The Pitfalls in System Design for Distributed Virtual Environments: A Case Study

Manuel Oliveira Jesper Mortensen Joel Jordan Anthony Steed Mel Slater

Computer Science Department

University College London

Gower Street WC1E 6PT London UK

+44 419 3876

{m.oliveira | j.mortensen | j.jordan | a.steed | m.slater}@cs.ucl.ac.uk

ABSTRACT

There is a common understanding amongst the games industry that the Internet is a poor network infrastructure that wreaks havoc on all game engines. This view is shared amongst the Distributed Virtual Environment (DVE) research community, who traditionally find the Internet as the main culprit for disruptions in the experience of end-users. In fact, the majority of the problems reside in the existing misconceptions regarding the nature of the Internet, leading to inefficient and inadequate network subsystems.

This paper reports the lessons learnt from the attempt of using an existing DVE system to support collaboration using haptic devices over the Internet2.

1. INTRODUCTION

The history of the Distributed Virtual Environment (DVE) is recent and brief [10], albeit rapid evolution from the text-based environments and Multi-User Dungeons (MUDs) to more sophisticated 3D immersive environments. With the computer being a mass-market commodity and the cheap network connectivity provided by the Internet, DVEs have abandoned the exclusivity of high technology research laboratories and become accessible by non-expert end-users at their homes.

Unfortunately, the existing DVE systems are fraught with technical difficulties that ultimately affect the user experience by disrupting both their sense of presence [2] and co-presence [8]. Most of these problems are rooted in the network support of the systems for effectively maintaining the illusion of a consistent environment that is distributed amongst geographically dispersed hosts. In fact, most of the DVE work is relatively successful within controlled conditions such as within a LAN, but are unsuccessful when deployed on the Internet demonstrating different failure thresholds [6]. The transition to the Internet is not the culprit; rather it magnifies the existing network problems contained within systems that otherwise may go unnoticed.

Much of the problem may be found in the fact that DVEs have emerged as part of the natural transition from single-user to multi-user systems when integrating networking capabilities. The development emphasis goes towards the graphical technical issues since it affects the visual experience of the user. However, the current implementations and associated problems demonstrate a poor understanding of the network and its behaviour. A classical example is the preliminary transition of the Doom™ game from single player to multiplayer. The first approach was to send a network data packet per keystroke, quickly saturating the network with the increase in traffic. In this particular case, updates were sent every n[th] frame, which could rapidly lead to network problems and overwhelm the world server since it was based on a client/server architecture.

This paper reports the failure of adopting a DVE system to support haptic collaboration across the Atlantic using the Internet2 infrastructure. The problems resided within the DVE system and not the network infrastructure. Although a specific system was used, the case is similar across other DVE systems.

The remainder of the paper is divided into an additional 4 sections, starting with explaining the importance of haptic feedback and its impact on the sense of co-presence. In section 3, the case study experiment is presented, along with the problems that compromised the initial design, requiring modifications. The experience garnered from this work has lead to design principles that are the basis of a prototype system. This system will be improved as discussed in the final section.

2. THE "SENSE" OF TOUCH

With single user virtual environment systems, the objective is to guarantee the immersion of the user so they feel present in the alternate reality at all times. However, when considering DVE, the sense of presence is not the sole factor to consider since a user is sharing the environment with other users. Thus, it is important to also guarantee the sense of co-presence. In fact, some studies [9] have demonstrated that co-presence contributes significantly to the sense of presence.

In [1], an experiment was carried out to validate the impact of haptics on the sense of co-presence. The experiment consisted of two users manipulating a ring along a curved wire without touching it. If the ring did touch the wire, the colors of the screen would change until the error was corrected. Two conditions were explored: Only visual feedback; Visual and haptic feedback. The results indicated that both task performance and co-presence increased when haptic feedback was used instead of just visual.

The study aimed to correlate haptics with co-presence, thus the infrastructure setup excluded the complexity of a network by having two haptic devices connected to a single host.

In [3], the study was done using a LAN as the network infrastructure. So all the users interacted through separate

ITP'02, December 6th, 2002, Juan Les Pin, France.

computers each with a haptic device connected to it. In order to emulate the Internet, an artificial delay was introduced, but this is not sufficient as demonstrated in [6]. The behavioral analysis of the Internet cannot be reduced to just latency. It is necessary to consider the loss and jitter, along with data packets out of order. Without considering these issues, once the system is deployed across the Internet, the results may be appalling even if the latency experienced is less than the one emulated on the LAN during development [6].

3. THE EXPERIMENT

The initial objective of the experiment [7] was to study the use of haptics in a DVE across transatlantic links. The target DVE system to be used was Distributed Interactive Virtual Environment (DIVEv3.3) [5] with minor modifications to integrate the haptic devices. The choice of DVE system was based on six-years experience using it in collaborative virtual environment research. Some of the strong advantages of the system include its simplicity of use. The possibility of loading different geometry file formats coupled with an easy scripting engine based on tcl/tk, allows a researcher to promptly setup experiments with minimal system tampering.

Previous experience was based on collaborative scenarios where the interaction amongst users was limited to social interaction based on speech and visual cues [9]. The inclusion of haptics revealed some serious limitations in the DIVE system design regarding the networking infrastructure.

The following subsections will provide an overview of the experiment design and describe the network infrastructure along with the problems encountered that limited the scope of the initial research objectives. The actual results of the experiment are reported elsewhere [7].

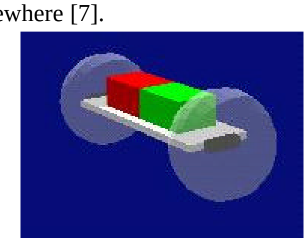

Fig. 1 - The "strecher" to be jointly carried by the users

3.1 Experiment Design Overview

The experiment involved two users that would collaborate in a joint task. The objective was to carry a flat object with two handles having objects resting upon it, as illustrated in Fig. 1. For convenience, the object was described to the subjects as a stretcher to be jointly carried with another user.

The environment itself was reduced to bare essentials, thus providing minimal distraction to the task being carried out. The result, depicted in Fig.2, consisted of a large building with a blue path starting outside and leading inside through a large entrance.

The path that subjects had to follow consisted of several axis-aligned segments. Although the subjects were encouraged to follow the path, they were briefed that the general direction of the path was more important than following it exactly.

The task would be initiated with the "stretcher" being located on the floor at the beginning of the path outside of the building. Successful completion of the task meant depositing the "stretcher" on a red square located within the building at the end of the blue path.

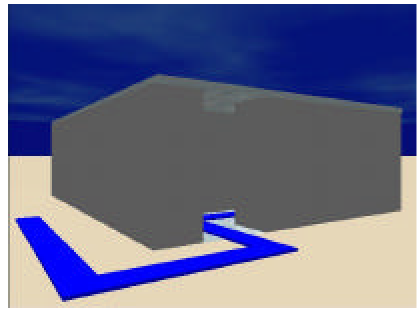

Fig. 2 - The environment with the simple building and the blue path indicating the trajectory to be followed

The study was conducted between UNC and UCL with a total of 17 subjects all recruited at UCL. The experimenter at UNC acted as a confederate throughout the experiment, aiding the subjects in their task. The assigned chore was terminated either upon successful completion or when the time limit of 8 minutes was reached.

To assist a user in discerning the position and orientation of their remote partner, users were represented by a simple block avatar. The avatar could only move its head and the pointer indicated the position of the hand. These minimal cues, in terms of body language, were complemented by the use of audio communication to allow participants to verbally negotiate their progress along the designated path.

The hardware setup used at UCL consisted of a ReaCTor system, consisting of four projection walls each with an area of 3x2.2 meters. The subjects would control their navigation and manipulation via a joystick with four buttons. They would wear, in addition to the goggles, the Intersense tracking device.

3.2 Network Infrastructure

A high level overview of the network topology used is depicted in Fig. 3.

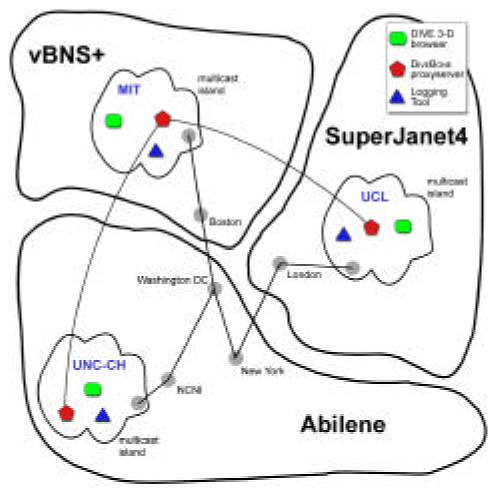

Fig. 3 - Network topology of the interconnecting sites

Although the actual experiment was conducted between UNC and UCL, initial testing and evaluation was carried out between all sites. The overall RTT between UNC and UCL stabilized around 80-90ms along with an estimated loss below 0.8%.

3.3 Problems Encountered

Although network performance was deemed adequate for the experiment via network monitoring, the initial objective of using haptics was compromised. An initial pilot study was carried out with a simplified scenario where two remote users would collaborate together to lift a box and maintain it above a certain threshold. Each user would be required to exert a reasonable force on opposite sides of the box in order to achieve the necessary leverage to lift it. The forces would be required to be at an angle to actually raise the box.

However, even though network conditions were reasonable, the system's performance was incredibly poor, making it unfeasible

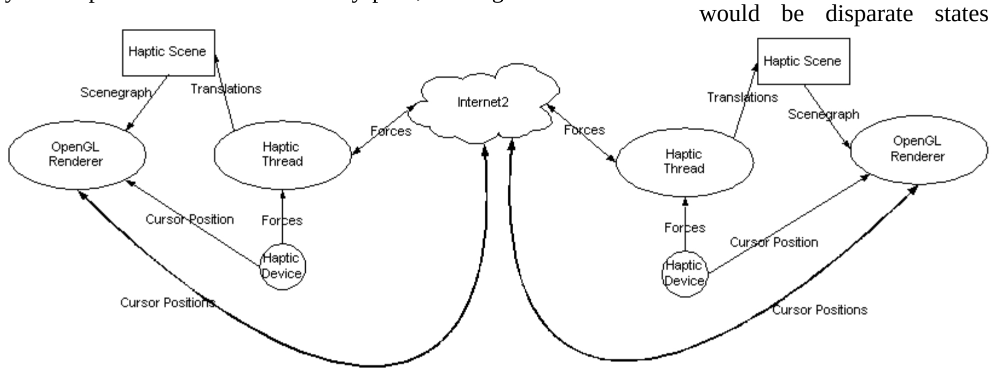

Fig. 4 - Overview of the parallel data structures within a DVE system with haptic device

to attempt the completion of the task. The nature of the problem is the different frequency of the updates of the processes as seen in Fig. 4.

The visual renderer handles the visual aspect of the DVE by performing the traversal of some data structure, typically an object scenegraph. To simulate the force feedback within a DVE, the haptic scenegraph is used to calculate all the forces applied to the objects and the reaction forces to be applied to the haptic device. Both the visual and haptic scenegraph are required to be consistent, so the feeling of touch reinforces what is visually perceived and vice-versa.

Now traditionally, the renderer performs frame updates between 10Hz to 60Hz depending on the performance of the computational resources (CPU, graphics card, etc). The haptic device on the other hand performs updates at 1KHz, which is two orders of magnitude greater than the renderer. This implies different processing threads to handle the divergent frequency rates. However, in DIVE the event loop is tightly coupled to the rendering loop, thus events may only be processed at a speed that is either the same or less than the frame rate. With remote collaboration, the local haptic events are required to be processed not only locally, but also need to be sent to other participating hosts. With the low frequency rate coupled with the delay of the network the visual and haptics scenegraphs would become unsynchronized quite quickly. Therefore, it was concluded that it would not be possible to use haptic devices with DIVE.

The experiment carried out was based on DIVE but without haptics, limiting the negotiation process to visual feedback and audio exchange between users. This incurs difficulties not normally found in the real-life manipulation of objects. The existence of gravity and the physical constraints of a similar object to the virtual "stretcher" would restrict movement of the participants. So for example, when one user pulls the object, the other user is pulled along being obliged to follow the movement.

The joint manipulation of the same object, identified as a "stretcher", posed several difficulties. It involved manipulating the local copy of the object and letting the DIVE system propagate translational and rotational changes to other remote copies of that object on the network thereby creating a sense of shared ownership of the entity in question. This mode of manipulation would only guarantee a synchronized environment as long as the changes were applied in the same order in both instantiations of the environment. In turn this would only be possible if the events were generated, sent and processed at a higher resolution than the frequency of the manipulation of the object. If not, there would be disparate states of the object in its various instantiations, each of which would then send updates of its global position resulting in significant jitter of the object. Also, it would continually swap between the local perceived state and the one received in the packets from other instantiations of the VE. Before any given frame is rendered the state of the object would be determined either by the local or the remote state due to processing of remote packets.

Until pilot experiments were run, the experiment was carried locally and as the LAN provided less than 10ms turn around times the system did not present any problems. As soon as a link up with UNC was carried out, thus experiencing a RTT of 80ms, the result was the shared object would jitter. This was resolved by implementing an alternative approach that employed distinct local copies of the contents of the stretcher and shared handles, each owned by a single avatar. A local TCL script then updated the distinct local object based on the state of shared global objects. The stretcher would then align locally based on the position/orientation of the handles. In this set-up direct manipulation of a shared object was avoided. The VE appeared synchronized and visually correct, even though the two instantiations would differ slightly due to the lag in updating the positions of the handles. The stretcher would align according to the position of the hand of the subject locally and the position of the rendered hand of the remote avatar. So the alignment of the stretcher was based on the information available locally at the time of rendering.

4. AN ALTERNATIVE SOLUTION

The initial objective of evaluating remote haptic collaboration was abandoned due technical insufficiencies of the DVE system. However, serious problems remained that occasionally affected the experience of the users [7]

In order to address the initial goal of the study, a prototype was developed addressing the problems that afflicted DIVE. The initial experiment consisted of a scene illustrated in Fig.5 where the users would have to manipulate together a blue box (dark shaded cube). The task was to try collaboratively apply opposing forces such that it was possible to lift the cube above a particular threshold and maintain it for a predetermined time.

No visual feedback was available regarding a representation of the remote user, neither was there any auditory feedback. Only a representation of the cursor indicating the contact point of the local and the remote user was portrayed. These limitations make the negotiation process extremely difficult without the usage of haptic feedback. To ameliorate the difficulties experienced, an enhancement to the user interface was implemented to explicitly denote where the remote user would perceive the position of the box by painting a transparent pink cube (lighter shaded cube). This improved the negotiation process, as the users could understand better what the other participant was doing when their models became unsynchronized.

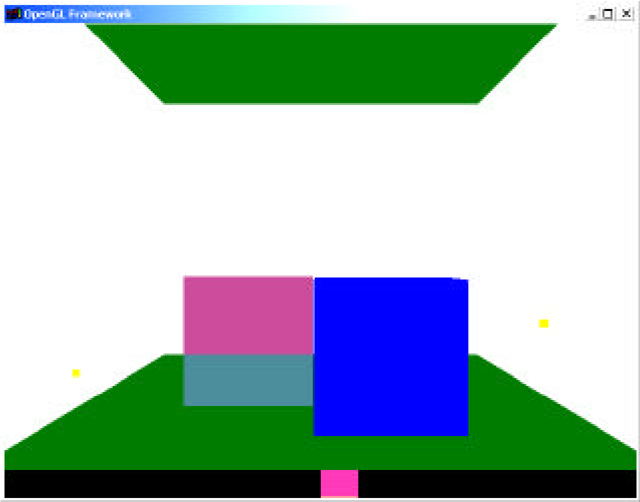

Fig. 5 - Screenshot of the user interface of the prototype system developed for collaborative haptics

Considering the limitations of coupling the event processing subsystem with the renderer subsystem, the system developed was based on a multi-threaded architecture. Therefore, it was possible for the rendering to be processed at 30Hz (or any other desired frame rate) whilst the haptic loop would run at 1KHz to fulfil the requirements of the haptic device. The network I/O code that related to the haptic subsystem was written into the haptic event loop.

Both disparate machines ran the code independently, and had their own copy of the environment. This meant that any environmental change made at a machine was communicated and then applied at the remote instance. The most effective method to connect the two distant instances of the environment was to send the reaction forces applied to the local cube to the remote system. Conversely, each instance applied any forces received over the network to its cube in addition to any forces applied by the local user. The network subsystem was implemented with two optional transport protocols, one based on TCP and another on UDP. The latter required a header definition to uniquely identify the packet by including a sequence number and timestamp.

Empirically it was found that TCP was inadequate to support user interaction successfully. This was due to the sensitivity in terms of the latency, loss and jitter regarding the haptic device. With the relatively high frequency rate, the TCP sliding window (waiting on positive acknowledgements for every packet) and Additive Increase Multiple Decrease (AIMD) behavior of the congestion control would wreak havoc in any attempt of collaboration.

The protocol was based on UDP, since the design did not require a total reliable protocol and consequently had better performance. The main two reasons were the low loss experienced on the network, along with the timeliness requirements of the data that would render pointless any retransmissions due to exceedingly large delays. If a loss did occur, it would affect the environment by reducing the power of a force, and therefore desynchronize the location of the cube instances in the shared environment. Although little packet loss was experienced, there would be considerable jitter at times that would lead to temporary inconsistencies with abrupt feedback forces.

A small study was carried out [4] to evaluate the prototype and the results are positive, with haptic feedback playing a major role in reinforcing the sense of co-presence. Based on the data collected from both the application and network monitoring, further refinements are underway to improve the system.

5. CONCLUSIONS AND FUTURE WORK

Both the games industry and the DVE research community view the network in an overly simplistic fashion, reduced to sending and receiving messages via different types of sockets.

This paper reported on the experience garnered during the study of collaborative haptics in trans-Atlantic experiments. The Internet-2 revealed to have adequate performance to support DVE interaction, but the strict requirements of haptic collaboration exposed pitfalls in an existing DVE system.

Ultimately, the initial objectives of the experiment were modified and the lessons learnt have lead to the development of a prototype system. The preliminary results of the system's usage are promising, but further work is required to improve the performance and haptic interaction. The next phase of the work will include a more sophisticated buffering mechanism coupled with an adequate physical model to provide further synchronization and smoothing.

6. ACKNOWLEDGEMENTS

This work was partially funded by the EPSRC Equator project.

7. REFERENCES

1. Basdogam, C., Ho, C., Srinivasan, M., and Slater, M., "An Experimental Study on the Role of Touch in Shared Virtual Environments", ACM TOCHI, Vol. 7, N. 4, December, 2000
2. Biocca, F., "The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments ", Journal of Computer Mediated Communication, Vol. 3, N. 2, September 1997
3. Choi, H., Choi, B. and Ryew, S., "Haptic Display in the Virtual Collaborative Workspace Shared by Multiple Users", Proc. IEEE International Workshop on Robot and Human Communication, 1997
4. Jordan, J., Mortensen, J., Oliveira, M., Slater, M., Tay, B., Kim, J., and Srinivasan, M., "Collaboration in a Mediated Haptic Environment", The 5th Annual International Workshop on Presence , Porto, October, 2002
5. Hagsand, O., "Interactive Multiuser VEs in the DIVE System", IEEE Multimedia, spring, 1996
6. Lincroft, P, "The Internet Sucks: Or, What I Learned Coding X-Wing vs TIE Fighter", Proc. Game Developer\s Conference, 1999
7. Mortensen, J., Vinayagmoorthy, V., Slater, M and Steed, A., "Collaboration in Tele-Immersive Environments", Proc. 8th Eurographics Workshop on Virtual Environments, Barcelona, June, 2002
8. Schroeder, R., Steed, A., Axelsson, A., Heldal, I., Abelin, A., Wideström, J., Nilsson, A. and Slater, M., "Collaborating in networked immersive spaces: as good as being there together?", Computer & Graphics, Elsevier Science Ltd, Vol. 25, N. 5, October 2001
9. Slater, M., Sadagic, A., Usoh, M. and Schroeder, R., "Small Group Behaviour in a Virtual and Real Environment", Presence: Teleoperators and Virtual Environments, Vol. 9, N. 1, 2000
10. Waters, R. and Barrus, J., "The Rise of Shared Virtual Environments", IEEE Spectrum March 1997