

Revealing the Realities of Collaborative Virtual Reality

**Mike Fraser, Tony Glover, Ivan Vaghi,
Steve Benford, Chris Greenhalgh**

Communications Research Group
School of Computer Science
University of Nottingham
University Park
Nottingham NG7 2RD, U.K.
+44 115 951 4227(226/227/203/221)
{mcf, atg, irv, sdb, cmg}@cs.nott.ac.uk

Jon Hindmarsh, Christian Heath

Work, Interaction and Technology Research Group
The Management Centre
King's College London
Franklin-Wilkins Building
London SE1 8WA, U.K.
+44 20 7848 4194(/4496)
{christian.heath, jon.hindmarsh}@kcl.ac.uk

ABSTRACT

We look at differences between the experience of virtual environments and physical reality, and consider making the technical limitations which cause these differences ‘visible’, aiming to provide resources to enhance communication between users. Three causes of such discrepancies are considered to illustrate this idea: field-of-view; haptic feedback; and network delays. For each, we examine ways of revealing the limitations of the virtual world as resources to better understand the intricacies of system and co-user behaviour. These examples introduce a broader discussion of design issues involved in producing interfaces for day-to-day collaboration through virtual environments. Issues include: the application and activity undertaken through the virtual world; the ability to focus on the business at hand rather than the system in use; and extent of users’ familiarity with application and system.

Keywords

Interaction techniques, realism, desktop and immersive interfaces, haptic feedback, network delays.

INTRODUCTION

“The screen is a window through which one sees a virtual world. The challenge is to make that world look real, act real, sound real, feel real.”
(Sutherland) [17]

“Virtual Reality creates a new objective level of reality ... If you’re ever confused about which reality you’re in, you put your hands on your eyes and see if you’re wearing Eye-phones or not.”
(Lanier) [13]

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee.

CVE 2000, San Francisco, CA USA

© 2000 ACM 1-58113-303-0/00/09...\$5.00

The compelling statements that heralded the emergence of Virtual Reality (VR) have greatly influenced the development of interactive graphics technologies. Research has been driven by a commitment to build environments and interfaces through a complex array of technologies which aim to provide the illusion of, simulate, or are even indistinguishable from the physical world. Even the recent concerns with identifying factors that influence the sense of presence in virtual worlds (e.g. [16]) derive from early writings on simulating physical realities and physical presence ‘for all practical purposes’. However, although computer graphics technologies are developing at incredible rates, their abilities to live up to these visions are very much in debate. Perhaps no single aspect of VR systems development throws these visions into doubt more acutely than the inclusion of multiple users into such spaces: the development of Collaborative Virtual Environments (CVEs¹).

Although some of Lanier’s predictions may be achieved in the longer term, we suspect that in the medium term a range of obstacles will inhibit such an all-encompassing experience. It may be that, whilst the overriding interest in achieving a real sense of ‘presence’ continues, the shorter-term potential of CVEs to provide an effective communications tool is being undermined. Despite numerous advances, there remain severe difficulties in developing environments that satisfactorily support interaction between individuals within a virtual, yet convincing, world. In particular, the current haptic, tactile and olfactory interfaces are extremely crude. Even though

¹ It should be noted that, when using the term ‘CVE’ in this paper, the systems we consider are those providing 3D graphical, and often audio, media to multiple users, or ‘multi-user virtual reality’, if you will. In recent times, the designation ‘CVE’ has been applied to an increasingly broad range of systems which are not discussed here, such as MUDs and 2D graphical spaces.

the realism of graphics technologies is rapidly advancing, interfaces to ‘realistically’ capture movement and real-time expression for such graphical illustration lag far behind. In this light, it may be worth reconsidering whether the successful design of CVE applications rests upon our ability to accurately simulate ordinary realities and physical worlds.

Even if we can imagine a technology that allows us to completely perceive and interact with a realistic virtual world, more fundamental problems remain when we introduce collaborating participants into the equation. One of these is the speed of light. Users might be standing next to one another in the virtual environment, perhaps even working with the same virtual object, and yet may be physically located on opposite sides of the planet. In this case, the speed of light alone will result in a network delay of at least tens of milliseconds between them. In practice the delay will be far greater due to network hardware and current congestion problems. Such delays make it extremely difficult to create the illusion of both users instantaneously interacting with the object in a perfectly consistent way. Either the system only shows the state of the object when different users’ updates have been resolved, slowing interaction while one or more messages pass between them, or it shows each user the effects of their local interaction as it happens, with the risk that they perceive different (virtual) realities.

It is apparent that it will prove difficult to create CVEs that adequately simulate the characteristics of conventional worldly environments, at least for some time. However, some researchers argue that this need not be an insurmountable problem. For example, Dourish et al. [6] dispute the extent to which it is worthwhile attempting to simulate certain characteristics of co-present, physical environments in media space technologies. Slater et al. [16] and others make a similar point with regard to virtual environments – that CVEs also do not necessarily need to reflect or embody the characteristics of conventional environments to enable them to support particular forms of activity or interaction. It may be the case that only certain properties of worldliness need to be represented or embodied in CVEs. Usoh et al. [18] assert that an individual’s feeling of ‘presence’ is critical to their experience – that factors such as being able to see one’s own body and having access to a certain mode of navigation may provide the resources with which a person can act and interact in the world. The difficulty, however, is that the experience of the user, and his or her ability to satisfactorily engage in the virtual world, is treated as independent of the activities in which participants will engage. A sense or perception of realism may be entirely different from, and unable to stand in stead of, the ability to locally accomplish specific cooperative tasks.

In this paper we would like to suggest an alternative, more pragmatic, approach. CVEs provide a technology which will, however crudely at the present time, support interaction between remote participants, and enable them to engage and cooperate in particular activities. Rather than assume that we need to simulate conventional realities, treating discrepancies as somewhat embarrassing but unavoidable problems or ignorable side effects, this paper examines the ways in which the limitations and constraints of the CVEs might be shown. We examine the alternative that providing individuals with a sense of the constraints of the system can enhance collaboration in such environments. In this way we can enable individuals to see, identify and manage the characteristics of the CVE system (and the problems and difficulties which arise therein) by providing resources to aid the practical accomplishment of cooperative tasks. In other words, we wish to provide individuals with a sense of the particular realities – the constraints and illusions – of virtual environments as they emerge within their interaction.

Virtual environments – models, avatars, interfaces and so on – are often designed with realism (or, at least, the approximation to realism) as the primary concern; yet aspects of a system that are less than ‘real’ are often concealed from the casual observer or participant. In the next three sections, we examine some problems that arise when particular characteristics of CVEs are hidden from users. Each section also considers potential ways of revealing the nature of these characteristics to the participants. We shall then go on to describe what relevance to interaction through CVEs these resources may provide, and consider what design issues need to be addressed in providing such resources in helpful forms. The characteristics we have selected, and developed potential solutions for, derive from technical problems associated with action and interaction through virtual environments: field-of-view; haptic feedback; and network delays.

FIELD-OF-VIEW

The display technology most frequently associated with virtual environments are Head-Mounted Displays (HMDs). Such tracked devices allow designers to provide and reproduce ‘realistic’ human actions (such as intuitive gaze direction). The wide availability of desktop computers, however, combined with the cost and limitations of HMDs, means that much of the mundane, practical use of CVE technology is performed through desktop displays. In this section we shall consider a common limitation of both kinds of display technology for the experience of realism in CVEs – their ability to provide a wide, ‘human-like’ field-of-view.

Both kinds of technology usually provide limited horizontal space in which to render a user’s view of the

virtual world. For a HMD, the field-of-view is a physical characteristic of the headset, and usually limited to under 60 degrees. On desktop displays, the field-of-view can be varied, but is usually limited to 90 degrees or less. Rendering 3D graphics on such screens uses a perspective projection [5] onto a limited-sized display screen. This space limitation means that rendering software projecting desktop views must accept a trade-off between causing perspective distortions and field-of-view on the world. It has been noted that CVE applications often attempt to simulate real worlds or tasks. Yet highly distorted views of such virtual worlds would render realism unclear. Consequently the field-of-view provided, whether in a HMD or on a desktop display, is likely to fall far short of a ‘realistic’ or human-like value.

Problems caused by limited of field-of-view

In a study exploring ‘object-focused interaction’ within CVEs, Hindmarsh et al. examined tasks that involved viewing, manipulating and making reference to shared objects through a virtual world [11]. One aspect of this study discussed how a user’s view of another’s action, a relevant artefact under discussion, and the relationship between the two, was often ‘fragmented’ by the limited field-of-view provided by a desktop interface to a CVE application (constructed with the MASSIVE-2 system [2]). This field-of-view caused users to encounter problems in discerning what fractions of the virtual environment other users could see when discussing and referring to features of the virtual world. In particular, using a ‘pseudo-humanoid’ avatar (an attempt to faithfully simulate reality) encouraged incorrect assumptions to be made about the field-of-view of the other. Evidence pointed to users assuming that a ‘human-like’ avatar equated to the other having a ‘human-like’ field-of-view. Referring to the scene that is shown in Figure 1, we contrive an example that is characteristic of the situations observed in this work – situations revealing how the field-of-view can make it difficult to ‘see what another sees’.



Figure 1: Can one user see what the other can see?

Figure 1 shows two participants interacting in a virtual world. The first user (whose avatar is to the right of the picture) wishes to point out the stereo to the second user (behind whose avatar we are viewing the scene). The first produces the utterance “you see the stereo?”. From the second avatar’s general orientation, the first user might infer that their own avatar, any pointing gesture they might produce, and the stereo artefact in question, are visually available to the second user. This would probably be the case for an equivalent situation in the real world, where both parties have human perceptual resources. Yet, as we discuss below, a limited field-of-view may cause these resources to be unavailable.

Revealing limitations in the field-of-view

Various methods might be employed to overcome this problem. One might be the use of wide-screen projections such as CAVEs, as these interfaces would allow a wide, undistorted field-of-view. However constraints such as resolution, cost, storage space and health and safety concerns render the potential benefits somewhat unclear in terms of finding a practical solution.

Fraser et al. [7] evaluate different solutions which may be applied to desktop CVE interfaces. These include explicitly displaying a user’s actual field-of-view by showing their view frustum as a graphical object. Other users might then be able to discern more explicitly what the other can see, rather than inferring this from, for example, what the other might see in a similar situation in the real world. Misunderstandings in interaction and resulting confusion might then be avoided.



Figure 2: Revealing the reality of field-of-view

Figure 2 shows the same scene as Figure 1, but with the second user’s field-of-view made explicitly visible as a semi-transparent volume, causing a ‘lighting’ effect. The first user may now discern the other’s rendered view, and thus be able to determine a course of action more easily. Pointing towards the stereo might indeed be problematic because the second user can see neither the stereo nor the first user’s avatar. Therefore, the first user could elaborate

on his/her actions over an audio channel, to compensate for the other's lack of visual cues, or perhaps move their avatar into the other's view before pointing (as shown in Figure 3).



Figure 3 – Revealing the field-of-view allows a pointing gesture to be designed so that another can see it

This example begins to show how making the limitations of a virtual field-of-view visible might enable users to anticipate problems and co-ordinate coping schemes more effectively.

HAPTIC FEEDBACK

Research is increasingly looking at how to provide users with haptic and tactile feedback – the sense of physical force and touch. Applications requiring such senses for training are some of the main drivers for producing haptic feedback in VR systems – touch is key to, for example, training medical skills [14]. Research is often focussed around single-user graphical systems, and developments have typically concentrated on the technical aspects of tactile augmentation or ‘force feedback’. In [12], an evaluation of tactile feedback techniques, Hoffman claims, “physically touching virtual objects using tactile augmentation enhances the realism of virtual environments”. However, discrepancies between the real and the virtual are also noted in this work – “the physical textures of the real objects (e.g. the fuzzy feel of a peach) are hard to reproduce in computer simulations.” [ibid.]. In general, haptic and tactile interfaces are in their infancy and technology is very far from any whole body realistic experience. Indeed, the typical experience of virtual environments today is unlikely to include any haptic feedback at all.

Problems caused by a lack of haptic feedback

There exist a myriad of situations in which the sensation of force will play a key role in collaboration. Again, consider the case where two users are working with virtual objects. In the real world, we feel a physical sensation when someone takes an object from our hand, even if we are not looking at it at the time, yet in current CVEs this is not the

case. Given the usual limited interface’s field-of-view, it is quite possible that a user does not see an object being taken from them. The result might be that, lacking the haptic cues for noticing this incursion, the user is baffled upon referring back to the object in their possession to discover that it is not there. In such situations, previous studies have shown that the incipient nature of technologies such as CVEs often causes users to assign system failure (rather than social and behavioural explanations) to unanticipated phenomena (e.g. [3, 19]). This leads us to propose that some more explicit representation should be used – to convey to the user both that an object is under competing forces for its possession; and that an object has been taken or ‘snatched’.

Compensating for a lack of haptic feedback

We might compensate for a lack of haptic and tactile feedback when grasping objects by portraying these forces in other media. For example, we might use audio to convey the nature and effects of the forces being applied. Another possibility would be to display users’ contesting forces by representing the resultant force direction and strength visually – e.g. as an arrow of particular size or length.

To demonstrate the display of lack of haptic feedback, we have extended the interface to MASSIVE-2 to allow multiple users to manipulate virtual objects in a variety of ways. The interface enables immersed users wearing HMDs to manipulate objects with two hands. Each hand can grasp and then pull or push an object. Joint manipulation can take the form of either a cooperative (e.g. combining strength to push) or competitive (e.g. pulling objects from one another) effort.

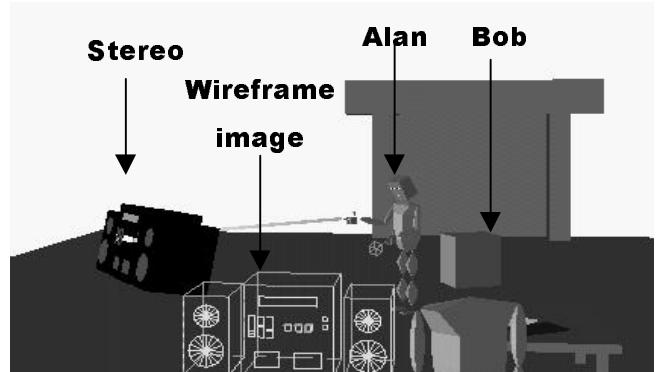


Figure 4: One user snatches an object from another

Our interface displays metaphors for the forces being applied to objects as well as the results of these forces. These metaphors are shown in both visual and auditory media.

Figure 4 shows an example of the result of snatching an object. Bob has previously been holding the stereo, but Alan has taken it, having a greater strength associated with his right arm. The line between Alan and the stereo indicates the association between the two – the fact that he

is in possession of it (users grasp objects by spearing them with the stick and pressing a button on a 3D mouse). Bob is left holding a wire frame representation of the stereo, showing that the object that was there has been snatched away and is now elsewhere. This wireframe image persists until Bob releases his grip, thus ensuring a visible representation until such time as he perceives the action.

The removal of the object from their hand is also accompanied by a sound. The reasons for this become clear when we consider the possibility of a user to have no immediate visual cues – audio representations compensate not only for the lack of haptic cues, but also for the potential lack of immediate visibility of another's action due to the small field-of-view.

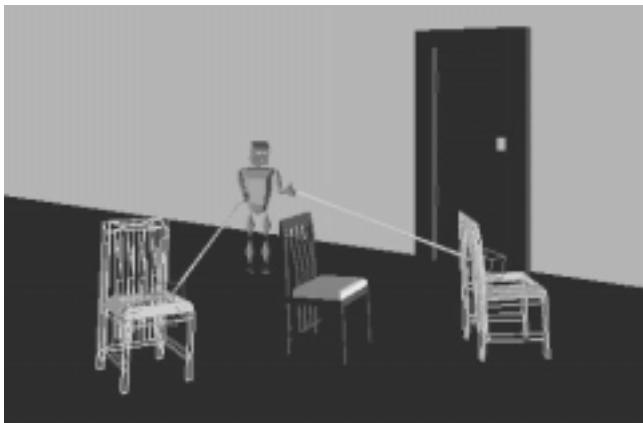


Figure 5: Competing forces on an object made visible

Figure 5 shows a second example. In this case the user is tugging a chair between their two hands to demonstrate a way of representing competing forces. Unlike the stereo, that is programmed to transfer from one grasp to another, the chair is programmed such that, when user or users tug it between two grasping hands one of them must eventually 'win' (i.e. the other hand lets go). The forces involved are shown by visually displaying a wire frame chair in each hand, representing the individual effect of each force, with a solid image being positioned at the average of the two. Therefore, considering a multiple user scenario, each can see that they have grabbed the object and are having a forcible effect upon it, but do not yet have sole possession of it. In this way, users may employ the visual cues to understand the kinds of collaborative object manipulation that they would normally feel through their hands.

NETWORK DELAYS

Any synchronous, collaborative, system has to address the reality of current rates of network communication. Delays in transmission of data in such communication technologies not only result in frustration for users, but can disrupt the very practices upon which face-to-face interaction rests. For example, when considering audio-video based communications technologies, Ruhleder and

Jordan found that delays in transmission are disruptive of turn-taking in the organisation of talk [15]. Such problems apply equally as critically, albeit in potentially different ways, to the disruption of sense-making and organisation in virtual environments [19]. We might also infer that relevant visible features of interaction – such as bodily comportment, gesture, gaze direction and so on – produced and, made sense of within a local, virtual environment context, may follow suit.

Unfortunately, complex graphics and audio media, commonly included in these high-end CVE systems, are particularly demanding of bandwidth. A common technique employed in CVE systems is to replicate, partially or totally, the 'objective' virtual world on each user's computer and to transmit updates between them. There are two reasons for this. First, a rapid real-time response is seen as being essential to the experience of CVEs. This can be achieved by allowing each user to update their local copy of the world, immediately seeing the results, and obtaining feedback on their operations and actions projected into the environment. Sending updates between them then synchronizes the different versions of the world. Second, this approach can reduce the amount of communication required between clients. Only changes in the state of the world need to be propagated between individual users, and thus the necessary application bandwidth is reduced.

Problems caused by network delays

If users are situated in disparate geographical locations, messages will often need to be sent across wide-area networks, causing updates to take some time to reach their destinations – delays resulting in a number of problems for social interaction. Users might apply incompatible updates to their local copies of the world that don't become apparent until later on. For example, they might both think that they have managed to move the same object. When the system finally resolves the conflict it will have to correct one of the copies to match an 'objective' state, causing a sudden jump in that user's view of the world. A second problem concerns the ordering of events. Even where multiple updates do not conflict, each user will receive each update at a different time and in a different order depending on where they are in the network. As a result, they may see events in different orders – or at least with differing timing – and constitute a different sense of local causality.

Vaghi et al. have experimentally investigated how network delay can affect users' behavior through trials in which two players engaged in a virtual ball game [19]. The first user participated on the computer hosting the virtual world and therefore suffered minimal delay, whilst the second user was placed on a second client subjected to artificially increasing network delays. As network delay increased,

divergence between the second user's perception and the first user's 'objective' context of interaction (seen on the game server) produced increasingly severe irregularities. Initially the second user saw the ball start to bounce away at slightly odd angles compared to the collision perceived by the first user. Later in the game the ball was seen to pass through the user's embodiment, occasionally bouncing back at some later time, having registered the collision with some delay. There were misunderstandings in conversation between the two players, since they could not always agree on statements regarding space and time relationships between events and objects.

Users reacted differently to these inconsistencies. The more successful players ascribed the inconsistencies to network delays (they had been warned about the possibility beforehand) and developed strategies for coping with them, for example waiting as long as possible before hitting the ball and trying to slow the pace of the game. They also reported that they had used a variety of strategies to assess how much delay was present.

However, not all users noticed the presence of delays. Instead, some formed other hypotheses to account for their problems: the ball was bouncing over them; the collision detection mechanism had stopped working correctly; they had an inherent lack of skill in using the application. Other subjects were able to identify the delay for what it was, but they did not always understand its consequences. In these cases, it seems that trust can be lost in the system, and in the reciprocal perspectives we assume in everyday interaction, because the presence of delays is hidden from users. In the face of increasing inconsistency it can become more difficult to understand what another perceives as happening.

Visibly presenting network delays

Transparently hiding the presence and nature of network delays causes problems for some participants interacting through the virtual world. We propose that we might help users to adopt coping strategies by explicitly showing delays within the world. Users may then view behaviour as being the product of 'delay induced phenomena' rather than for reasons such as those mentioned above.

We have implemented an example of this approach in the MASSIVE-2 system. The amount of delay associated with an object or a user is displayed visually. The system estimates the maximum difference between the 'objective' position of a user's avatar, and where another might perceive it to be. This disparity is based on two factors: measures of current network delay times between the users; and speed, direction and predictability of movement. For example, consider a user moving their avatar at 5 m/s over a network delaying updates to our view of the world by 420 milliseconds – a value representative of a modem connected to a relatively congested, low bandwidth wide-

area network. The resulting uncertainty of position is a volume of radius 2.1 metres, as shown in the representations in Figure 6. The avatar displays two representations of this example delay. The volume around the perceived avatar position shows potential uncertainty in their spatial position, and a three-dimensional sliding widget hovering above the avatar indicates current levels of network delay for the associated user.

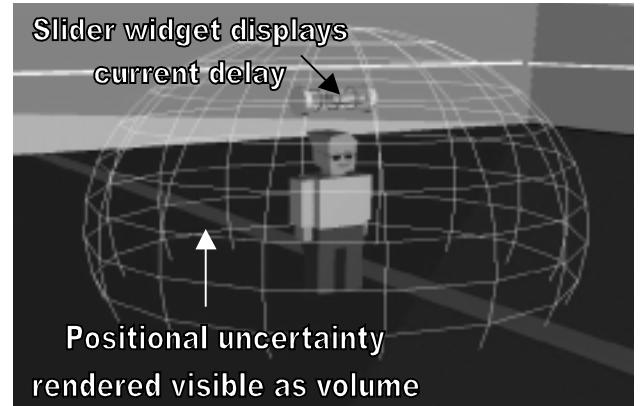


Figure 6: Network delay shown as uncertainty volume and widget

The potentially inconsistent position of an avatar renders its association with an object such as the ball in the game unclear; yet through this extended representation it is likely that influence on objects within the volume may be seen as related to the correct avatar. This method of showing delay might prove useful for accomplishing actions in a scenario such as the ball game, enabling other's understanding of the relationship between object and avatar. The widget above the avatar displays the current delay levels calculated for the associated user. This representation may be useful for coping with delay in more focused and intricate encounters.

Sudden increases in another's delay can be noticed in the variation of these indicators. Such visual cues could allow participants to assign tasks to other (less delayed) users. We might also expect users to understand a participant's sudden inability to perform or continue with the enterprise at hand.

DISCUSSION

We have attempted to show an alternative approach to designing virtual environments, by shifting focus away from the general trend of 'hiding' their properties and towards providing an enabling technology for communication.

Our three different examples have displayed how limitations in virtual reality might cause problems for interaction when attempting to faithfully display the pretence of realism. Limitations in field-of-view renders users' understanding of what others can see problematic. Lack of haptic feedback removes natural cues by which

users may understand collaborative object manipulations. Interacting with local copies of a world in the presence of network delays can result in apparent inconsistencies in social and world behaviour. In each case we have demonstrated specific examples of how these limitations might be more explicitly presented to users in order to help them achieve practical communication: showing a user's view frustum; leaving visual traces of objects that have been snatched away; showing a level uncertainty in an object's position.

These representations have been described in the context of specific instances, implemented within a specific CVE system. Thus it may be noted that there is a question over their generic applicability. Indeed, it might be argued that they themselves are incompatible – the use of both an uncertainty volume in combination with visually displaying a user's field-of-view, for example, may produce so much occlusion of the avatar that interaction becomes significantly impaired. It is clear that a design to show the field-of-view in another when network delay is present and significant would need to be revised and sensitised in particular ways. Thus, in the remainder of this paper, we go on to describe the kinds of design issues which will need to be considered in selectively and sensibly revealing the actual reality of virtual collaboration.

Taking into account the specificity of the instances and systems we have discussed thus far, we would argue that the three problems highlighted, and our more general approach to addressing those problems, are highly generic, and important considerations for CVE systems design. In any case, we have not, by any means, attempted to present an exhaustive taxonomy of such problems; there are many others that could be pursued. An interesting fourth instance is that discussed in [4]. Bowers et al. reveal the ways in which activity in the real world can influence how actions and responses in the virtual world are perceived and oriented to by virtual co-participants. Users constitute their environment, not just of the mutual resources available in the virtual world, but of all potential incursions from the real world. Moreover, there are similar problems with wires and cabling attached to immersive tracking technologies that intrude upon the movement of users, and force attention away from the virtual 'business at hand' [18]. One (simple, early) technique to address this 'reality' of CVE interaction, has been the provision of avatars which can 'sleep' in the virtual world, displaying to virtual co-users that one is distracted by events within the physical environment [1].

Revealing the practical reality of CVEs provides participants with a sense of the constraints and limitations of the system; it renders visible particular characteristics as resources with which users may recognise, identify, and

use to manage difficulties. Yet configuring CVE systems in this way raises certain issues that have to be addressed to enable us to provide satisfactory and 'user sensitive' solutions. In this section, we discuss some of these issues to clarify the ways in which such pragmatic solutions might be approached.

Each of the examples we have noted proposes particular solutions to how we might reveal the underlying characteristics of the CVE system at the interface. For instance, the co-participant's field-of-view is portrayed through a semi-transparent pyramid, or the removal of an object leaves a trace in the individual's hands. Of course, advances in the technology may alleviate some of these problems. Wide screen immersive displays such as CAVEs may become more widely available and haptic feedback technologies may become more sophisticated. However, these problems seem unlikely to disappear in the foreseeable future.

Defining a solution involves presenting particular information to users in particular ways, and presupposes that the information presented is relevant to the particular individual, occasion, and activity in which he or she is engaged. We are currently undertaking empirical studies to examine how individuals orient to and use the particular solutions that we have built. In advance of this, however, and drawing on previous experiences with studies of interaction through CVEs and other communication technologies, the remaining sections discuss what are felt to be three key aspects of designing representations when revealing the reality of CVEs.

Application and activity

Our earlier research suggests that a critical consideration as to how information concerning the system might be presented to the participants, is related to the character of the application domains (e.g. control room, meetings, performance and entertainment) and activities (e.g. surveillance, discussing documents, acting) in which they are engaged.

At a gross level, different applications may require or suggest that particular information about the system is critical, whereas for other applications it is peripheral. For example, the presentation of information about objects that have been 'snatched' may only be relevant to applications where specific conflicts between users are likely or necessary. On the other hand, the display of field-of-view may be initially required in most populated virtual environments where intricate or focused collaborative work may occur.

More likely, however, is that the demands of certain activities will be common to a number of application domains. Different sorts of activities (e.g. referencing objects, following the movement of objects across a virtual scene, using a whiteboard, etc.) will require or demand the

availability of different types of information. For example, displays of network delays may be more useful when engaged in an activity like referencing objects, but less important when listening to a lecture or monologue. Therefore, the system could be designed to be sensitive to the activity underway and respond to changes between activities. The display of system constraints could then be altered as the activities develop and emerge.

Involvement and alienation

There is a long-standing tradition in studies of interaction through computer technologies to recognise the importance of the tacit in the organisation of human activities. Highly complex actions and interactions rest upon a ‘seen but unnoticed’ body of practices and procedures. Indeed a field of research, ‘workplace studies’, is directed towards the investigation and explication of these fundamental social practices. A central feature of these procedures is that they are taken for granted and ‘glossed’; that is, in doing what we do, we rarely attend to the ways in which we do it. When we explicitly attend to such practices, it changes the nature of our involvement, and can undermine our very participation in complex, and even mundane, activities [8, 9, 11].

Crudely revealing the characteristics and constraints of a VR system in the course of participants’ interaction within the virtual environment could serve to undermine the free flow, flexible and collaborative accomplishment of activities. Indeed, it may be that the illustrative representations we have described in this paper might achieve such an undesired objective – it is hard to say without empirical investigation. Of issue here, however, is that poor use of representation may re-direct users’ involvement away from the activity itself, and to the ways in which the system is or is not supporting them. In other words, the method through which the characteristics of the system are revealed may encourage participants to reflect upon how an activity is being organised and accomplished. Their attention may be turned away from the business at hand and towards their mode and method of participation.

For example, in work that attempts to furnish VR users with accurate representations of particular gestures (e.g. [10]), explicit interface actions are required to initiate those gestures. However, the impact and import of those gestures is often tied to the interactional or temporal position in which they are produced. By the time an appropriate button is selected and depressed, and the avatar is animated on another’s display, the gesture may be seen as a comment on a totally different turn at talk or activity. Provision of so-called such ‘symbolic’ gestures is one example of the ways in which users might become involved in accomplishing simple activities. In this case, we might see disruption of the flow of a collaborative task

whilst users become engaged in attempting to categorise their gestural expression into the available possibilities.

Competence and familiarity

The ways in which participants become engaged in virtual tasks bear upon a related issue – that of their familiarity with the system. For someone who is new to a technology, and unaware of its facilities and limitations, a configuration revealing certain characteristics may well be very useful. Indeed, such information may be critical to becoming competent with the system. These resources might enable a user to engage in the activities without attending to the problems, complexities, limitations, and idiosyncrasies of the system and how it must be used to accomplish certain actions. Moreover, the ability to understand the system more quickly as a novice could enable a user to enter an established system or application with a relatively shallow ‘learning curve’.

In the case at hand therefore, we can begin to consider how the competence and familiarity of the participants, is relevant to how, and indeed whether, particular system characteristics are displayed. It is likely, through use of the system, that individuals will become increasingly familiar with its limitations and will also develop a body of practices and procedures with others in order to address those limitations and accomplish particular tasks (cf. [6]). As familiarity develops, therefore, the presentation of the system’s limitations may become frustrating, or worse, threaten the activities at hand. Alternatively changing the appearance of the CVE and removing resources that have been relied upon could be equally disruptive. Moreover, when participants of different degrees of competence and familiarity with the system are collaborating, further issues arise with regard to which resources should be provided, to whom, and when. Therefore, the ways participants may usefully and easily configure information concerning the system for themselves and others must be explored.

Furthermore, it is not difficult to imagine situations where it might be useful for representations of technical aspects of the CVE system to become more obvious. These are likely to involve unusual occurrences or difficulties. For example, if network delays increase above a ‘bearable’ threshold, users might wish to be aided by representations; yet occasions involving more acceptable values could remain transparent. Similarly, if a group of CVE users are *au fait* with the fact that all participants have limited desktop-based fields-of-view, then representations may become less necessary. However, consider the case of an immersed CAVE user being introduced into such a situation. The potentially confusing impact of an unusual, yet invisible, field-of-view on interaction might be attenuated through more explicit representation of its extent. Thus, the challenge is to provide representations of actions and orientations that fit within the smooth flow of

interaction, diverting attention away from the business at hand only when users' understanding and coping strategies may be likely to fail.

FUTURE WORK AND CONCLUSIONS

We are currently coordinating a program of research concerned with the issues, themes and suggestions raised within this paper. In particular, we feel it will be necessary to explore the relative success of displaying field-of-view, haptic feedback and network delays in the ways described. We intend to investigate how individuals, in interaction with others, use and orient to these resources, and examine how appropriate the different kinds of representation may be in certain situations. Finally, we hope to initiate a longitudinal study to assess the influence of an individual's familiarity with a particular system and activity.

This paper has discussed a pragmatic approach to CVEs; one that encourages content over style. If virtual environments are to provide effective media for interaction then the practicalities of collaborative systems as they exist must be considered. This suggests that empirical studies of interaction in and through CVEs will be consequential and key to unlocking the variety of ways in which they will eventually be used to support communication.

REFERENCES

1. Benford, S. D., Bowers, J. M., Fahlén, L. E., Greenhalgh, C. M. and Snowdon, D. N., User Embodiment in Collaborative Virtual Environments, in *Proc. CHI'95*, Denver, CO, USA, pp. 242-249, 1995, ACM press.
2. Benford, S., Greenhalgh, C. and Lloyd, D., Crowded Collaborative Virtual Environments, in *Proc. CHI'97*, Atlanta, GA, USA, pp. 59-66, 1997, ACM Press.
3. Bowers, J., Pycock, J. and O'Brien, J., Talk and Embodiment in Collaborative Virtual Environments, in *Proc. CHI'96*, Vancouver, Canada, pp. 58-65, 1996, ACM Press.
4. Bowers, J., O'Brien, J. and Pycock, J., Practically Accomplishing Immersion: Cooperation in and for Virtual Environments, in *Proc. CSCW'96*, Cambridge, MA, USA, pp. 380-389, 1996, ACM Press.
5. Carlbom, I. and Paciorek, J., Planar Geometric Projections and Viewing Transformations, *Computing Surveys*, 10(4), December 1978, pp. 465-502.
6. Dourish, P., Adler, A., Bellotti, V. and Henderson, A. Your Place or Mine? Learning from Long-Term Use of Audio-Video Communication, *Computer-Supported Co-operative Work*, 5(1), pp. 33-62, 1996, Kluwer.
7. Fraser, M., Benford, S., Hindmarsh, J. and Heath, C., Supporting Awareness and Interaction through Collaborative Virtual Interfaces, in *Proc. UIST'99*, Asheville, NC, USA, pp. 27-36, 1999, ACM Press.
8. Garfinkel, H. *Studies in Ethnomethodology*, Polity Press, Cambridge, 1967.
9. Goffman, E. *Interaction Ritual: Essays on face-to-face behaviour*. New York: Pantheon Books, 1967.
10. Guye-Vuillème, A., Capin, T. K., Pandzic, I. S., Thalmann, N. M. and Thalmann, D. Nonverbal communication interface for collaborative virtual environments, *Virtual Reality Research, Development and Applications*, 4(1), pp 49-59, 1999.
11. Hindmarsh, J., Fraser, M., Heath, C., Benford, S. and Greenhalgh, C., Fragmented interaction: establishing mutual orientation in virtual environments, in *Proc. CSCW'98*, Seattle, WA, USA, pp. 217-226, 1998, ACM Press.
12. Hoffman, H. G., Physically Touching Virtual Objects Using Tactile Augmentation Enhances the Realism of Virtual Environments, in *Proc. VRAIS '98*, Atlanta, GA, USA, pp. 59-63, 1998, IEEE Press.
13. Lanier, J. Virtual Reality is the Telephone of the Future, in Rucker, R. (ed.) *Mondo 2000: A user's guide to the new edge*, New York: Harper Collins.
14. Logan, I., Ward, J., Wills, D., Sherman, K. and Mohsen A., Arthroscopic Training in a Virtual Environment: Is it a reality?, in *Proc. MED TEC '97*, Vienna, VA, USA, 1997.
15. Ruhleder, K. and Jordan, B., Meaning-Making Across Remote Sites: How Delays in Transmission Affect Interaction, in *Proc. ECSCW'99*, Copenhagen, Denmark, pp. 411-429, 1999, Kluwer.
16. Slater, M., Usoh, M. and Steed, A., Depth of Presence in Virtual Environments, *Presence: Teleoperators and Virtual Environments*, 3(2), pp. 130-144, MIT Press.
17. Sutherland, I. The ultimate display, *Proc. IFIP Congress*, 2, 1965.
18. Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M. and Brooks, F. P. Jr., Walking > Walking-in-Place > Flying, in Virtual Environments, in *Proc. SIGGRAPH'99*, Los Angeles, CA, USA, pp. 359-364, 1999, ACM Press.
19. Vaghi, I., Greenhalgh, C. and Benford, S., Coping with Inconsistency due to Network Delays in CVEs, in *Proc. VRST'99*, London, UK, pp.42-49 1999, ACM Press.