Supporting Transoceanic Collaborations in Virtual Environment

Tomoko Imai *, Andrew E. Johnson **, Jason Leigh **
David E. Pape **, Thomas A. DeFanti **
* Aoyama&Morikawa Laboratory
2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-8656,University of Tokyo, Japan
** Electronic Visualization Laboratory (EVL), University of Illinois at Chicago, USA
E-mail: timai@mlab.t.u-tokyo.ac.jp, ajohnson@eecs.uic.edu, jleigh@eecs.uic.edu
Tel : +81-3-5841-8980 Fax : +81-3-3818-5706

Abstract

In this paper we present a virtual mail (v-mail) system. The prototype system is developed to realize efficient "handing-off" of work between participants that are separated by great distances and/or time zones.

1. Introduction

Businessmen, researchers, and artists conduct international collaborations. They often discuss over phones without visiting their colleagues because traveling is both expensive and time consuming. There are may researches on supporting synchronous distributed collaborations [8, 2, 4]. However, when participants are in different time-zones, asynchronous collaboration is very attractive. In asynchronous collaboration, the most effective computer-based collaboration tool today is still e-mail. Despite this, there are few applications which support asynchronous collaboration in virtual environments [4, 8].

Asynchronous collaboration is the most attractive and difficult form of collaboration. It is attractive because this mode reduces traveling cost and frees people from tight meeting schedules. Its main drawback is that participants cannot quickly obtain feedback. In international collaborations, it often takes a day to receive a response. Collaborators can easily waste days just clarifying questions, plans, and instructions. Therefore, elements which can lead to ambiguity should be minimized as much as possible so that collaborators can concentrate on the tasks at hand.

What are the possible sources of misunderstanding? In the case of international collaboration, the language barrier is the largest factor. Other factors are related to the domain of the collaboration itself. For exam-ple, if the task is to design products, buildings, or art work, the explanation or clarification of ideas usually involves the description of objects and features in a spatial context. This is difficult to accomplish with traditional text-only e-mail systems, and even advanced e-mail systems that allow the inclusion of pictures, animation and audio as attachments. In these systems, the e-mail messages are still detached from the environment in which they were initially recorded. That is, although an animation sequence was recorded by pointing a video camera at a real space. The eventual viewing of the recording occurs in a detached space. There is still the need for the viewer to correlate the environment in which the recording was originally made with what is viewed on the monitor. In VE we have the opportunity to bridge this by allowing the same environment to be the medium over which a recording was originally created *and* played back. When a play-back involves the recording of the remote participant's virtual presence (an avatar) and actions (such as gestures), one can walk around the play-back of the avatar and make adjustments to the environment even while the play-back is occurring- hence treating the avatar as a surrogate of the original recorder of the message. We have designed the **virtual mail (v-mail) system** to explore mainly these issues. The goal is to develop a system that allows collaborators to enter a persistent environment and leave expressive v-mail messages (consisting of audio and gestural recordings) for one another. In the following sections we will describe our initial implementation of the v-mail system and its user testing. Finally, future directions for this research are discussed.

2   The Virtual Mail System

V-mail is a prototype system that supports both synchronous and asynchronous collaboration. The participants' mail messages and the ongoing modifications of the VE are maintained by a central server. The v-mail system is designed to be used in the CAVE virtual environment [1]. The CAVE is a 10 foot by 10 foot by 10 foot room constructed of translucent walls that are rear-projected with stereoscopic images. The participant in the CAVE sees the head tracked stereo image through a pair of LCD shutter glasses. The input device is a wand equipped with a magnetic tracker and a joystick. The interaction with the VE is done through the user's head position and orientation; wand's position and orientation; and the wand's three buttons and analog joystick. The graphics for the CAVE are driven by a Silicon Graphics Onyx2^{tm}.

V-mail is built on top of the CAVE library, a higher level graphics library, called XP (eXtended Performer), and the CAVERNsoft networking library. These libraries are explained in the following subsections.

2.1   Audio/Gestural Recording and Playback

V-mail's user-interface is embodied in a virtual friend or pet that follows the participant as he/she interacts with the VE. Touching the pet pops up an interface for recording, reviewing, and deleting messages. This pop-up menu may either follow the user as with the virtual pet, or it may be set to a stationary location in the environment. Allowing the pet to follow a participant constantly has the advantage that the interface for recording and stopping the recording can be reached anywhere in the VE. However in some instances the pet may actually occlude part of the VE and hence the option is provided to leave the pet behind. At a touch of a button the pet can be instantly recalled to the participant's side.

To record a message, a user begins by selecting the virtual pet. In response, a microphone, mailbox, and delete icon appears. If the user selects the microphone, the recording, stop, and playback icons appear. The user's voice is picked up by a wireless microphone worn by the user and digitized by the Onyx2. A sound server stores the digitized data in an audio file. At the same time, tracked head and wand data, consisting of position and orientation information, coupled with timestamp data are stored in a gesture file. This timestamp information is used during playback to ensure that the audio is synchronized with the re-animation of the gestures.

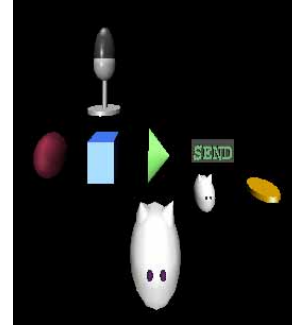The recording of both voice and gesture stops when

Figure 1. Receiver list under the send icon.

the user selects the stop button. The selection of the send button pops up a list of the virtual pets belonging to the other participants (Fig. 1). V-mail can be sent to the owner of any of the pets by selecting the pet under the "SEND" icon. The sent v-mail is stored at the central server to await download from its intended recipient.

2.2   Graphics Support for V-mail

The v-mail client applications use XP (Extended Performer System) for rendering the graphics [9]. XP builds on a set of existing toolkits (the CAVE library, IRIS Performer^{tm}, OpenGL^{tm}) and provides basic modules for user-interaction, lighting, sound, navigation, and loading geometric objects, and their behaviors. A performer scene graph is constructed from XP's **scene file**, which is written in a readable script language. The scene text specifies position and behavior of geometric and sound objects. First, the root class parses the scene file and constructs a Performer scene graph. Then, constructed nodes start sending messages between them. An event which triggers messages can be defined by a programmer, such as a user's button press and proximity. Different classes recognize different messages.

2.3   Network Support for V-mail

The clients communicate with the server using CAVERNsoft[6], which facilitates the construction of persistent collaborative virtual environments [5]. CAVERNsoft manages distributed data-stores. The data distribution between a client and a server application is realized by the Information Request Broker(IRB). The IRB is a repository of persistent data that is accessible by a variety of networking interfaces. It emulates a persistent distributed shared memory system which can deliver data over both reliable and unreliable networking protocols.

2.4 Message Exchanging across Network

The messages exchanged across the network are an extension of messages sent among local XP classes. The server class recognizes messages required to keep persistence of the VE. Here, the persistence of the VE means consistency among the clients' states. XP classes understand messages to control local events. The server dispatches messages among the clients through communications channels provided by CAVERNsoft, then, XP delivers these messages to the appropriate XP classes locally. The message consists of a **command** and a **argument** fields. On receiving a message, appropriate actions are taken according to the command type.

There are three kind of messages, initialization, termination, and run-time messages. The initialization message is sent when a user starts a client process. It uses the reliable TCP channel of CAVERNsoft and each message consumes 256 bytes. First, the user's avatar is switched on at all the currently participating clients and a list of active participants at the server is updated. Then, mail messages stored in the server are downloaded. Finally, a list of account names is sent. When the user leaves the VE, his/her avatar is switched off and the client's name is deleted from the active participants list.

Run-time messages are sent after the initialization messages. Currently there are two kinds of run-time messages, a **mailing message** and a **tracker message**. The mailing message consists of the sender's name, size of the audio and gesture files, followed by the raw data of the recording. When a client sends a mailing message to the server, the server delivers the message to the receiver's mail box on the server. The mail is downloaded when the receiver enters the VE. The tracker message is sent through a dedicated unreliable UDP channel. Its arguments are head position and orientation, right hand position and orientation. The arguments can be easily extended to include left hand information.

3 User Testing

3.1 Conducted User Test

To evaluate v-mail, a qualitative user test was conducted employing the observation and the thinking-aloud method[7]. The thinking-aloud method requests that the user describe their actions while they are being performed. Three pairs of test users used the v-mail system. The current version of the v-mail system supposes a class of users who have a prior experience of working in a VE.

In the test, two people collaborate asynchronously. One person is assigned to play the role of a designer and the other person is assigned as the implementer of the design. The scenario is to design a virtual museum using the eight provided geometric objects. Firstly, the designer places the geometric objects in the virtual museum. Then, he/she sketches the design on paper. The sketch is given to the implementer and the designer leaves the test room. The VE is re-initialized and the implementer is asked to recreate the design from the sketch. When the implementer finishes moving the geometric objects he/she records messages for the designer using v-mail and leaves the environment. Next, the designer returns to the VE to view the v-mail and to examine the implementor's re-creation. This time, the designer is not allowed to move any objects. The designer examines the virtual museum arranged by the implementor and leaves a v-mail to suggest alterations. Then, the implementor returns to the virtual museum to make the corrections that were suggested in the designer's v-mail. This entire process of critique and adjustment repeats for a total of 3 cycles.

After this role playing, users answer a questionnaire and are interviewed. The entire process, including the interaction in the CAVE, was recorded on video tape.

3.2 User Test Results

3.2.1 Recording V-mail

An example of the gesture usage observed is as follows; A user pointed to the object which she/he wanted to move. The user acknowledged the proper placement of an object by circling the area around the object. The following is an excerpt from a recorded message of an implementer attempting to clarify what she is seeing to the designer: "You indicated two doors but I can see three doors, One is here (pointing to the first door) (Fig. 2). One is here (pointing to the next door). The third is here (pointing to the third door). So, I chose this area (moving her virtual hand over the area), between these two doors. Under the globe (pointing the place) you showed another instrument but I can't find it."

3.2.2 Replaying V-mail

When a user replayed the v-mail, a messenger avatar talked and moved to re-enact the message. There were three ways users interacted with the replayed v-mail. The common behavior was to start moving objects after listening/watching the entire v-mail. However when the designer talked slowly, the implementor could use

Figure 2. An avatar pointing to the door and a receiver looking into the direction.

the opportunity to move the object to newly prescribed location. Finally, one user used the playback to confirm what he did to the environment. The users used two kinds of relative standing positions to the messenger avatar. While the avatar was talking to the user, the user stood in front of the avatar. When the messenger avatar started explaining how to move objects, the user stood next to the avatar so that he/she could see the same object from the same view point. These two modes are also seen in daily life and are very natural behavior. The users treated the messenger avatars as if their collaborators were actually present in the collaboration. Inadvertently collaborators would find themselves talking to the messenger avatar.

4   Future Work

We were encouraged by the favorable results of our v-mail prototype. Participants seemed to be able to effectively use voice and gesture to asynchronously communicate ideas to one another. It will be interesting to observe the degree of gesture usage amongst international collaborators from differing cultures.

The most interesting finding in v-mail was the way viewers of the v-mail messages tended to treat the messenger avatar as if they were real participants. This might suggest that v-mail can have a significant impact in bridging time-zone differences.

Other possible improvements could be to allow v-mail to be attached to particular objects in the space, as in the earlier work done by Harmon [3]. Also the recording techniques created in v-mail can be used to create virtual training environments where appropriate pre-recorded sequences can be played back at appropriate times based on user input. We will continue to conduct transoceanic collaboration to develop and evaluate software for both synchronous and asynchronous collaboration.

References

[1] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti. Surround-screen projection-based virtual reality: The design and implementation of the $CAVE$. In *Proceedings of ACM SIGGRAPH '93*, pages 135–142. ACM Press, 1993.

[2] J. M. S. Dias, R. Galli, A. C. Almeida, C. A. C. Belo, and J. M. Rebordão. *mWorld* : A multiuser $3D$ virtual environment. *IEEE Computer Graphics and Applications*, 17:55–65, March-April 1997.

[3] R. Harmon, W. Patterson, W. Ribarsky, and J. Bolter. The virtual annotation system. In *Proceeding of the IEEE VRAIS'96 Conference*, pages 239 – 245. IEEE Computer Society, IEEE Computer Society Press, March 1996.

[4] W. Lamotte, E. Flerackers, and F. V. Reeth. Visinet:collaborative $3D$ visualization and VR over ATM networks. *IEEE Computer Graphics and Applications*, 17:66–75, March-April 1997.

[5] J. Leigh, A. E.Johnson, and T. A. DeFanti. $CAVERN$: A distributed architecture for supporting scalable persistence and interoperability in collaborative virtual environments. *Journal of Virtual Reality Research*, 2.2:217–237, December 1997.

[6] J. Leigh, A. E.Johnson, and T. A. DeFanti. The CAVERN web page: http://www.evl.uic.edu/cavern, 1998.

[7] J. Nielsen. *Usability Engineering*. Academic Press, Inc., 1993.

[8] A. Pang and C. Wittenbrink. Collaboraive $3D$ visualization with $CSpray$. *IEEE Computer Graphics and Applications*, 17:32–41, March-April 1997.

[9] D. E. Pape, T. Imai, J. Anstey, M. Roussos, and T. DeFanti. Xp: An authoring system for immersive art exhibitions. In *Virtual Systems and Multi-Media*, November 1998.