Effects of Network Characteristics on Human Performance in a Collaborative Virtual Environment

Kyoung Shin Park and Robert V. Kenyon
Electronic Visualization Laboratory
University of Illinois at Chicago, Chicago, IL 60607, USA
(312) 996-3002 voice, (312) 413-7585 fax
kpark@eecs.uic.edu, kenyon@uic.edu

Abstract

We assessed the effects of network latency and jitter on a cooperative tele-operation task in a collaborative virtual environment. Two remote partners worked together to manipulate shared virtual objects over a network. The task was to minimize the time to transfer a ring through one of four paths with the least number of collisions. The performance of human subjects was measured and analyzed quantitatively as a function of network latency: 10 and 200 msec delays with and without jitter. Jitter had the greatest impact on coordination performance when the latency was high and the task was difficult. These results are discussed in light of current and future CVE tasks.

1. Introduction

Collaborative virtual environments (CVE) are designed to allow people in remote locations to work together over networks. People can share collaborative experiences, learn from their colleagues or teachers, work together on designing systems, or perform a complex group task through these shared virtual environments.

CVEs are applicable to cooperative spatial tasks, such as 3D architectural design and environment planning, car design and modeling, and training to repair the Hubble space telescope [7, 8, 10]. They are also useful for supporting natural spatial social skills [1, 14] (e.g. face-to-face negotiations) - people can make use of the shared virtual space as a mean of interactive negotiations with one another. They can be applied to complex cooperative tele-operation task, such as telesurgery.

In many current CVE applications, users spend most of their time navigating in 3-D space in order to move to a particular location but spend little time in performing an action, such as manipulating objects or interacting with others. Such navigation tasks generally require neither highly intensive two-/multi-way interactions nor conflict resolution among participants (e.g. colliding each other). In contrast, more complex manipulative tasks need higher interactivity, which requires high network bandwidth and short latency.

Current CVEs transmit information about their local entities to remote sites through the network so that all sites can share the same information [5, 15]. To maintain this consistency, it is important to render remote entities in real-time so that the user will not notice any difference between local and remote entities in the environment. Thus, CVEs demand a high quality of service (QoS) on the network to maintain natural and real-time interactions among users. For example, users expect an accurate visual scene of the remote object's movements to avoid collisions between their objects and those controlled by the remote partners.

As CVEs become widespread, network QoS will be a major issue [4, 5, 9, 12]. Quality of service refers to the performance guarantees on the throughput (bandwidth), network latency, and jitter. Network latency is the time it takes to get information from one site to others through the network. For example, when a user performs an action, the information about the action is transmitted over the network, and remote users will receive the result of the action after some amount of delay. Often, networks exhibit variability in delay, called jitter can which result in a jerky presentation of remote participant's actions. Hence, CVEs mostly run on a local area network (LAN) to insure the required QoS [14, 16]. Some CVE applications use the wide area network (WAN) [10, 11], but the interactivity is reduced. A high-speed network like ATM (Asynchronous Transfer Mode) has been used in CVEs [7, 8], yet access to such a high performance network is still limited to a few research institutes and companies.

This study compares two commonly used networks, Ethernet and Integrated Services Digital Network (ISDN), to examine the tradeoff of network QoS and interactivity in a CVE. Ethernet is relatively fast and routinely used to connect CVEs. ISDN is a slow but inexpensive WAN to connect CVEs. To evaluate the effects of jitter in these networks, a constant (no jitter) latency network was simulated using a fiber-optic local network, called Scramnet.

A set of motor control tasks was developed to measure the coordination between two participants in the CVE (Figure 1). The task required cooperative manipulation of objects and conflict resolution. The task contained four levels of difficulty aimed at providing the subjects with easily negotiable interactions, and inten-

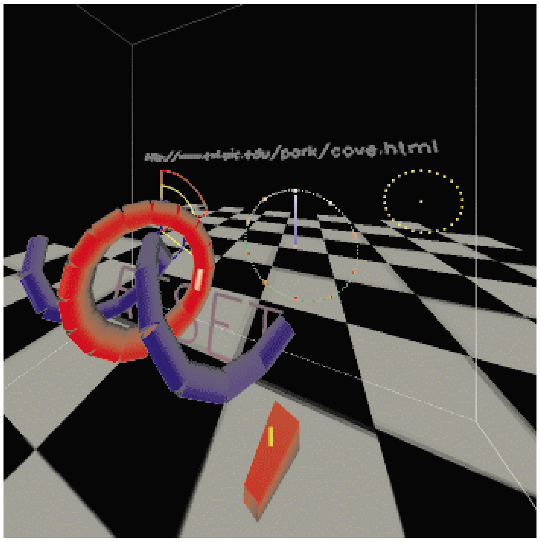

Figure 1. A cooperative tele-operative task in a networked virtual environment

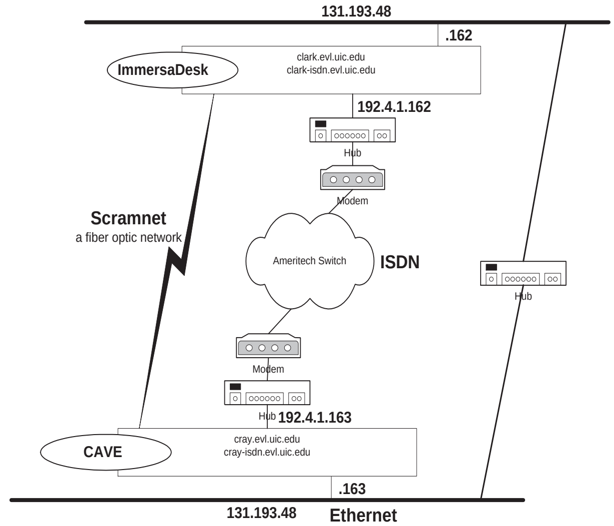

Figure 2. Illustration of computers and networks used for the experiment. CAVE and ImmersaDesk were on the same Ethernet subnet, and the IP address for ISDN was separated from Ethernet IP address. Scramnet was connected with a fiber optic.

sive hand-guidance interactions between users to assess the impact of the network on cooperative performance.

The purpose of this study is to measure changes in human performance generated by latency and jitter in the network connecting the CVEs. Furthermore, our focus is to identify human behaviors, adoptions and adaptations in relation to the constant or variable latency, and to understand the dynamic nature of the human-to-human coordination process.

2. Methods

Five teams of two subjects were formed. Each subject within the team controlled one object in each VE. All virtual objects were visible within both sites. However, the manipulation of a particular object was under the strict control of only one individual, thereby eliminating interference between the two subjects and allowing simultaneous manipulation of their object.

2.1. Subjects

Ten subjects from the University community volunteered as participants in the experiment. Their ages ranged from 23 to 60 with a median age of 27. All subjects were right-handed, had normal visual acuity and stereo perception. Subjects from the participant pool were randomly assigned to two person teams according to their schedules or were allowed to choose to work with a friend. All subjects were naive to the task and the purpose of the experiment.

2.2. Apparatus

Two tele-immersive environments (CAVE and ImmersaDesk) were connected using one of three different kinds of networks: Scramnet, Ethernet, and ISDN (Figure 2). The configuration of ISDN was 2B-channels and the dedicated services, provided by two Ameritech ISDN phonelines. The CAVE and the ImmersaDesk are high resolution, large field of view, and projection-based immersive virtual environment systems [2]. The current configuration of CAVE displays 1028 x 768 resolution stereoscopic images at 96 Hz on each surface. The ImmersaDesk is a drafting table format VE display. It features a 67x50-inch rear-projected screen at a 45-degree angle. The screen has a sufficiently wide-angle view - e.g. 110 degrees horizontal field of view when the user stands close (within 1-foot away) to the screen.

The scenes were rendered on a Silicon Graphics Onyx Infinite Reality Engine, and the position data for the user's head and hand was obtained by using Ascension Motion Star Extended Range tracking system. Users use a wand (3D equivalent of a mouse) to interact with and control virtual entities in the CAVE and ImmersaDesk. A direct voice communication was established using wireless headset microphones and speakers in the VE systems.

The shared visual scene within each VE system was rendered at the frame rate of that individual VE system. Each system maintained consistent local and remote object models by transmitting state updates across the networks. The system used the network communication of TCP/IP based client-server distributed model for Ethernet and ISDN. However, the exchange of information over Scramnet was different from the TCP/IP network protocols. Scramnet is fiber optic reflective memory network, which consists of