

*To what extent is cyberspace
really a space?*

Collaborative Virtual Environments

The technology of Collaborative Virtual Environments (CVEs) aims to transform today's computer networks into navigable and populated 3D spaces that support collaborative work and social play. CVEs are virtual worlds shared by participants across a computer network. Participants are provided with graphical embodiments called avatars that convey their identity, presence, location, and activities to others. They are able to use these avatars to interact with the contents of the world and to communicate with one another using different media including audio, video, graphical gestures, and text.

These worlds can take many forms, and the representations used for the virtual world and avatars can vary tremendously. In this article we wish to briefly review the emergence of CVEs and outline the challenges and future for this significant technology.

The Emergence of CVEs

CVEs can be seen as the result of a convergence of research interests within the VR and computer-supported cooperative work (CSCW) communities. Within the CVEs represent a natural extension of current commercial single-user VR technology to support multiple participants. This extension allows better support for a range of applications. For example, the communication between

instructors and trainees central to simulation and training applications can be supported. Visualizations may also be shared and discussed by teams of scientists or decision-makers. Finally, the ever-expanding variety of multiplayer games and simulators demonstrates the potential of CVEs in leisure and entertainment, the most notable examples being games such as Doom and Quake. In all of these examples, participants are often physically dispersed and communicating over a computer network.

Within the CSCW community CVEs represent a technology that may support some aspects of social interaction not readily accommodated by technologies such as audio and videoconferencing and shared

Steve Benford, Chris Greenhalgh, Tom Rodden,
and James Pycock

desktop applications. Studies of cooperative work in real-world environments have highlighted the important role of physical space as a resource for negotiating social interaction, promoting peripheral awareness, and sharing artifacts [2]. The shared virtual spaces provided by CVEs may establish an equivalent resource for telecommunication. CVEs also have the potential to support crowded online situations where tens or hundreds of participants negotiate social engagement by dynamically forming subgroups. Crowded virtual trading floors, shopping malls, and crush halls are potential examples. Finally, CVEs may enable participants to discuss and manipulate shared 3D models and visualizations in such a way that each can adopt their own viewpoint and can naturally indicate to others where they look and point.

The research interest in CVEs has been complemented by considerable commercial activity, including proposed extensions to the Virtual Reality Modeling Language (VRML97) standard, a widely adopted IEEE standard for single-user interactive 3D graphical worlds. Several companies are currently building Internet-oriented CVE systems, either based on extensions to the VRML standards, or alternatively on proprietary technologies [4] (see this article's sidebar for examples).

Research Challenges for CVEs

The current research and commercial interest in developing CVEs that support rich social interaction in densely populated virtual worlds is an ambitious goal, one that requires addressing a variety of technical challenges. In this section we briefly review some of the various research challenges currently facing CVEs.

Challenge 1: Scalability and interest management.

The requirement to support real-time interaction between large numbers of simultaneous participants distributed over a wide area network makes CVEs a challenging class of application, especially with regard to scale. The scalability of CVEs can refer to the graphical and behavioral complexity of virtual worlds and their contents, especially avatars, as well as the number of simultaneous participants that can be supported. Limitations on scalability arise from a variety of system bottlenecks. Large numbers of active participants generate high volumes of network traffic, especially movement updates and audio packets. Servers on the network may have to process this data, for instance, in computing a consistent copy of the world from many update messages. Even if the core network and server facilities can sustain a CVE, the "last mile" network connection to each

participant's machine can easily become a bottleneck, especially for domestic access via the Internet using dial-up modems. Finally, even if the information can be delivered to participants, their local computer must be able to process it and render the shared virtual world at a satisfactory quality while maintaining a sufficiently rapid response to the participants' movements and other actions.

Human perceptual and cognitive limitations provide a significant guide in developing responses to the problems of scale. By arranging the virtual environments so that each participant is not overloaded and sees and hears "enough" of the world but no more, the problems of scale can be diminished. In this case "enough" is defined in terms of interest in the world and its contents and may be constrained by features such as solid boundaries. For example, a participant need not receive audio packets from objects too far away to be heard or update messages from objects behind a nearby wall or that are deemed to be uninteresting. However, participants' interests will dynamically change as they explore a world. So the challenge for CVE developers is to design flexible and dynamic interest management schemes (see Sandeep and Zyda's *Networked Virtual Environments—Design and Implementation*, ACM Press, July 1999). The best-known examples define interest through the division of virtual space.

The NPSNET system (see the sidebar) divides the virtual world into fixed-sized hexagonal cells [10]. Participants send their information (movement updates) to their current local cell but can choose to receive information from potentially many cells that fall within their area of interest. The use of fixed-sized cells is appropriate to applications such as battle simulations where objects move with predictable speeds and trajectories.

The Scalable Platform for Large Interactive Networked Environments (SPLINE) system (see the sidebar) divides a virtual world into different-sized locales that are stitched together [1]. Each locale defines its own coordinate system and participants receive information from their current locale and its immediate neighbors. Using variable-sized locales provides additional flexibility in coping with less predictable objects.

MASSIVE-2 (see the sidebar) divides a world into regions whose boundaries can provide different degrees of permeability in different directions [6]. For example, a region boundary may attenuate audio information but allow visual information to pass unhindered. Regions can also provide aggregate representations of their contents and can move around the world to follow them. For example, a

region may move around with a crowd of participants, providing a low cost representation of them when seen at a distance. Figure 1 shows the difference between the use of fixed-sized cells and variable-sized bounded regions in managing interest.

Cells and regions exploit space, especially distance, to manage scale and reduce traffic and interaction effects. Higher-level spatial semantics can also be exploited to tackle the problems of scale. MASSIVE-2 provides users with a spatial model of interaction based on two distinct spatial fields—focus and nimbus. Participants' focus is a field whose shape describes their allocation of attention across space. Participant's nimbus is a similar field whose shape describes how they are projecting their information into space, and that may be used to model directional information sources or social behaviors such as shouting (large nimbus) or whispering (small, narrow nimbus). A participant's awareness of another is a function of his or her focus and the other's nimbus. Figure 2 shows how simple instantiations of focus and nimbus as discrete volumes can be used to negotiate different levels of mutual awareness, including two different ways of establishing partial awareness.

Challenge 2: Distributed architectures. CVEs support varying numbers of geographically distributed users and keep participants up to date with changes in the world and other forms of communication and interaction (such as network audio). Supporting these users represents a major challenge, and CVE systems vary significantly in the ways in which they handle the issues of distribution. Essentially, three basic architectures are exploited by CVEs:

• **Client/server:** Each participant's application communicates only with a common server program that is responsible for passing messages on to other clients as appropriate. MASSIVE, SPLINE, and Distributed Interactive Virtual Environment (DIVE), to a lesser extent, use servers to coordinate initial world joining. Client/server approaches are also the norm for public Internet CVEs, because the server can tailor its communication to match the network and machine capa-

Figure 1. Comparing fixed-sized cells (NPSNET) and variable-sized bounded regions (MASSIVE-2) for determining interest.

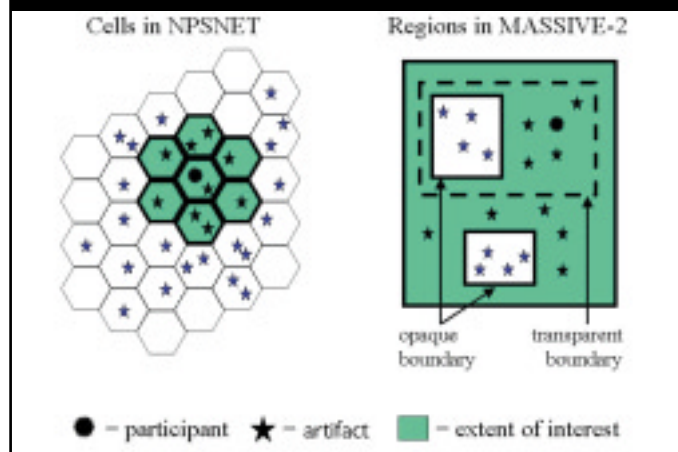
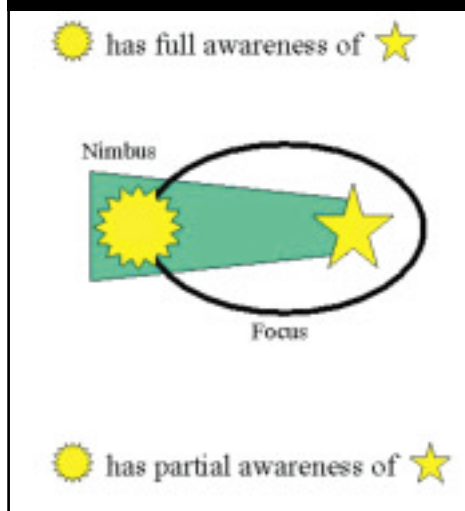


Figure 2. Negotiating awareness (interest) using focus and nimbus.



bilities of each client. For example, SPLINE supports dial-up users through specialized servers that connect them to the core system over a compressed and optimized application protocol.

- **Peer-to-peer unicast:** Each individual client program sends information directly to other client programs, as appropriate. Typically, this is the most bandwidth-intensive of the three approaches, but it avoids placing additional load on particular server machines and typically introduces lower network delays. This was used

for communication in MASSIVE-1, and is also used to deliver tailored video streams to observers in the Freewalk system [11]. It is also commonly used to provide initial world state information to new participants.

- **Peer-to-peer multicast:** Similar to peer-to-peer unicast (except the same information is sent simultaneously and directly to many other client programs) normally using a bandwidth-efficient network mechanism such as IP multicast, this approach is used exclusively in NPSNET, and is used for all updates in DIVE and MASSIVE-2. It is also used for audio in many systems, including SPLINE, even when a client/server approach is used for graphical data. Multicasting is not currently available on all networks or operating systems, and wide-area availability is particularly limited. Consequently, some systems now include

application-specific multicast bridging and proxying servers, which simplify use over wide-area and nonmulticast networks.

Although we have presented these as alternatives, a single CVE system can combine multiple approaches—for different media, different stages of participation or for different groups of participants. A key area of research is exploring new methods of

combining these architectures to effectively support a range of applications and media over mixed infrastructures (networks and computers). An example of this is SPLINE's combination of a peer-to-peer networked "core" with specialized support for low-bandwidth, client/server access. Other groups are working on distributing and coordinating multiple servers in a client/server framework.

Challenge 3: Migrating lessons from 2D inter-

Example CVEs

The growth of CVEs has seen a rapid rise in the development of systems that allow a number of users to share a 3D virtual space. These range from research prototypes to commercial products and include the following systems and applications

- DIVE: a CVE development environment; www.sics.se/dive

One of the first CVEs was the Distributed Interactive Virtual Environment (DIVE) system [7]. DIVE has been continuously developed since the early 1990s and is now accessible from a typical desktop PC and can also be used immersively (with a head-mounted display), or with projected displays. It allows distributed participants to be embodied in and freely navigate 3D graphical worlds and to communicate using a combination of real-time network audio, video, text chat, simple graphical gestures and meeting support tools such as shared whiteboards. Users can also manipulate objects in the virtual world (moving and turning objects or adding new objects). The DIVE "universe" comprises many disjoint virtual worlds, linked by portals, and world and object descriptions can be downloaded on demand from the Web. Objects' behaviors can be described in tcl and the environment provides a Java Interface. DIVE is freely available for research use from the Swedish Institute of Computer Science.

- NPSNET: simulating battles; movesinstitute.org/~npsnet
NPSNET-IV [10] is a 3D networked virtual environment system developed at the Computer Science Department of the U.S. Naval Postgraduate School. It is designed to support large-scale military training and simulation exercises. Each participant's machine can be configured to simulate a military vehicle or dismounted infantry personnel, and thousands of these interacting entities can be supported (see Figure 3). As a result, NPSNET-IV has made a

major contribution to the development of scaleable network architectures and interest management techniques.

- MASSIVE and inhabited television; www.crg.cs.nott.ac.uk/research/systems

The MASSIVE CVE was developed at the University of Nottingham and has been used to support an extended series of experimental trials of teleconferencing over wide area networks [6]. MASSIVE-2 (developed in 1996) and MASSIVE-3 (developed in 1999) have been used to create a variety of applications, especially in the area of public participation in online art and performance. This work has led to the idea of inhabited television as an application area for CVEs. This combines CVEs with broadcast television to create a new entertainment medium in which the public participates in TV shows that are staged within a shared virtual world [5]. The action from these shows is captured by virtual cameras within the world, mixed down and broadcast to more passive viewers as conventional television. Figure 4 presents screenshots from an inhabited television gameshow called "Out of this World" created for the "isea98:revolution" digital arts festival. During this event the public controlled two competing teams of avatars, led by two immersed actors; four virtual cameras were used to create a broadcast that was shown to a nearby theater audience.

- SPLINE: An environment for online communities; www.merl.com/projects/spline/index.html
Scalable Platform for Large Interactive Networked Environments (SPLINE) [1] was a CVE system developed by Mitsubishi Electric Research Laboratories. SPLINE was presented as a middleware set of facilities for supporting the creation of end-user applications. In particular, it defined a Java API that was

faces and CSCW. The dominant approach to collaboration in CVEs assumes each participant sees the same content, albeit from a different perspective. However, the experience of the CSCW community in designing collaborative 2D graphical user interfaces suggests this approach may actually hinder people's ability to collaborate. Early experiences of shared interface systems based on the principle of "What You See Is What I See" (WYSIWIS) led to a

reexamination of some of the principles of sharing and the need for public and private interfaces and different views on shared data. Some of these early lessons concerning the nature of cooperation have been migrated to CVEs and some systems now offer users "subjective" views on shared worlds [12]. These subjective views can reflect the different interests and roles that users inhabiting shared worlds may have. For example, participants inside a 3D

used to create and interface with virtual worlds. One of the most interesting features of SPLINE is that, rather than defining complete virtual worlds, it defines smaller regions (called "locales") that could be joined together to form complete worlds or net-

Alphaworld is one of the original online virtual communities and is now one of many virtual worlds supported by Active Worlds (www.activeworlds.com). In addition to providing the support common to most CVEs, Active Worlds virtual cityscapes can be built by

Figure 3. Images of a virtual soldier and a stretcher team from the NPSNET system (courtesy of U.S. Naval Postgraduate School).



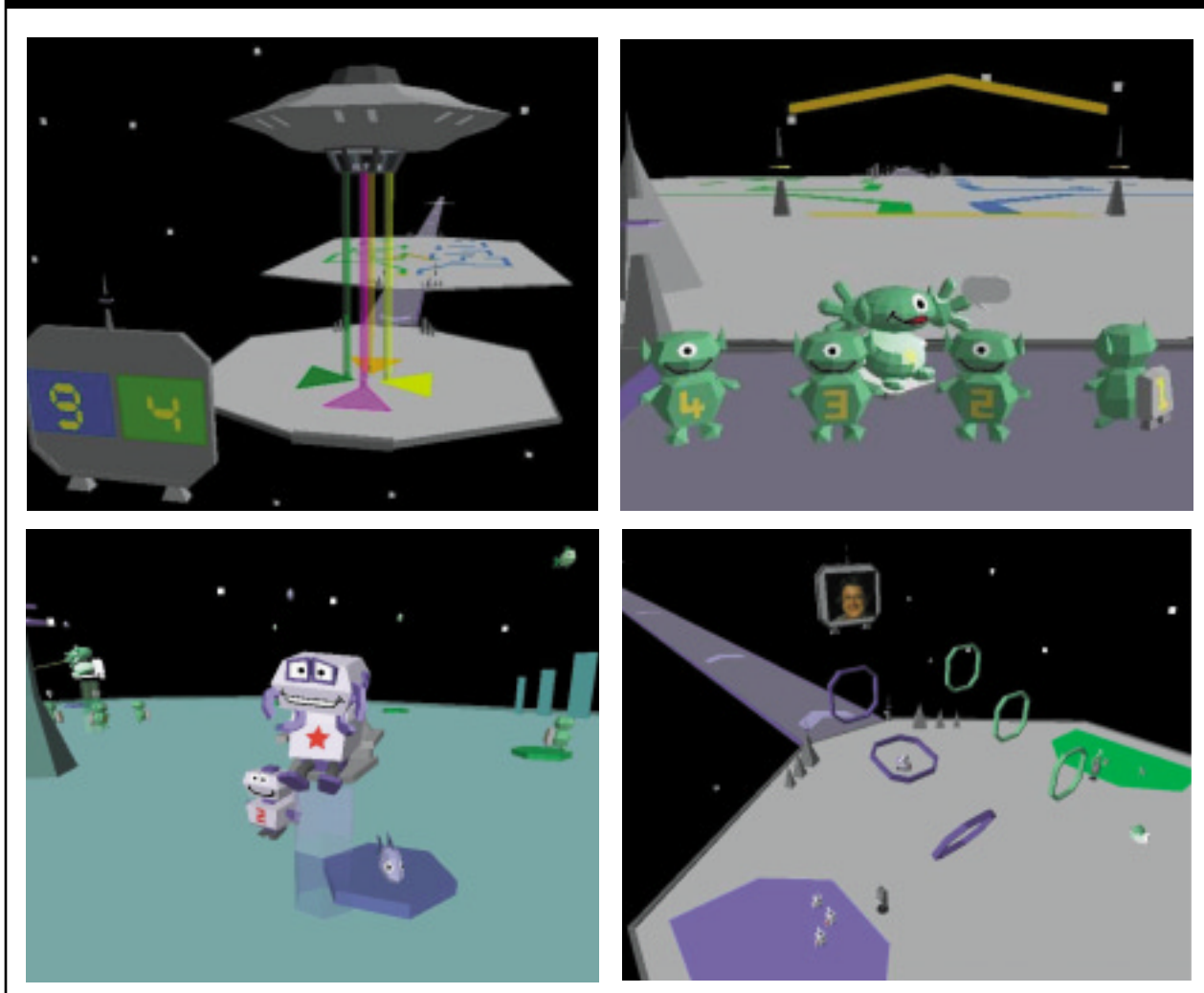
works of spaces. Locales could be joined together in very flexible ways, for example, to create a building that is larger on the inside than the outside.

- The Worlds of the Contact Consortium; www.ccon.org
The systems described here primarily exist in research establishments. Recently there has been a rapid growth in CVEs that are more readily accessible through the Internet and dial-up channels. Many of these are commercial platforms. Some of these systems have their roots in text-based virtual worlds (MUDS and MOOS are examples) and essentially offer graphical visualizations of such environments. Many still focus on using text chat to support communication although a number are starting to provide some audio facilities. Some popular and illustrative worlds are mentioned here.

its inhabitants in a Lego-like manner. The resulting environments are large and the hundreds of thousands of users of Active Worlds have filled a huge space with over twenty million objects.

Blaxxun Community Platform (www.blaxxun.com) is a modular client/server system that provides the full range of core functions for operating and visiting communities on the Web. In addition to the services that manage users, worlds, and chat communication, the system also offers a variety of modules that offer, for example, interfaces to database products and other server systems for advertising management or e-commerce, mechanisms for introducing pre-programmed community participants who can perform a variety of roles, and modules that help users behind firewalls to access the Blaxxun Community Services. **C**

Figure 4. Images of inhabited TV created using the MASSIVE-2 system. Top left: the set; top-right: the Alien team; bottom left: the robot team in action; bottom right: the quiz.



architectural model may see different overlays for wiring, plumbing, and networking; the virtual cameras used to capture the action in inhabited TV (see sidebar) might be visible to the performers, but not to other online participants or viewers.

The relationship between CVEs and other forms of shared interface has also been explored as part of a “space-versus-place debate” within the CSCW community. Those arguing for “space” propose it is independent movement within a shared coordinate system, combined with the representation of others’ positions through avatars, that underpins CVEs support for social interaction (for example, in allowing participants to point at objects). Those arguing for “place” maintain that social behavior is engendered by other important aspects of an environment beyond the provision of a shared coordinate system [8]. Although not exclusive, these points of view have led to different emphases in the design of online shared

environments; space leading to navigable CVEs with avatars, and place leading to more generalized abstractions suggesting conventions of conduct or that support ease of navigation.

Challenge 4: New kinds of human factors. In considering how user studies may inform the development of CVEs, it has become clear that new methods may be required to those typically used in evaluating single-user VR systems. Studies of single-user VR have tended to draw upon individual perceptual psychology for their orientation in exploring issues such as immersion, usability, and motion sickness. The use of CVEs to support cooperative work and social interaction presents new challenges for human factors: how can we understand the nature of social interaction within a CVE?

Recent studies have turned to a broader range of social scientific methods and to inform CVE design choices. Studies of early trials with the MASSIVE-1

system used ethnographic techniques to characterize how conversational mechanisms were exploited or adapted in shared virtual worlds. This work contributed to avatar design by showing that even graphically simple avatars could effectively represent participants on some occasions, while on others trust in avatars would break down when found to be unoccupied, for example when their owners were away attending to events within their local environments [3]. Further studies highlight difficulties with the use of humanoid-style avatars. Participants would assume others could see objects in the virtual world normally visible in the periphery of human vision when in fact their field of view was severely restricted by the CVE technology [9]. Perhaps avatars should more accurately convey their perceptual capabilities in the virtual world.

The Future of CVEs

Where next for CVEs? The next few years should see them move closer to mainstream commercialization. One route to this is the extension of current single-user VR applications to multiuser support in such areas as design, simulation, and training. Another engaging possibility is to exploit the entertainment and social nature of CVEs, extending today's Internet virtual communities with richer content and greater interactivity. Early demonstrations of inhabited TV (see the sidebar) hint at the potential of media fusion where the social and community strengths of CVEs are combined with the interactive narrative strengths of multimedia and the production strengths of broadcasting for the benefit of each. As CVEs reach commercial maturity, so too will issues of content and the nature of the social experience come to the fore and interdisciplinary collaborations with sociologists and artists become increasingly important.

Of course, many of the technical challenges outlined in this article persist. For example, scalability will continue to be a core challenge for CVEs. Moreover, new challenges will emerge, especially where the relationship between the virtual world and the everyday physical world is concerned. Early studies of social interaction in CVEs stressed the interdependence between virtual and physical space [3]. In turn, recent technical developments offer new ways of bridging the two. Tangible interfaces allow the manipulation of digital information through physical objects. Augmented reality has used VR technologies and techniques, but has turned them around so that information from the digital world is pulled into the physical world rather than the user

being immersed in the digital world. Ubiquitous, mobile, and wearable computing promises to make access to digital information universal and continual. A future research challenge concerns the relationship between the shared digital world, manifested through CVEs, and a shared physical world enhanced with digital information. What new techniques will enable a densely populated collaborative virtual universe to exist in parallel with the existing physical universe, and how will people manage their ongoing participation in both? ■

REFERENCES

1. Barrus, J.W., Waters, R.C. and Anderson, D.B. Locales: Supporting large multiuser virtual environments. *IEEE Comput. Graph. App.* 16, 6 (Nov. 1997), 50–57.
2. Bentley, R., Rodden, T., Sawyer, P., Sommerville, I., Hughes, J., Randall, D. and Shapiro, D. Ethnographically informed systems design for air traffic control. In *Proceedings of CSCW'92*, Toronto, (Nov. 1992), pp. 123–129.
3. Bowers, J., Pycock, J., and O'Brien J. Talk and embodiment in collaborative virtual environments. In *Proceedings of CHI'96*, Boston, ACM Press, 1996.
4. Damer, B. Demonstration and guided tours of virtual worlds on the Internet. In *CHI'97 Supplementary Proceedings*, Atlanta, ACM Press.
5. Greenhalgh, C.M., Benford, S.D., Taylor, I.M., Bowers, J.M., Walker, G. and Wyver, J. Creating a live broadcast from a virtual environment. In *Proceedings of SIGGRAPH'99*, Los Angeles, Aug., 1999.
6. Greenhalgh, C. *Large Scale Collaborative Virtual Environments*. C.J. van Rijsbergen, Ed. Springer-Verlag, Distinguished Dissertation series, June 1999.
7. Hagsand, O. Interactive Multiuser VEs in the DIVE system. *IEEE Multimedia* 3, 1, IEEE Computer Society. (Spring 1996), pp. 30–39.
8. Harrison, S. and Dourish, P. Re-place-ing space: The roles of place and space in collaborative systems. In *Proceedings of CSCW'96*, Boston, 1996, pp. 67–76.
9. Hindmarsh, J., Fraser, M., Heath, C., Benford, S. and Greenhalgh, C. Fragmented interaction: Establishing mutual orientation in virtual environments. In *Proceedings of CSCW'98*, Seattle (Nov. 1998).
10. Macedonia, M.R., Zyda, M.J., Pratt, D., Brutzman, R., Donald, P., and Barham, P.T. Exploiting reality with multicast groups: A network architecture for large-scale virtual environments. In *Proceedings IEEE Virtual Reality Annual International Symposium (VRAIS'95)*, North Carolina, March 1995.
11. Nakanishi, H., Yoshida, C., Nishimura, T. and Ishida, T. FreeWalk: A three-dimensional meeting place for communities. *Community Computing—Collaboration over Global Information Network*. T. Ishida, Ed. Wiley, 55–89.
12. Smith, G. Co-operative virtual environments: Lessons from 2D multi-user interfaces. In *Proceedings of CSCW'96*. Boston, November, 1996, pp. 390–398.

STEVE BENFORD (sdb@cs.nott.ac.uk) is a professor of Collaborative Computing at the School of Computer Science and IT at the University of Nottingham.

CHRIS GREENHALGH (cmg@cs.nott.ac.uk) is a reader in Interactive Systems at The School of Computer Science and IT at the University of Nottingham.

TOM RODDEN (tar@cs.nott.ac.uk) is a professor of Computing at The School of Computer Science and IT at the University of Nottingham.

JAMES PYCOCK (james.pycock@xrce.xerox.com) is a program manager for office applications research across Xerox Research and Technology and a project manager for emerging office environments at the Cambridge Lab of Xerox Research Centre, Europe.