently, the inhabitants of the physical space would see the synthetic space as an extension of their physical environment and vice versa. Given an additional audio link between the two spaces, the inhabitants of each would be able to communicate directly with one another.

This example relies on two underlying techniques. First, the texture mapping of live video streams is required in the virtual environment. This has already been demonstrated in a number of existing collaborative virtual environments including *Interspace* [Suzuki 1995], *Freewalk* [Nakanishi et al. 1996], and *DIVE* [Fahlén et al. 1993]. Second, some form of projected interface must be available within the physical space such as the use of a portable video projector, a multiprojector display, or a CAVE [Cruz-Neira et al. 1992].

Of course, this simple example only involves establishing a boundary between a physical space and a synthetic space. We might also create boundaries between two physical spaces or between two synthetic ones. Thus, in terms of the dimension of artificiality, there are three general types of boundary:

—*physical-synthetic*: as illustrated in Figure 6.

—*physical-physical*: for example, the Portland Experience or the wide variety of existing "real-world" boundaries such as doors, windows, curtains, etc.

—*synthetic-synthetic*: boundaries between synthetic spaces, including portals that link distinct virtual worlds and bounded subregions of a single world that have different effects on mutual awareness [Benford et al. 1997a].

In terms of the dimension of transportation, the broad aim of introducing such boundaries is to reduce the separation between local and remote (i.e., to make two remote spaces feel as if they are local to one another).

Having introduced the concept of mixed-reality boundaries, we now explore their properties in more detail.

4.1 Awareness, Transparency, and Privacy

The first property considers the kinds of effects that a boundary might have on awareness and communication. Will the boundary be completely transparent (i.e., freely allow information to pass across it), or will it alter awareness in some way, perhaps being more opaque? For example, mixed-reality boundaries might attenuate awareness in the same way that many real-world boundaries do (e.g., the effects of frosted glass windows). They might also amplify awareness, acting as a kind of lens.

4.2 Medium-Specific Effects

Boundaries might apply their effects differently in different media. For example, a boundary might have no effect on visual awareness but might attenuate audio awareness (rather like a window in the real world). Awareness attenuation in different media might be implemented through different level-of-detail techniques such as controlling the volume of audio or the level of detail of graphical rendering. For example, Hudson and Smith [1996] describe a technique for introducing various degrees of privacy into live video data of real-world scenes such that general activity is conveyed, but specific details are masked from view.

4.3 Directionality and the Balance of Power

The effects of boundaries might be directional. Thus, one could construct boundaries which were transparent in one direction but opaque in the other, or boundaries which supported interaction in one direction but not in the other (e.g., where the occupants of a physical space could manipulate the contents of a synthetic one but not vice versa). The poetry experience suggests that this property should be treated with caution. It also raises issues of mutuality, privacy, and balance of power. Of course, there may be circumstances in which one way boundaries are useful (e.g., the relatively rare uses of one-way mirrors as physical-physical boundaries). However, the effects of such boundaries on the power relationships between the occupants of different spaces must be carefully thought out.

4.4 Interaction with and across Boundaries

So far, we have described boundaries as being passive conveyors of information. We also need to consider how participants might interact with and across boundaries.

Interaction with a boundary might involve changing the properties of the boundary itself. Participants on either side of the boundary might dynamically control its transparency and other effects so as to reconfigure the possibilities for observation and communication and hence establish variable degrees of privacy. Navigation is another form of interaction with a boundary. In this case, the boundary might become mobile within one or both of the connected spaces (e.g., it might represent a moveable viewpoint within a synthetic space or be attached to a mobile camera within a physical one). Participants might then steer the boundary through the connected spaces. As an additional note, an interesting approach to interaction with projected displays in physical environments is demonstrated by the HoloWall system [Matshusita and Rekimoto1997]. This uses reflected infrared light from users' bodies to capture their movements in front of the display and allows them to interact by gesturing or by touching it.

Interaction across a boundary involves participants in one space manipulating objects in the other space. Manipulating digital objects in remote synthetic spaces is relatively easy to support, as this is one of the fundamental properties of CVEs (and similar technologies). On the other hand, manipulating objects in remote physical spaces is more difficult, typically relying on the use of remote effectors such as mobile robots and robot arms, as are being developed in the telepresence community. This would seem to be an important area for future development given the above comments on directionality and the balance of power.

4.5 Group Interaction

Most traditional user interfaces have been designed for single users. However, as a consequence of being situated in socially shared physical and synthetic spaces, mixed-reality boundaries are public and hence inherently shareable. The question then arises as to how a group of participants can usefully interact with and across a shared boundary? For example, how can a group of people standing in front of a projected display in a physical space (such as a public auditorium) usefully and meaningfully interact with it? Simple solutions to this problem might involve some form of floor control so that only one person at a time can interact with the boundary. This approach is commonly used with CAVE-like environments where one participant controls the display and the others are reduced to being passive passengers. More complex solutions might involve tracking and then combining the actions of many participants. For example, people might be given simple voting consoles, or a video camera might be used to track the general spatial distribution of a group in front of a projected display in order to steer a common viewpoint (e.g., so that the viewpoint would steer left if the center of mass of the group were located on the left side of the display).

4.6 Location within Space

In the simple example of Figure 6, the boundary was located on the "walls" of the corresponding physical and synthetic spaces in order to make them appear as direct extensions of one another (or at least to provide a window between the two). However, mixed-reality boundaries could be located at other positions within a space. For example, they might be laid flat in order to create a kind of table. This approach to projection into physical spaces has been demonstrated in the *Digital Desk* from Rank Xerox Research [Wellner et al. 1993] and by the *Responsive Workbench* from GMD [Fleischmann 1995]. The *Digital Desk* provided an enhanced desk environment for working with physical, digital, and hybrid (i.e., physical and digital) documents. Images of digital documents could be projected onto a horizontal desktop and registered with physical documents. Video recognition of users' hand gestures was used as the basis for interaction with the digital information. The Responsive Workbench consisted of a large horizontal screen onto which 3D graphics were back-projected. The use of stereo shutter glasses allowed these graphical images to be seen as 3D objects that had been placed on the table, and users' could interact with these using 3D manipulation devices such as a wand. In *Clearboard* [Ishii and Kobayishi 1992], the boundary was located on a drawing surface, similar to an architect's drawing board. Synthetic information could then be created or accessed on this boundary.

4.7 Synchronization of Spatial Frames

The synchronization of spatial frames refers to whether the spatial coordinate system of one space can be easily related to that of the other via the boundary. As observed in the poetry performance, this may be a major contributing factor to successful communication between the connected spaces. Of course, this also depends on the degree of spatiality of the two spaces as defined previously. For example, given two highly spatial spaces, each with their own Cartesian frame of reference, to what extent does the boundary allow these two frames to be joined into a single frame? In other words, can the participants of one space naturally reason about the positions and orientations of those in the other space with respect to themselves?

4.8 Multiple and Fragmented Boundaries

Given that boundaries might be located in different places within a space, why not use several such boundaries to link two spaces? For example, one might create a first common boundary as a window between two spaces and, at the same time, create a second boundary as a horizontal document display and drawing surface. This approach might be particularly appropriate where multiple cameras are installed in a remote physical space (e.g., in later versions of media-spaces that use multiple cameras to provide an overview of a room, a document viewer, and a close up view of a participant's face). The boundary between the two spaces therefore becomes fragmented. Consequently, the issue of synchronizing spatial frames becomes more problematic as the frames of reference now relate to two different spaces and several locations.

The use of multiple boundaries in this way also raises the possibility of joining together more than two spaces to create a potentially complex topology of inter-linked physical and synthetic environments as we shall describe later on.

5. THE INTERNET FOYER—A DEMONSTRATION OF A MIXED-REALITY BOUNDARY

We now present a demonstration of a mixed-reality boundary called the Internet Foyer. This joins a visualization of an organization's World Wide Web home pages in a CVE to its physical reception area so as to create a combined entry point into its physical and electronic manifestations.

5.1 The Goal of the Internet Foyer

Foyers are important areas of physical buildings for a variety of reasons. First, they present the public face of a building and the organization(s) that it houses. The significance of this function alone should not be underestimated; large sums of money are spent on making foyers interesting and impressive places. Second, foyers provide a context for locating useful information for visitors such as maps, directories, and displays. Third, from a cooperative point of view, they may be home to various people whose job it is to help these visitors (e.g., receptionists). Fourth, they enhance security by providing a single point of entry into the organization within which incoming and outgoing people are made publicly visible and hence accountable. Indeed, foyers may often contain security staff and areas where visitors sign in. Fifth, they provide public meeting places, either for arranged rendezvous or for chance encounters. Some larger foyers contain shops, cafés, and other facilities that reflect this kind of social function.

Turning away from physical space and toward virtual space for a moment, we see that it is clear that many organizations are making increasing use of computer networks, and many have established a public network presence through services such as the World Wide Web (WWW). Thus, in parallel to the physical manifestation of buildings, organizations are increasingly acquiring an electronic manifestation through computer networks. In such cases, the public home pages of an organization on the WWW could be considered to be a kind of foyer—the public entry point into the organization's network manifestation.

However, when considered as foyers, WWW pages leave much to be desired. The people who pass through them are generally not visible to one

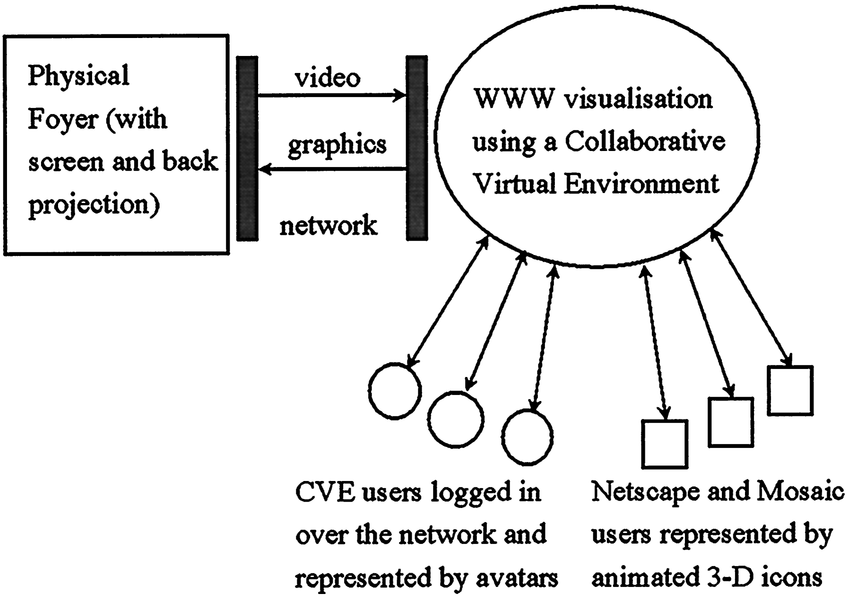

Fig. 7. An overview of the Internet Foyer.

another or to other observers (a major criticism of current WWW technology in general). Security may be compromised, and there are no opportunities for rendezvous and social encounters. Consequently, there are two main goals to the Internet Foyer:

(1) To construct a virtual foyer based on a populated visualization of an organization's WWW space that might come closer to the functionality of a real foyer. On its own, this virtual foyer represents an application of CVE technology—it is a shared synthetic space.

(2) To join this virtual foyer to a real foyer through a simple mixed-reality boundary of the sort shown in Figure 6. Thus, visitors entering an organization's WWW space would be able to communicate with those entering its physical space and vice versa.

5.2 The Functionality of the Internet Foyer

The overall concept of the Internet Foyer is summarized in Figure 7. On the left of the diagram we see a traditional physical foyer. Projected onto the wall of this physical foyer for its inhabitants to see is a graphical visualization of the virtual foyer. On the right of the picture we see the virtual foyer, a 3D graphical visualization of an organization's home pages and their visitors, realized as a CVE. Users of this CVE are mutually embodied and able to communicate with one another through open audio channels. A textured video window in the virtual foyer affords these users a corresponding view back into the real foyer, and an open audio channel allows communication between the two spaces.

There are three ways to experience the Internet Foyer: as a visitor to the physical foyer, as the user of a CVE, and as the user of a traditional WWW browser. The following summarize the functionality of the Internet Foyer as seen by each of these types of user.

As a Visitor to the Physical Foyer. Such users enter a normal physical space. Projected onto a wall of that space is the graphical representation of the virtual foyer. This representation shows a number of linked WWW pages drawn as a 3D network structure. It also shows the presence of virtual reality users in the virtual foyer via graphical user embodiments that move around this visualization. These graphical embodiments may include live video-textured faces in order to convey facial expression. In addition, the visualization shows graphical representations of traditional WWW browser users who are currently accessing the pages. Animated movements show the progress of these users as they flit from page to page. Finally, an audio link allows communication with the virtual reality users in the virtual foyer.

As the User of a Collaborative Virtual Environment. CVE users see the same basic visualization as those in the physical foyer. However, they are able to freely navigate around the visualization, homing in on specific details or backing off in order to obtain an overview. They are also able to select objects in the visualization (both representations of WWW pages and of other users). At present, selecting another object launches a Web browser to display its contents. A real-time video window that is texture-mapped onto a wall of the virtual foyer allows CVE users to look out into the physical foyer and to see its occupants looking back at them. Finally, an open audio link supports communication with other CVE users and with those in the physical foyer.

As the User of a Traditional WWW Browser. These users see the Internet foyer as a series of WWW pages. On entering the foyer they are asked to register themselves using a simple form so that they can voluntarily provide information to personalize their graphical representation (e.g., their name, the URL of their home page, and the location of an image to be used for its face). As they browse the pages in the foyer, they are provided with additional information about current and recent visitors to the virtual foyer, displayed as simple textual lists.

Figures 8, 9, and 10 present images of the Internet Foyer as it appears to different users. Figure 8 shows an overview of the Internet Foyer as it appears to a CVE user. The image shows a visualization of several interlinked WWW pages (spheres connected by arrows) with the video window into the physical foyer in the background.

Figure 9 shows a view from the same user when they have homed in on a specific part of the WWW visualization. This image shows how the presence of the Web browser users is represented in more detail. Selecting one of the spheres or one of these user representations would result in a Web browser being launched in order to display its contents (a WWW page).

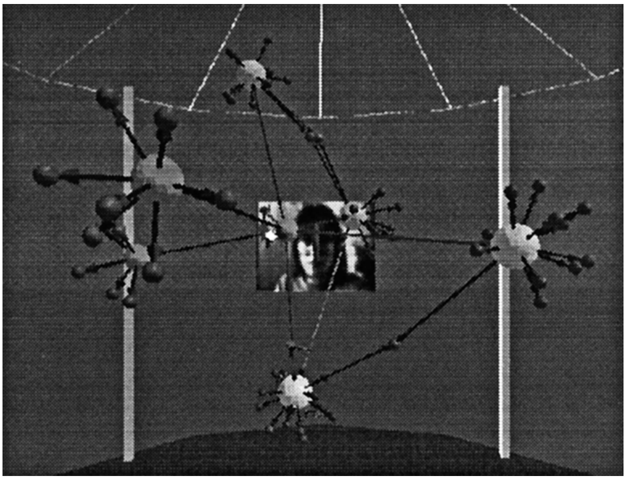

Fig. 8. The Internet Foyer as seen by a CVE user.

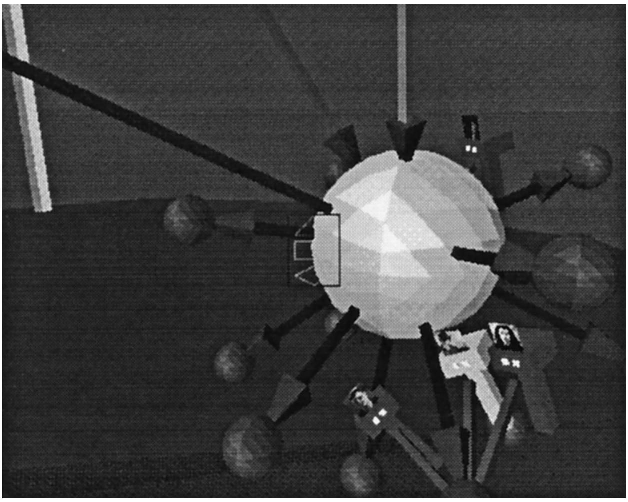

Fig. 9. Close-up to a Web page.

Finally, Figure 10 shows how the Internet Foyer appears to visitors in the physical foyer.

5.3 The Implementation of the Internet Foyer

The implementation of the Internet Foyer relies on the integration of several existing technologies. The following subsections highlight the key techniques used.

5.3.1 *The Collaborative Virtual Environment*. The virtual foyer component of the Internet Foyer has been implemented using the DIVE Collaborative Virtual Environment platform. DIVE is a general-purpose toolkit which has been developed by the Swedish Institute of Computer Science

Fig. 10. The Internet Foyer seen from the physical foyer.

[Fahlén et al. 1993]. DIVE is based on a distributed database model where local copies of a virtual world database are maintained in a consistent state as a result of the transmission of updates between them.

5.3.2 *Constructing the Visualization.* The visualization of the connected WWW pages has been produced by an application called FDP-Grapher (which has been implemented in DIVE). FDP-Grapher dynamically constructs 3D visualizations of network structures using the Force Directed Placement (FDP) technique [Fruchterman and Rheingold 1991]. This technique is based on a physical simulation model that treats the nodes of the network as masses and the arcs as springs. The whole structure is placed in a random initial configuration, and a series of iterations are performed in which the physical effects of the springs on the masses are simulated. This continues until the visualization settles in a stable state. The resulting visualizations typically group strongly interlinked nodes into spatial clusters.

In the Internet Foyer, FDP-Grapher has been connected to a simple WWW robot that explores a region of the WWW as defined by an initial URL and a link adjacency distance. Thus, the visualization is capable of charting arbitrary regions of the WWW (up to approximately a hundred nodes before scaleability problems set in with the FDP implementation). The FDP algorithm has also been adjusted to treat single outlying pages as lighter nodes and strongly linked pages as heavier ones in order to produce a more legible final visualization.

5.3.3 *Video and Audio Support.* The Internet foyer includes the use of texture-mapped video streams, both to provide CVE participants with a view of the remote physical foyer and to introduce facial expressions onto their embodiments. This involves a real-time video stream being attached to a surface (currently a single polygon) in the *DIVE* environment and being constantly retextured as new frames arrive. The video data stream is transmitted over a multicast protocol and is currently capable of supporting a video stream plus audio channels and virtual world updates on a standard Ethernet and achieves a video frame rate in excess of 10 frames per second. The audio data stream is supported by dedicated audio server processes that run over UDP.

5.3.4 *Tracking WWW Browsers*. The final software component of the Internet Foyer is called FollowWWW. This is a general package for tracking the presence of WWW browsers as they pass through a server and for multicasting this information to interested parties (e.g., the virtual foyer visualization). These users are then represented by simple graphical embodiments in DIVE which are animated to reflect movement between different WWW pages in the visualization. FollowWWW can use either live data or WWW server log files as its input.

5.4 Properties of the Internet Foyer Boundary

We now briefly reflect on the properties of the Internet Foyer boundary. The Internet Foyer boundary is transparent, operates identically across the visual and audio media, and is fully bidirectional. Once established there is no opportunity for interacting with or across the boundary beyond talking to remote participants. In both the physical and virtual foyer, the boundary takes the form of a window that is located on a convenient wall. The spatial frames on either side of the boundary are synchronized in that left and right have consistent and expected meanings. Thus, the Internet Foyer demonstrates what is probably the most basic boundary configuration possible.

5.5 Classifying the Internet Foyer

So where does the Internet Foyer fit into our classification scheme? First, it clearly combines aspects of telepresence (looking into the remote physical foyer); media-spaces (support for peripheral awareness and chance encounters); CVEs (shared 3D data visualizations); and spatial video-conferencing (support for video faces attached to graphical bodies within a common spatial frame of reference).

With reference to our various dimensions, the Internet Foyer combines aspects of both the physical and synthetic worlds into a single system, placing it along the midpoint of the artificiality dimension. As the Internet Foyer can be accessed either though a physical foyer or remotely using CVE technology, it might be located toward either of the local or remote points of the scale of transportation. This raises the issue of heterogeneity. A key aspect of the Internet Foyer is that it provides different kinds of access to different users. Thus, visitors to the real foyer experience a reasonably familiar local environment extended through projected graphics, whereas virtual users experience a new remote one.

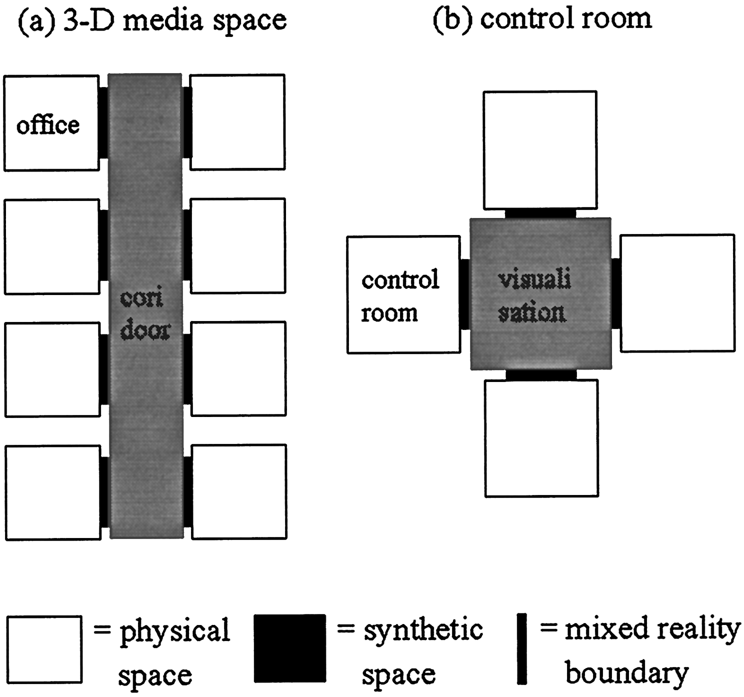

Fig. 11. Synthetic spaces linking multiple physical spaces.

In terms of its spatiality, the Internet Foyer is strongly spatial. It provides a global shared spatial frame of reference and therefore supports containment, topology, orientation, and movement.

6. TESSELLATED MIXED REALITIES

Having demonstrated the use of a single mixed-reality boundary to join together two spaces, we now further extend the boundary concept by considering how multiple physical and synthetic spaces might be linked together through the use of multiple boundaries in order to create a structured mixed reality. We propose that this ability to join together many spaces is a powerful and distinguishing feature of our boundary-based approach to mixed reality.

In the case where all of the spaces were physical, the result would be a form of media-space. In the case where they were all synthetic, the result would be a structured CVE. The hybrid case, however, raises some interesting new possibilities and leads to the idea of tessellated mixed realities.

From the point of view of a virtual space, different video views could be texture mapped into different locations. For example, a new form of navigable 3D media-space might be created by linking several video views into a navigable synthetic corridor (i.e., one could navigate through a 3D structure in order to access different video views into people's physical offices). Alternatively, a single synthetic space such as an information

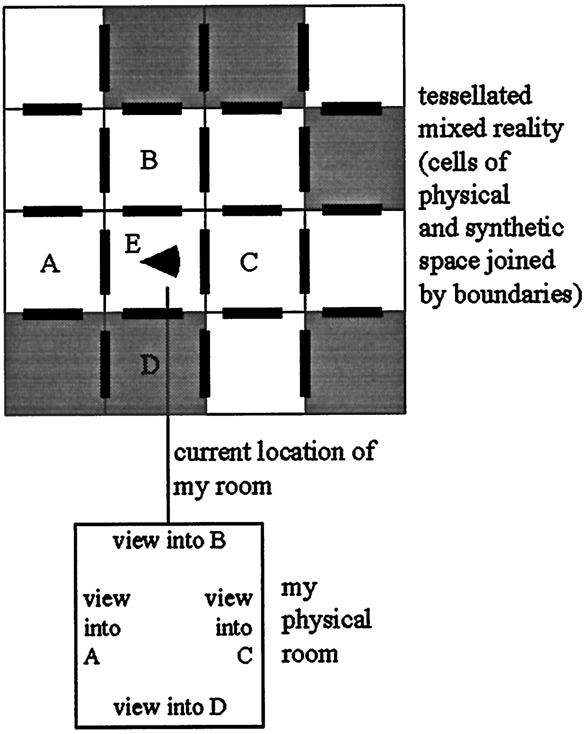

Fig. 12. Using a physical room as a vehicle to move through a tessellated mixed reality.

visualization, might be used as a central linking point for several physical spaces. A team of distributed software engineers might link their offices through a common visualization of a software structure, or a series of separate Air Traffic Control Rooms might be linked through a common 3D visualization of air-space.

These possibilities are briefly summarized in Figure 11, where 11(a) shows the possible use of a 3D corridor in a media-space, and Figure 11(b) shows the use of a shared 3D visualization to link physical control rooms.

From the point of view of a physical space, multiple linkages might involve locating different boundaries on different surfaces of a room. In this case, a new form of navigation might be achieved by dynamically switching some of these views to portray new scenes.

Given a consistently defined overall topology of tessellated cells linked by mixed-reality boundaries, a local physical room could become a vehicle to be navigated through a tessellated mixed reality. Figure 12 depicts a mixed-reality constructed from square cells, each of which is linked to its neighbors by an appropriate mixed-reality boundary. A physical room might logically be located in any one of these cells and given an orientation. This would determine the projected views to be shown on each of its four walls. Thus, in Figure 12, the user's physical room (vehicle) is currently located in cell E, facing cell A, thereby causing views into cells A, B, C, and D to be projected onto its walls.

The ability to occupy a given cell would be subject to various constraints, such as there being a match between the properties of the boundaries of the cell and the available equipment within the particular room.

Given that a room can occupy a cell, it may be possible to move the room through the overall structure. For example, the room in cell E might be moved into one of the adjacent cells A, B, C, or D; might be rotated to face another cell; might take a guided tour through several cells; or might jump to some other location altogether. We propose that these ideas of structured and tessellated mixed realities open up new possibilities for a general merging of the physical and synthetic worlds and represent a key direction for future research into mixed realities.

7. SUMMARY

This article has focused on the nature of shared spaces, a topic that has gathered considerable interest over recent years. It has contributed to the growing debate about the interpretation of shared space, with a particular focus on reconsidering the relationship between real and virtual spaces. This discussion has motivated the development of new techniques for creating shared mixed realities based upon the approach of mixed-reality boundaries.

The first contribution of the article has been to review current spatial approaches to CSCW (i.e., media-spaces, spatial video-conferencing, CVEs, telepresence applications, and collaborative augmented environments) with a view to understanding the fundamental differences among them. This review classified current approaches along the three dimensions of transportation, artificiality, and spatiality. Transportation concerns the degree to which users are transported into some new space or remain in their local space. Artificiality concerns the degree to which the shared space is based on real-world information or is synthesized. Spatiality concerns the degree to which the shared space exhibits key spatial properties such as containment, topology, movement, and a shared frame of reference. This classification has established the general relationships between physical spaces, augmented realities, telepresence, and CVEs. It has also led us to think about hybrid approaches that combine different kinds of shared space and that might lie at the center of our classification. These hybrid spaces represent forms of mixed reality—shared spaces that combine the physical and synthetic and the local and remote.

The second contribution of the article has been to present a practical experience of a mixed-reality collaboration. We staged a poetry performance that occurred simultaneously in a physical and virtual theater, where these were joined together using a combination of projected displays, tracking of performers' movements, and live audio. Although each of these spaces functioned well as a social environment in its own right, an integrated mixed reality was not successfully established. In particular, the poets in the physical theater were unable to engage the attention of the on-line audience members in the virtual theater. We discussed two factors that may have contributed to these problems. First, there was no consistent global spatial of frame of reference that spanned both spaces. Second, the boundary between the two spaces was not symmetric in terms of awareness and communication. These observations have motivated a more systematic approach to creating shared mixed realities.

The third contribution of the article has been to explore one particular style of mixed reality, based on the idea of creating transparent boundaries between the physical and the synthetic. This contrasts with previous approaches that have focused on how a pair of spaces may be superimposed through the use of augmented-reality technologies. Our article has proposed how a simple mixed-reality boundary might be created through a combination of video projection into a physical space and video texturing in a synthetic one. It has also identified some general properties of such boundaries, including their degree of transparency, the possibilities for interaction with and through them, and the location of multiple boundaries within a single space. We have presented an initial demonstration of a mixed-reality boundary called the Internet Foyer. This joins a visualization of an organization's home pages on the World Wide Web to its physical reception area, so as to create a single entry point into the organization that can be shared by both its physical and virtual visitors. Finally we have speculated on the possibility of using multiple mixed-reality boundaries to link together many physical and synthetic spaces to create a form of tessellated mixed reality.

There are many possible applications of shared mixed realities. The most obvious candidates are those which involve knowledge of events in the real world combined with distributed access to electronic data, such as

—doctors whose diagnoses might involve the ability to see a real live patient (e.g., to sit in on a remote clinical session) and to visualize 3D scan data captured by various medical imaging techniques,

—construction engineers who need to discuss engineering plans and data within the context of an emerging physical building,

—environmental planners who wish to discuss geographical data captured through environmental remote sensing techniques in the context of some real-world environmental development, and

—distributed control rooms (e.g., Air Traffic Control) where different physical control rooms might be joined through a common visualization.

The extent to which such applications might be located along the different dimensions of transportation, artificiality, and spatiality will no doubt depend upon the nature of the data being discussed, whether remote collaboration is required and other factors such as available technologies. However, it seems unlikely that any of them should fall at the extremes of any of our dimensions, or could be fully supported by any of the current spatial approaches. In conclusion, we argue that a fruitful direction for future research into shared spaces is to consider techniques for integrating physical and synthetic environments into varying forms of mixed reality and that mixed-reality boundaries are a potentially powerful tool for achieving this.

ACKNOWLEDGMENTS

We would like to thank the European Commission and the UK's EPSRC for funding this work through the eRENA project (ESPRIT IV I3 programme) and the *Multimedia Networking for Inhabited Television* project (Multimedia Networking Applications programme) respectively.

REFERENCES

Benford, S., Bowers, J., Fahlen, L. E., Mariani, J., and Rodden, T. 1994. Supporting co-operative work in virtual environments. *Comput. J. 37*, 8.

Benford, S. D., Brown, C. C., Reynard, G. T., and Greenhalgh, C. M. 1996. Shared spaces: Transportation, artificiality and spatiality. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '96, Boston, MA, Nov. 16–20). ACM, New York, NY, 77–86.

Benford, S. D., Greenhalgh, C. M., and Lloyd, D. 1997a. Crowded collaborative environments. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '97, Atlanta, GA, Mar. 22–27). ACM, New York, NY, 59–66.

Benford, S. D., Greenhalgh, C. M., Snowdon, D. N., and Bullock, A. N. 1997b. Staging a public poetry performance in a collaborative virtual environment. In *Proceedings of the 5th European Conference on Computer-Supported Cooperative Work* (ECSCW '97, Lancaster, UK, Sept.). Kluwer B.V., Deventer, The Netherlands.

Benford, S. D., Snowdon, D. N., Brown, C. C., Reynard, G. T., and Ingram, R. J. 1997c. The populated Web: Browsing, searching and inhabiting the WWW using collaborative virtual environments. In *Proceedings of the 6th IFIP Conference on Human Computer Interaction* (INTERACT '97, Sydney, Australia, July 14–18). IFIP, Laxenburg, Austria, 14–18.

Benford, S., Greenhalgh, C., Brown, C., Walker, G., Regan, T., Rea, P., Morphett, J., and Wyver, J. 1998. Experiments in inhabited TV. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '98, Los Angeles, CA, Apr. 18–23). ACM Press/Addison-Wesley Publ. Co., New York, NY.

Billinghurst, M., Weghorst, S., and Furness, T., III 1996. Shared Space: An augmented reality interface for computer supported collaborative work. In *Proceedings of the Workshop on Collaborative Virtual Environments* (CVE '96, Nottingham, UK, Sept.).

Bly, S. A., Harrison, S. R., and Irwin, S. 1993. Media spaces: Bringing people together in a video, audio, and computing environment. *Commun. ACM 36*, 1 (Jan.), 28–46.

Bowers, J., Pycock,, J., and O'Brien, J. 1996. Talk and embodiment in collaborative virtual environments. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '96). ACM Press, New York, NY.

Cook, S., Birch, G., Murphy, A., and Woolsey, J. 1991. Modelling groupware in the electronic office. In *Computer-Supported Cooperative Work and Groupware*, Greenberg, S., Ed. Harcourt Brace & Co., Orlando, FL.

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., and Hart, J. C. 1992. The CAVE: Audio visual experience automatic virtual environment. *Commun. ACM 35*, 6 (June), 64–72.

Dourish, P. and Bly, S. 1992. Portholes: Supporting awareness in a distributed work group. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '92, Monterey, CA, May 3–7), P. Bauersfeld, J. Bennett, and G. Lynch, Eds. ACM Press, New York, NY, 541–547.

FAHLÉN, L. E., BROWN, C. G., STÅHL, O., AND CARLSSON, C. 1993. A space based model for user interaction in shared synthetic environments. In *Proceedings of the ACM Conference on Human Factors in Computing* (INTERCHI '93, Amsterdam, The Netherlands, Apr. 24–29), B. Arnold, G. van der Veer, and T. White, Eds. ACM Press, New York, NY, 43–48.

FITZPATRICK, G., KAPLAN, S., AND MANSFIELD, T. 1996. Physical spaces, virtual places and social worlds: A study of work in the virtual. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '96, Boston, MA, Nov. 16–20). ACM, New York, NY, 334–343.

FLEISCHMANN, M. 1995. Imagination systems—Interface design for science, art and education. In *Conference Proceedings of Imagina '95* (Imagina '95, Monte Carlo, Feb. 1–3). 257–261.

FRUCHTERMAN, T. M. J. AND REINGOLD, E. M. 1991. Graph drawing by force-directed placement. *Softw. Pract. Exper. 21*, 11 (Nov.), 1129–1164.

GAVER, W. W. 1992. The affordances of media spaces for collaboration. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '92, Toronto, Canada, Oct. 31–Nov. 4), M. Mantel and R. Baecker, Eds. ACM Press, New York, NY, 17–24.

GAVER, W., MORAN, T., MACLEAN, A., LÖVSTRAND, L., DOURISH, P., CARTER, K., AND BUXTON, W. 1992. Realizing a video environment: EuroPARC's RAVE system. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '92, Monterey, CA, May 3–7), P. Bauersfeld, J. Bennett, and G. Lynch, Eds. ACM Press, New York, NY, 27–35.

GAVER, W. W., SELLEN, A., HEATH, C., AND LUFF, P. 1993. One is not enough: Multiple views in a media space. In *Proceedings of the ACM Conference on Human Factors in Computing* (INTERCHI '93, Amsterdam, The Netherlands, Apr. 24–29), B. Arnold, G. van der Veer, and T. White, Eds. ACM Press, New York, NY, 335–341.

GREENHALGH, C. AND BENFORD, S. 1995. MASSIVE: A collaborative virtual environment for teleconferencing. *ACM Trans. Comput. Hum. Interact. 2*, 3 (Sept.), 239–261.

HARRISON, S. AND DOURISH, P. 1996. Re-place-ing space: The roles of place and space in collaborative systems. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '96, Boston, MA, Nov. 16–20). ACM, New York, NY, 67–76.

HEATH, C. AND LUFF, P. 1991. Collaborative activity and technological design: Task coordination in London underground control rooms. In *Proceedings of the European Conference on Computer-Supported Cooperative Work* (ECSCW '91), L. Bannon, M. Robinson, and K. Schmidt, Eds.. Kluwer B.V., Deventer, The Netherlands.

HUDSON, S. E. AND SMITH, I. 1996. Techniques for addressing fundamental privacy and disruption tradeoffs in awareness support systems. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '96, Boston, MA, Nov. 16–20). ACM, New York, NY, 248–257.

HUGHES, J. A., RANDALL, D., AND SHAPIRO, D. 1992. Faltering from ethnography to design. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '92, Toronto, Canada, Oct. 31–Nov. 4), M. Mantel and R. Baecker, Eds. ACM Press, New York, NY, 115–122.

ICHIKAWA, Y., OKADA, K., JEONG, G., TANAKA, S., AND MATUSHITA, Y. 1995. MAJIC videoconferencing system: Experiments, evaluation and improvement. In *Proceedings of the European Conference on Computer-Supported Cooperative Work* (ECSCW '95, Stockholm, Sweden, Sept.). Kluwer B.V., Deventer, The Netherlands.

ISHII, H. AND ULLMER, B. 1997. Tangible bits: Towards seamless interfaces between people, bits, and atoms. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '97, Atlanta, GA, Mar. 22–27). ACM, New York, NY, 234–241.

ISHII, H., KOBAYASHI, M., AND GRUDIN, J. 1992. Integration of inter-personal space and shared workspace: ClearBoard design and experiments. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '92, Toronto, Canada, Oct. 31–Nov. 4), M. Mantel and R. Baecker, Eds. ACM Press, New York, NY, 33–42.

KENDON, A. 1990. *Conducting Interaction: Patterns of Behaviour in Focused Encounters*. Cambridge University Press, New York, NY.

KUZUOKA, H., ISHIMODA, G., AND NISHIMURA, T. 1995. Can the GestureCam be a surrogate?. In *Proceedings of the European Conference on Computer-Supported Cooperative Work* (ECSCW '95, Stockholm, Sweden, Sept.). Kluwer B.V., Deventer, The Netherlands.

MATSUSHITA, N. AND REKIMOTO, J. 1997. HoloWall: Designing a finger, hand, body, and object sensitive wall. In *Proceedings of the ACM Symposium on User Interface Software and Technology* (UIST '97, Banff, Alberta, Canada, Oct. 14–17). ACM Press, New York, NY, 209–211.

MILGRAM, P. AND KISHINO, F. 1994. A taxonomy of mixed reality visual displays. *IEICE Trans. Inf. Syst. E77-D*, 12 (Dec.).

NAKANISHI, H., YOSHIDA, C., NISHIMURA, T., AND ISHIDA, T. 1996. Freewalk: Supporting casual meeting in a network. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '96, Boston, MA, Nov. 16–20). ACM, New York, NY, 308–314.

OLSON, M. H. AND BLY, S. A. 1991. The Portland experience: A report on a distributed research group. In *Computer-Supported Cooperative Work and Groupware*, Greenberg, S., Ed. Academic Press Computers and People Series. Academic Press Ltd., London, UK, 81–98.

REYNARD, G., BENFORD, S., AND GREENHALGH, C. 1998. Awareness driven video quality of service in collaborative virtual environments. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '98, Los Angeles, CA, Apr. 18–23). ACM Press/Addison-Wesley Publ. Co., New York, NY.

RODDEN, T. R. 1996. Populating the application: A model of awareness for cooperative applications. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '96, Boston, MA, Nov. 16–20). ACM, New York, NY, 87–96.

ROOT, R. W. 1988. Design of a multi-media vehicle for social browsing. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '88, Portland, OR, Sept. 26–28), I. Greif, Ed. ACM Press, New York, NY, 25–38.

SACKS, H., SCHEGLOFF, E., AND JEFFERSON, G. 1974. A simplest systematics for the organisation of turn-taking in conversation. *Language 50*, 696–735.

SARIN, S. AND GREIF, I. 1985. Shared-data conferencing: Computer-based real-time conferencing systems. *IEEE Comput. 18*, 10 (Oct.), 33–45.

SCHMALSTIEG, D., FUHRMANN, A., SZALAVARI, Z., AND GERVAUTZ, M. 1996. Studierstube—An environment for collaboration in augmented reality. In *Proceedings of the Workshop on Collaborative Virtual Environments* (CVE '96, Nottingham, UK, Sept.).

SELLEN, A., BUXTON, B., AND ARNOTT, J. 1992. Using spatial cues to improve videoconferencing. In *Proceedings of the ACM Conference on Human Factors in Computing Systems* (CHI '92, Monterey, CA, May 3–7), P. Bauersfeld, J. Bennett, and G. Lynch, Eds. ACM Press, New York, NY, 651–652.

SHERIDAN, T. B. 1992. Musings on telepresence and virtual presence. *Presence: Teleoper. Virtual Environ. 1*, 1 (Winter), 120–126.

SHU, L. AND FLOWERS, W. 1992. Groupware experiences in three-dimensional computer-aided design. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work* (CSCW '92, Toronto, Canada, Oct. 31–Nov. 4), M. Mantel and R. Baecker, Eds. ACM Press, New York, NY, 179–186.

SLATER, M., USOH, M., AND STEED, A. 1994. Depth of presence in virtual environments. *Presence: Teleoper. Virtual Environ. 3*, 2 (Spring), 130–144.

SLATER, M., USOH, M., AND STEED, A. 1995. Taking steps: The influence of a walking technique on presence in virtual reality. *ACM Trans. Comput. Hum. Interact. 2*, 3 (Sept.), 201–219.

STONE, R. J. 1991. Advanced human-system interfaces for telerobotics using virtual reality and telepresence technologies. In *Proceedings of the 5th International Conference on Advanced Robotics in Unstructured Environments* (ICAR'91).

SUZUKI, G. 1995. Interspace: Toward networked virtual reality of cyberspace. In *Conference Proceedings of Imagina '95* (Imagina '95, Monte-Carlo, Feb. 1–3).

TAKEMURA, H. AND KISHINO, F. 1992. Cooperative work environment using virtual workspace. In *Proceedings of the 10th International Conference of the Computer Graphics Society on Visual Computing: Integrating Computer Graphics with Computer Vision* (CG

International '92, Tokyo, Japan, June 22–26), T. L. Kunii, Ed. CG International Series. Springer-Verlag, New York, NY, 171–181.

WELLNER, P., MACKAY, W., AND GOLD, R. 1993. Guest editors' introduction to the special issue on "Computer-Augmented Environments: Back to the Real World". *Commun. ACM 36*, 7 (July), 24–26.

ZYDA, M. J., PRATT, D. R., FALBY, J. S., LOMBARDO, C., AND KELLEHER, K. M. 1993. The software required for computer generation of virtual environments. *Presence: Teleoper. Virtual Environ. 2*, 2 (Spring), 130–140.

Received: June 1997; revised: May 1998 and July 1998; accepted: September 1998