

Supporting Social Human Communication between Distributed Walk-in Displays

David Roberts, Robin Wolff, Oliver Otto
The Centre for Virtual Environments,
University of Salford,
Manchester, M5 4WT, UK
d.j.roberts@salford.ac.uk
r.wolff@salford.ac.uk
o.otto@salford.ac.uk

Dieter Kranzlmüller,
Christoph Anthes
Institute of Graphics and
Parallel Processing,
Joh. Kepler University Linz,
A-4040 Linz, Austria
dk@gup.jku.at
canthes@gup.uni-linz.ac.at

Anthony Steed
Dep. of Computer Science,
University College London,
London, WC1E 6BT, UK
A.Steed@cs.ucl.ac.uk

ABSTRACT

Future teleconferencing may enhance communication between remote people by supporting non-verbal communication within an unconstrained space where people can move around and share the manipulation of artefacts. By linking walk-in displays with a Collaborative Virtual Environment (CVE) platform we are able to physically situate a distributed team in a spatially organised social and information context. We have found this to demonstrate unprecedented naturalness in the use of space and body during non-verbal communication and interaction with objects.

However, relatively little is known about how people interact through this technology, especially while sharing the manipulation of objects. We observed people engaged in such a task while geographically separated across national boundaries. Our analysis is organised into collaborative scenarios, that each requires a distinct balance of social human communication with consistent shared manipulation of objects.

Observational results suggest that walk-in displays do not suffer from some of the important drawbacks of other displays. Previous trials have shown that supporting natural non-verbal communication, along with responsive and consistent shared object manipulation, is hard to achieve. To better understand this problem, we take a close look at how the scenario impacts on the characteristics of event traffic. We conclude by suggesting how various strategies might reduce the consistency problem for particular scenarios.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces - Interaction styles;

H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces - Computer-supported cooperative work;

J.4 [Computer Applications] Social and Behavioral Sciences - Psychology, Sociology;

Permission to make digital or hard copies of all or part 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, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

VRST'04, November 10-12, 2004, Hong Kong.
Copyright 2004 ACM 1-58113-907-1/04/0011...\$5.00.

General Terms

Human Factors, Performance, Experimentation, Measurement, Design

Keywords

CVE, event traffic, consistency control, human interaction

1. INTRODUCTION

Walk-in displays (e.g. CAVEs [5]) physically place the user in an intuitively interactive information context. Linking such displays with a Collaborative Virtual Environment (CVE) additionally situates people in an intuitively social context. This technology lends itself well to Social Human Communication (SHC) by supporting its four primary elements: verbal and non-verbal communication, references to objects and references to the environment [3, 12]. In the real world, people combine all of these elements, often without realising it. A goal of this study is to verify if the same is true for distributed users interacting through this technology. Another is to see if nuances are maintained across the virtual link. Previous studies [13, 17, 20] have shown both an increased team performance and an increased feeling of co-presence between remote users of walk-in displays compared to those of desktop displays. The extent to which people naturally use SHC and to which nuances are interpretable might play a more or less part in this topic.

From a technical viewpoint, supporting both intentional and unintentional body movement within a distributed system requires a large volume of update-event traffic across the network. This may result in reductions in responsiveness and consistency of interactions with objects. This can be addressed with either of two options: developing consistency mechanisms within CVE systems or laborious fine-tuning of event traffic during application development. In order to attempt either, it is necessary to have an understanding of how SHC may be used within given collaborative scenarios and to know about the characteristics of event traffic generated within the given forms.

This paper attempts to identify the communication requirements of various event types and how often events of each type occur in various collaborative scenarios. An experimental task is undertaken between two walk-in displays, one in the UK and the other in Austria. The aim is to identify the levels of each element of SHC used in each scenario, describe the requirements on consistency and detail resultant event traffic. Our findings draw some basic requirements on consistency for supporting SHC within a variety of scenarios.

The paper is organised as follows. The next section clarifies some of the terminology used throughout the document and gives a short overview about CVE technology and related work. In section 3 we describe our experimentation and in section 4 the results of it organised in four distinct collaboration scenarios and the influence of human and system factors. Section 5 concludes.

2. SOCIAL HUMAN COMMUNICATION

Social Human Communication (SHC) encompasses a taxonomy of interaction that includes verbal and non-verbal communication, and the role of objects and the environment on communication [3, 12]. Speech, and other sounds, such as whistling and booing, belong to verbal communication, whereas body language, including posture and gesture, belong to non-verbal communication. Verbal and non-verbal communication are often inextricably linked through nuances such as lip-synch, clapping and unintentional gesturing and posture changes while speaking.

The subject of communication is not always abstract and often relates to our surroundings and artefacts within it. Both provide a context for understanding. We may discuss our surroundings or an object through both verbal and non-verbal communication but in addition we can move around the environment and manipulate objects within it.

A nuance might arise from the synchronisation of concurrent elements of SHC. For example, a user might point to an object saying “lets pick that up” and then turn and point to a place in the environments saying “and take it over there”, thus relating verbal and non-verbal communication in relation to an object and the environment.

2.1 Supporting Characteristics of Walk-in Displays

Although walk-in displays, such as a CAVE [5], share many characteristics with other display types, they have some unique features that support SHC. The user is physically placed within the virtual environment has a one-to-one scale view of remote users and objects within it. Natural body movement may be used to move around the immediate environment and to interact with users and objects [10, 23]. The user is aware of his body movement in relation to the environment through the linked senses of vision, proprioception, hearing and sometimes haptics. Linking of the visual and proprioception senses is not affected by the technology, as it is in other display types that replace the body with some graphical representation. Body movement, often just head and predominant hand, is continuously tracked, allowing both conscious and unconscious non-verbal communication to be captured. The wide field of view encourages natural head and body movement for both focussed and general observation. Spatial sound may be mapped to correlate with the relative position of remote users and objects.

2.2 Balancing Consistency with Responsiveness

A key requirement of Virtual Reality (VR) is the responsiveness of the local system. For example, delays in representing a perspective change following a head movement are associated

with disorientation and feelings of nausea. The collaboration in distributed virtual environments requires sufficient consistency of information represented to each user. The key goals of a CVE infrastructure are thus to maximise responsiveness, consistency and scalability in the face of limited and variable network bandwidth. This is primarily achieved through scaling and localisation. Scaling, such as awareness management, is concerned with allowing the complexity of the environment to scale, including the number of users and objects, without reducing the quality of experience of any user [7]. Localisation is achieved through replicating the environment, including shared objects and avatars, on each user’s machine. Sharing experience requires that replicas be kept consistent. This is done by sending changes (in the form of events) across the network. Localisation may go further than simply replicating the state of the environment and can also include the predictable behaviour of objects within it. Networks may, unfortunately, introduce dynamically changing delays and possibly loss and mis-order into event propagation. This can adversely affect the synchronisation, concurrency, causality and responsiveness of events. Synchronisation ensures that events are replicated within real-time constraints. Causal ordering ensures that causal relationships are maintained. Concurrency defines the ability of the system to allow events to occur simultaneously. Finally, responsiveness is the delay as perceived by the user after an action on the system. Concurrency and therefore responsiveness are reduced as the level of consistency is increased. This all leads to the need for consistency management, the role of which is to provide sufficient synchronisation and ordering whilst maximising concurrence and thus the responsiveness of the system.

2.3 Related Work

Several studies have investigated the effects of linking various combinations of display system for collaboration. It was found that immersed users naturally adopt dominant roles [21] versus desktop users. Recent studies by Schroeder [20] and Roberts et al. [17], using DIVE, investigated the effect of display type on collaboration of a distributed team. Greenhalgh et al. have investigated a number of assumptions concerning how networked CVEs operate and how they can support many users in the tasks they are trying to perform [8]. Hindmarsh et al. showed that the origins of disturbed interaction are mainly in visual discontinuities during activities caused by the desktop screen [11]. These problems mainly arose from a poor field of view and cumbersome and slow view changes. Another study compared effects of network delays over ISDN and Ethernet connection [14], where the object manipulation was under the strict control of only one individual, thereby eliminating interference between the two subjects. Pinho et al. describes a framework supporting the development of collaborative manipulation techniques in an immersive virtual environment [15]. Further, it has been found that when performing similar action the symmetric action interaction is superior to asymmetric action interaction and visa versa for diverse action [18]. These tests also found that a high quantity of verbal communication is used to compensate for the fragmentation of the work place on a desktop [19].

3. EXPERIMENTATION

Our analysis is based on experiments on remote users trying to solve a collaborative task via linked walk-in displays. We now describe the experimentation in terms of CVE platform, interface and task.

3.1 CVE

The DIVE system is an established test bed for experimentation of collaboration in virtual environments [4, 6, 9, 13] and, after three major revisions, it remains an effective benchmark. The advantages of this system are the support of rapid prototyping, cross-device support, and lower latency through point-to-point communication. DIVE was ported to walk-in display systems [22]. Subsequently an experiment on a loosely-coupled collaborative task with two users in different walk-in displays was found to be very successful [20]. We used DIVE version 3.3x5 for our experiments.

We extended DIVE with a plug-in for event monitoring [1]. The plug-in listens for event occurrences of chosen objects and records information, such as event-type, origin, involved objects and the current wall clock timestamp. The recorded information was then used to produce graphs, some of which are presented in the results section.

Another extension we added was event filtering, which reduced the frequency of events generated by the tracking system. Throughout our tests, the tracking system was filtered to only produce events for movements greater than one centimetre. In extensive testing, this level of filtering was found to produce the optimal balance between system performance and usability.

3.2 Task

For our experimentation, we used the virtual gazebo application, which was purposely developed to examine a set of distinct forms of interaction within a structured task [17]. This application simulates a construction site requiring shared manipulation of objects.

The application contains materials and tools for construction, both of which must be manipulated in a variety of ways. Screws fix beams in place and planks may be nailed to beams. Tools are used to drill holes, tighten screws and hammer nails. Although some aspects of the construction can be undertaken independently, the simulation of gravity ensures that collaboration is necessary for others. For example, a single person can place a metallic foot on the ground or drill a hole in a beam while it lies on the ground, whereas two people are required to carry or fix a beam.

Within our application, each user was represented by a human-like avatar representation. Such an avatar reflected articulations of a user's head and predominant hand based on input from a walk-in display's motion tracking system.

3.3 Procedure

We wanted to test the impact of the application and task when using immersive walk-in display devices (Figure 1). Data were collected in a series of six user trials between two walk-in displays at the University of Reading, UK and the Johannes Kepler University in Linz, Austria, over a period of two days. Table 1 summarises the configuration of the two setups. Both

remote sites have been connected via the standard Internet connections.

Table 1. Configuration of both walk-in displays.

Location	UK (Reading)	Austria (Linz)
Display	4 wall cubic display	4 wall cubic display
Tracking	Ultrasonic/Acoustic Intersense IS900	Magnetic Ascension MotionStar
Computer	2 pipes 6 dedicated processors SGI Origin 2000	4 pipes 128 shared processors SGI Origin 3800
Audio	Microphone headset	Microphone headset



Figure 1. A user (left) collaborating with a remote user (represented by avatar, right) via linked walk-in displays.

The participants were MSc and PhD students from one of the two mentioned universities. A team was built with one student on each site and the only time they met was through the CVE. The team is told the aim of the construction task (Figure 2) as well as intermediate goals but they are not told the particular method to achieve these goals. During a team's collaboration, data for the event traffic analysis was gathered by the event monitoring plug-in. Furthermore, we observed and video-captured how the two people in each system made use of SHC while undertaking various tasks. At the end of a test the participants were asked to fill out a post-questionnaire contributing to our observations.



Figure 2. The nearly completed task.

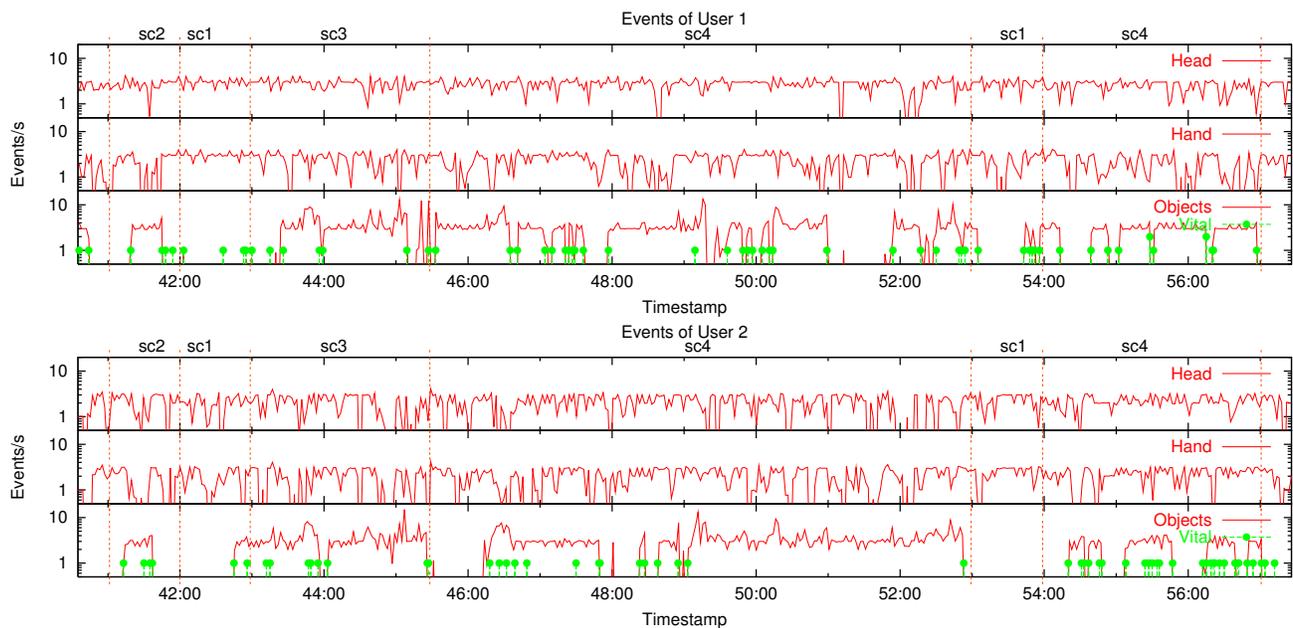


Figure 3. Remote event occurrences of users 1 and 2 for scenario sc1-sc4

4. RESULTS

Figure 3 shows a comparison of the frequency of event occurrences over the network between both remote sites. The figure shows typical graphs of avatar updates, caused by navigation through the virtual environment using the joystick and motion tracking of a user's head and hand within the walk-in display; updates of manipulated objects, caused directly by a user manipulating an object or indirectly by application-level consistency control; as well as particular event types that have been considered as 'vital' within the sub-tasks of our application. Such event types include collisions, grasps and releases as well as flags signalling an object state. Based on such events, the object behaviours trigger particular actions, modifying their own or another object's attributes. Although such vital events can be redundantly repeated, such as drilling a hole again, it would be very hindering for the task's performance if such an event gets delayed or lost. In the experiment, we measured overall proportions of about 34% head events, 31% hand events, 34% object events and 1% vital events. Within our trails, the frequency of avatar movement events was fairly continuous and constitutes of about 64% of event throughput. Object movement events only occurred during or shortly following interaction. Avatar events appeared more continuously, whereas object events occurred in burst and show occasional peaks. The average transmission of an event was under one second, whereas the average network delay between Reading to Linz was around 40ms.

In the next subsections that describe the various collaboration scenarios, we will look more closely into the characteristics of the event occurrences in the graphs and state our observations we made on SHC between the collaborating users.

4.1 Scenario 1: Planning and Instruction

Planning is necessary to determine method and responsibilities. Instruction occurs when a person demonstrates to the other how to undertake a given operation, such as using a tool to fix two construction objects together.

4.1.1 Typical use of SHC

The process of planning and instructing requires that everybody involved sees and hears the discussion. Verbal communication is essential to describe the upcoming task and to agree on locations and coming steps. Other cues, such as gestures, are widely used to point out directions and to underline the verbal communication. Earlier studies showed that simple embodiments contribute to the interaction and a more realistic humanoid avatar representation may support better collaboration [2]. Thereby the faithfulness of the avatar gestures is as important as the realism of the environment.

Observations show that in the planning phase, users mainly stay close together or use body-centric gestures such as facing each other, while they discuss their further action. The use of walk-in displays supports this kind of user behaviour by allowing the user to turn and move naturally within the spatial context. Planning can also involve the use of objects. For example, when explaining how or where to use a tool, it is easier to take the object and to demonstrate it. Using objects to describe an action is typical in real world interaction, even if the object is only used to mimic the action. When the environment is designed to support such communication, it contributes to the planning and also later to the task. For example, the use of different textures on similar objects can help to make verbal references to those objects.

From video footage taken during user trials, we have observed numerous nuances that link verbal and non-verbal communication while referring to objects and places within the environment. For example, a user might point to an object and say “lets pick that up” and then turn and point to a place in the environment saying “and take it over there”, or the user simply takes a object and tells the other user to do the same. Much of the non-verbal communication identifiable during planning consisted of turning, pointing and nodding.

4.1.2 Event logs

In the event logs, planning phases can be seen where objects are not manipulated at either site, as can be seen around timestamps: 41:00, 42:00-43:00 and 53:00-54:00 (marked with “sc1” in Figure 3). Transitions between planning and doing are clearly seen as rises in avatar updates. Any occurrences of vital events are not usually essential in this scenario.

It is interesting to see that the head generates updates more frequently than the hand at these times. Unfortunately, event occurrences of gestures and locomotion are not easily distinguishable in our event logs. Therefore, we reconstructed the path (trace) of the user’s head and hand movement (Figure 4) based on the logged events. An analysis of these paths revealed that the events mostly contained position updates for short distances (<5 cm for head and <10 cm for hand), but with a large deviation in changes of direction along the path between each update. This indicates that users mainly reside in one place and just move their head and hand back and forth. Simply filtering transformation update events for distances less than 10cm in this scenario would have reduced the event traffic caused by head updates by 45% and hand updates by 20%.

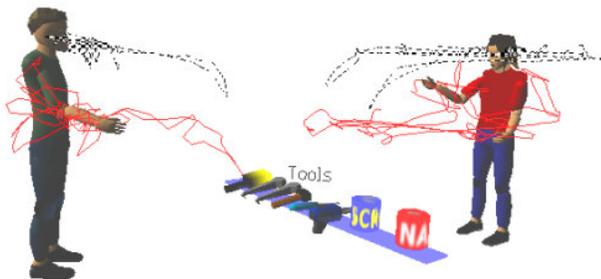


Figure 4. Traced head (dotted line) and hand (solid line) events in a planning scenario.

4.2 Scenario 2: Working Separately

The team temporarily split to undertake separate but related tasks, such as gathering material objects, while keeping an eye on each other in case help is needed.

4.2.1 Typical use of SHC

Most collaborative tasks involve some independent work. At such times, communication reduces naturally due to independence and concentration on one’s own task. It is interesting to note that, at such times, small talk replaces work related verbal communication, and this seems to maintain the social feeling and increases the feeling of co-presence. Changes in level or type of activity or level of communication of the other person can signal a need to bring collaboration closer

together again and communicate directly. We often observed people looking over to see how their partner was getting on and offering assistance when necessary, for example, by fetching a tool for the other to use. Representing interaction with an object through natural body movements, driven by motion tracking, made such changes in activity easier to spot. Furthermore, the naturalness of view change offered by motion tracking as well, as the wide field of view in a cubic display, simplifies keeping a watch on others.

The ease and naturalness of collaboration between the test subjects was very noticeably improved above that in earlier trials that linked a desktop and walk-in display [17]. This was particularly apparent for parts of the task where both participants concurrently positioned one or more objects. Unlike previous desktop trials [11], we did not observe delays in identifying objects referenced through verbal and non-verbal communication. This suggests that the combination of wide peripheral vision and ability to glance using unconstrained head movement overcomes this problem.

4.2.2 Event logs

An example of independent work is illustrated between timestamps 41:00-42:00 in section “sc2” in Figure 3. The graphs show continuous bursts of object events as a result of fine-grained remote representation while the respective object is being manipulated. Vital events occur rarely. Some of them, such as collisions, appear asynchronously to object manipulations, whereas others, such as grasp or release, occur synchronously with start and end of bursts of object events. However, the resolution of object updates remains constant, independent of any obvious interest. As all sorts of sub-tasks, such as navigation, moving and positioning of objects are involved in this scenario, we expected a wide range of angular updates and distances. However, an analysis of reconstructed traces of avatar head and hand updates showed mostly small differences in distance and angle between the update events. When users travelled long distances, they travelled quickly, and thus generated only a few event updates. Most events occurred during object manipulation. Although fine nuances of object updates supported the naturalness of interaction on the local site, these were often overseen on remote sites. Hence, only large changes, such as waving the hand when attracting attention or nodding, would need to be transmitted to replicas.

4.3 Scenario 3: Moving an Object Together

Here two users work together to carry an object to a given place in the environment (Figure 5). Simulated gravity ensures that larger construction objects cannot be lifted alone and simulated friction restricts dragging.

4.3.1 Typical use of SHC

Supporting various forms of SHC is particularly important when people come to collaboratively manipulate an object. For example, when moving an object from point A to B, users must agree on roles and responsibilities as well as actions. Initially they must agree on where are they going, how are they going to get there, who will take the lead and what problems may be encountered. It is very important that the users can see which end of the object is picked up. Highlighting the manipulated part

of the object supports this. Once underway, an agreement on speed and adaptations to path must be communicated and acted upon. Walk-in displays are well suited to support SHC in tasks such as collaborative manipulation of objects. This is because, in addition to supporting gesturing and voice as described above, each user interacts with the object in a spatially natural way with important aspects of this body movement represented remotely with respect to the object.

Video footage shows that users combine speech with a complex set of gestures to communicate initial planning. This set includes nodding and shaking of the head and pointing, not only with the hand but also the head and body. For example a user may face his partner, point with the hand to the object and move the head between the two. Once the object has been lifted the users keep each other in sight, either by walking side by side or the one in front walking backwards. As the primary hand is used to carry the object, gesturing plays less of a role and users make more use of body and head orientation as well as the object, the environment and verbal communication, in order to determine changes in movement and role.



Figure 5. Two remote users carrying a beam.

4.3.2 Event logs

This scenario is demonstrated in the events logs in Figure 3 between timestamps 43:00-45:30, “sc3”. User 2 starts by moving a damper object, but the shared object moves only when user 1 joins the shared manipulation. The actual manipulation of the shared object can be seen in the peaks shown in the object event graphs, as a result of our supporting concurrency mechanisms. Vital events often bracket the manipulation of a shared object attribute and seldom coincide with it. On average, this scenario shows clearly more event traffic for object updates (40%) than for avatar updates (around 30% each for head and hand). However, again we observe slightly more head updates than for the hand. The reconstructed traces show mostly straight paths with small angular changes when navigating through the environment, as seen on the right side in Figure 7. Dead reckoning may be applied here to reproduce motions between path changes. However, high responsiveness is required to provide the users with quick feedback while communicating through the shared object.

4.4 Scenario 4: Assembling Objects Cooperatively

In our application, objects are connected by holding them together, while drilling a hole through both and inserting and tightening a bolt. Typically, a construction will be built from the ground up, fixing one item in place at a time. Simulated gravity requires that one user must hold an object in place while the other fixes it (see Figure 2 for an example).

4.4.1 Typical use of SHC

Similar to scenario 3, during an assembling task it is necessary to communicate with the other partner. Verbal communication helps to agree on the next step, but also provides the other user with further information. For example, a user positions an object while the other is told to gather the necessary tools to fix it. The first user can be of assistance by verbally directing how to fix it or informing the other person where a tool can be found (“look behind you”). Nevertheless, they can help each other by pointing out where to place an object, how to use the object itself, and what to do next. Again users can support each other’s work with a complex set of gestures. For example, nodding with the head or looking to the object instead of to the partner shows the other user agreement or a direction. As before, the technology of the walk-in display can support this kind of interaction. It allows the users to move with their around the object of interests, without time consuming locomotion to get the right view position as would be required on a desktop. Furthermore, it is helpful when objects do support the communication between two users interacting around it, e.g. through a change of texture or other feedback when a “sub-goal” is archived. Thereby, all participants can see how far the progress is and they may or may not continue with their work.

4.4.2 Event logs

The period over timestamps 45:30-53:00 and 54:00-57:00 in sections “sc4” in the Figure 3 corresponds to a scenario of holding and fixing. Comparing the graphs, one can recognise elements of scenario 2 and 3. A synchronisation of vital events can be clearly seen at the start and end of most bursts of object updates. However, similar to scenario 2, head and hand updates occur more frequently at these times. The traces in Figure 6 show many short paths around the manipulated objects. As a high resolution of object movement is of importance for both collaborating users, it is hard to find a fixed threshold for event filtering in this scenario. Adaptive event filtering that controls the update rate based on an object’s context (e.g. reduce when held at target position) would be of advantage here. Nevertheless, due to the increased load of event traffic it is likely that vital events may get delayed. For example, the very last vital event at 57:00, generated by user 1, appears a fraction earlier than the actual trigger event of user 2 at 57:10. Such incidents can be observed here and there throughout the scenario and represent a risk of disorder. In this specific example, the object would fall down and change to “fixed” state on the ground.

During our experiments, we observed several occurrences of collaborators realising inconsistencies between each other’s view of the environment that arose through loss of vital events. Each time the team were able to bring the environment back to a

sufficiently constant state through communicating the differences and re-enacting the activity that should have resulted in the change. A common example is that a drilled hole would only appear at one site, requiring someone to drill a second hole so that it would usually appear at both. Such repair work took time and eroded the performance of the shared task. Interestingly, however, few people mentioned this as problem when asked what factors effected collaboration.

The performance of our application could have been improved if we had been able to send vital events reliably while maintaining responsiveness through unreliable transmission of movement events. In our task, more vital events occur around fixing operations than carrying. As one might expect from this, most of the losses of synchronisation were observed during fixing. Although the frequency of human movement events remains fairly constant, as shown in Figure 3, traces of the movement clearly show that significant changes in path occur frequently during fixing operations and comparatively seldom during carrying. This is seen in the movement of avatar hand and head in Figures 6 and 7. It seems likely that a dead-reckoning algorithm could be devised to greatly reduce event frequency during carrying. However, applying a similar approach during the fixing task would be far more challenging, due to the frequency and magnitude of changes in path.

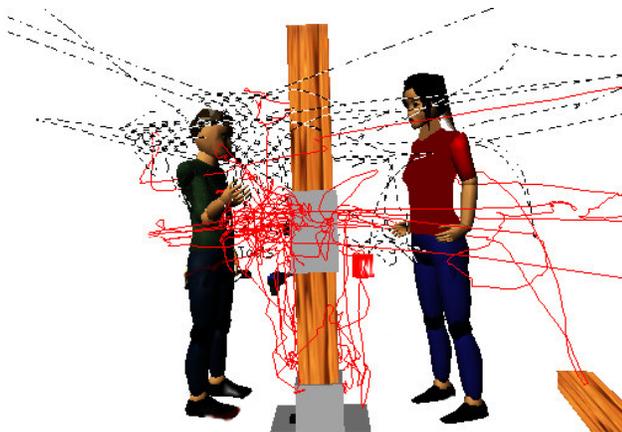


Figure 6. Traced head and hand events while assembling objects.

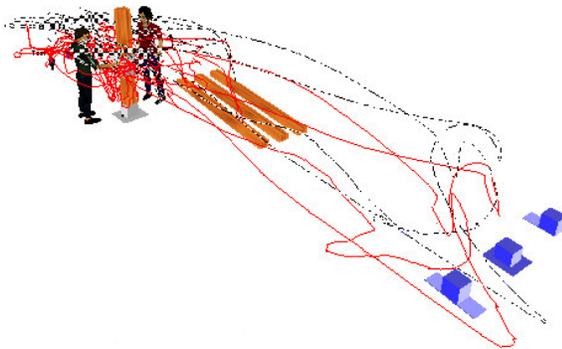


Figure 7. Traced head and hand events over the whole task.

5. CONCLUSIONS

The use of walk-in displays significantly improved the observable naturalness and performance of a distributed team, compared to the exclusive use of desktop, or of an asymmetric pair of both display types, as used in previous trials. Delays in identifying objects referenced by others were not observed, indicating that the characteristics of walk-in displays overcome the problems reported in earlier trials such as [11].

Supporting natural interaction implies increased event traffic for detailed movement representation of avatar and objects. The magnitude and importance of movements depends on the kind of collaborative scenario. Within our trials, the frequency of avatar movement events is fairly continuous and constitutes 64% of event throughput. Object movement events only occur during or shortly following interaction. The highest peaks in frequency of events come from shared object attribute manipulations and are caused by the added effect of the consistency mechanism. Vital events, which synchronise or trigger actions and are essential for steering the application, often bracket the manipulation of a shared object attribute and sometimes coincide with it.

It is very likely that the overall event traffic could be substantially reduced by applying dead reckoning to scenarios such as carrying. Applying this technique to other scenarios involving a higher degree of social human communication is likely to be difficult and unproductive, due to the frequency and magnitude of significant movement and the importance of such movements in conveying nuances. In our application, points of intense movement often coincided with causally significant changes to the shared object. This exacerbates the likelihood of significant inconsistencies arising from the loss of vital events. Responsive communication of these movements would be hard to achieve reliably. These results reinforce the need for multiple quality of service event delivery as demonstrated in [16]. Much of the non-verbal communication consisted of turning, pointing and nodding. We are investigating the use of Hidden Markov Models to categorise and identify such behaviour allowing its communication to be reduced to behaviour identifiers.

These early trials suggest that linked walk-in display technology will be of great benefit to spatially unconstrained teleconferencing, especially where users collaborate through and around shared information artefacts. We believe that two pressing issues are the integration of better consistency management and of input devices capable of detecting the majority of non-verbal communication.

6. ACKNOWLEDGMENTS

This work has been supported by the British Council in a bilateral cooperation between the University of Reading, UK, the University Collage London, UK and the Johannes Kepler University Linz, Austria.

Several of our colleagues contributed to this work, most notably Ali Al-Khalifah and Detlef Krischker at the University of Reading, Prof. Jens Volkert at the GUP and Johann Messner at the ZID.

7. REFERENCES

- [1] Anthes, C., "Event and latency measurement in collaborative Virtual worlds", The University of Reading, Reading, UK, MSc Thesis 2002.
- [2] Bowers, J., Pycock, J., and O'Brien, J., "Talk and embodiment in collaborative virtual environments", *In Proceedings of the ACM SIGCHI'96*, Vancouver, British Columbia, Canada, pp. 58-65, 1996.
- [3] Burgoon, M., Hunsaker, F. G., and Dawson, E. J., *Human Communication*, 3rd ed. London: SAGE Publications, 1994.
- [4] Carlsson, C. and Hagsand, O., "DIVE - A platform for multi-user virtual environments", *Computers & Graphics*, vol. 17 (6), pp. 663-669, 1993.
- [5] Cruz-Neira, C., Sandin, D., DeFanti, T., Kenyon, R., and Hart, J., "The CAVE: Audio Visual Experience Automatic Virtual Environment", *Communications of the ACM*, vol. 35 (6), pp. 65-72, 1992.
- [6] Frécon, E., Smith, G., Steed, A., Stenius, M., and Stahl, O., "An Overview of the COVEN Platform", *Presence: Teleoperators & Virtual Environments*, vol. 10 (1), pp. 109 - 127, 2001.
- [7] Greenhalgh, C. M., "Analysing Awareness Management in Distributed Virtual Environments", *Presented at the 2nd Annual Workshop on System Aspects of Sharing a Virtual Reality, part of CVE'98*, Manchester, UK, 1998.
- [8] Greenhalgh, C. M., "Understanding the network requirements of Collaborative Virtual Environments", in *Collaborative Virtual Environments: Digital Places and Spaces for Interaction*, D. Snowdon, E. F. Churchill, and A. J. Munro, Eds. London, UK: Springer Verlag, pp. 56-76, 2001.
- [9] Greenhalgh, C. M., Bullock, A., Frécon, E., Lloyd, D., and Steed, A., "Making Networked Virtual Environments Work", *Presence: Teleoperators and Virtual Environments*, vol. 10 (2), pp. 142-159, 2001.
- [10] Heldal, I., Steed, A., Spante, M., Schroeder, R., Bengtsson, S., and Partanan, M., "Successes and Failures in Co-Present Situations", *forthcoming in Presence: Teleoperators & Virtual Environments*, 2004.
- [11] Hindmarsh, J., Fraser, M., Heath, C., Benford, S., and Greenhalgh, C., "Object-Focused Interaction in Collaborative Virtual Environments", *ACM Transactions on Computer-Human Interaction (ToCHI)*, vol. 7 (4), pp. 477-509, 2000.
- [12] Knapp, M. L., *Nonverbal Communication in Human Interaction*, 2nd ed: Holt, Rinehart and Winston, 1978.
- [13] Mortensen, J., Vinagayamoorthy, V., Slater, M., Steed, A., Lok, B., and Whitton, M. C., "Collaboration in Tele-Immersive Environments", *Presented at the Eighth Eurographics Workshop on Virtual Environments*, Barcelona, 2002.
- [14] Park, K. and Kenyon, R., "Effects of Network Characteristics on Human Performance in a Collaborative Virtual Environment", *In Proceedings of IEEE VR '99*, Houston, TX, pp. 104-111, 1999.
- [15] Pinho, M. S., Bowman, D. A., and Freitas, C. M. D. S., "Cooperative Object Manipulation in Immersive Virtual Environments: Framework and Techniques", *In Proceedings of ACM Symposium on Virtual Reality Software and Technology (VRST)*, Hong Kong, China, pp. 171 - 178, 2002.
- [16] Roberts, D. J., "A Predictive Real Time Architecture for Multi-User, Distributed, Virtual Reality". Ph.D. Thesis, University of Reading, UK, 1996.
- [17] Roberts, D. J., Wolff, R., and Otto, O., "Constructing a Gazebo: Supporting team work in a tightly coupled, distributed task in virtual reality", *Presence: Teleoperators & Virtual Environments*, vol. 12 (6), pp. 644-657, 2003.
- [18] Ruddle, R. A., Savage, J. C., and Jones, D. M., "Symmetric and asymmetric action integration during cooperative object manipulation in virtual environments", *ACM Transactions on Computer-Human Interaction (ToCHI)*, vol. 9 (4), pp. 285-308, 2002.
- [19] Ruddle, R. A., Savage, J. C., and Jones, D. M., "Verbal communication during cooperative object manipulation", *In Proceedings of Proceedings of the ACM Conference on Collaborative Virtual Environments (CVE'02)*, pp. 120-127, 2002.
- [20] Schroeder, R., Steed, A., Axelsson, A.-S., Heldal, I., Abelin, Å., Wideström, J., Nilsson, A., and Slater, M., "Collaborating in networked immersive spaces: as good as being there together?" *Computers and Graphics*, vol. 25 (5), pp. 781-788, 2001.
- [21] Slater, M., Sadagic, A., Usoh, M., and Schroeder, R., "Small Group Behaviour in a Virtual and Real Environment: A Comparative Study", *Presence: Teleoperators and Virtual Environments*, vol. 9 (1), pp. 37-51, 2000.
- [22] Steed, A., Mortensen, J., and Frécon, E., "Spelunking: Experiences using the DIVE System on CAVE-like Platforms", in *Immersive Projection Technologies and Virtual Environments*, B. Frohlicj, J. Deisinger, and H.-J. Bullinger, Eds. Wien: Springer-Verlag, pp. 153-164, 2001.
- [23] Steed, A., Spante, M., Schroeder, R., Heldal, I., and Axelsson, A. S., "Strangers and Friends in Caves: An Exploratory Study of Collaboration in Networked IPT Systems for Extended Periods of Time", *In Proceedings of ACM SIGGRAPH 2003 Symposium on Interactive 3D Graphics*, Monterey, California, USA, pp. 51-54, 2003.