

Development of a Tool to Facilitate Observational Analysis of Collaborative Virtual Environments

Abstract: The principal purpose of this project was to develop a tool to facilitate observational analysis of collaborative virtual environments. The project also included a detailed observational analysis using this tool, which was part of an evaluation of a virtual environment for training first responders to weapons of mass destruction. Observational requirements were established that necessitated a tool that met the following requirements: 1) Simultaneous dynamic observation of both participants in and out of the virtual world; 2) Observations of interactions between participants; and 3) Real time view of autonomic response synchronized with other observational data. A review of existing tools indicated that no tools met these criteria, so a tool was developed that met these criteria. This paper will discuss the process and issues encountered in developing the tool, as well as a summary of the results of the analysis of the exercises.

Background and Rationale

Context

Researchers at the University of Missouri – Rolla (UMR) developed a collaborative virtual environment to train first responders to deal with terrorist attacks involving the use of weapons of mass destruction. This is referred to as the First Responder Simulation and Training Environment (FiRSTE) project (Berry & Hilgers, 2004; Hall et al., 2004; Hilgers et al., 2005; Leu et al., 2003; Misra, Decker, Barker, & Hilgers, 2004; Tichon, Hall, Hilgers, Leu, & Agarwal, 2003; Wilfred et al., 2004) (Also, see <http://campus.umar.edu/firste>). The project was funded by the U.S. Army's Tank-automotive and Armaments Command (TACOM, grant # DAAE07-02-C-L068). As a part of this project, first responders were given scenario based exercises to test the environment and its interfaces. An evaluation was performed to determine if the first responders were able to effectively perform their tasks using the system. The main goal of the evaluation was to address the following research questions: 1) Did the environment provide meaningful and realistic training?; and b) What factors mediated the degree to which the environment provided meaningful and realistic training?

Three sets of exercises were carried out using ten participants from the surrounding community who would serve as first responders if an emergency were to occur. The collaborative environment supported two responders interacting in the same virtual environment; therefore, each exercise consisted of two participants in the environment and one incident commander who communicated with the participants while viewing a screen that displayed the activity in the virtual environment. Responders wore emergency gear including oxygen tanks and masks. They interacted with the environment via an interface that included a head mounted display, locomotion devices (either a treadmill or a navigation pad), other devices for controlling and detecting location within

the virtual environment, devices that simulated chemical detection based on models of gas release in the environment, and gauges that indicated oxygen levels in their tank. The environment was developed using the *Half Life*[®] game engine. More details regarding the interfaces and virtual environment are available elsewhere (e.g., Hilgers et al., 2005). (Figure 4 includes a photo of the participants in the interface and their representation in the virtual environment).

During the evaluation field notes were taken, which provided detailed real-time information regarding responder interactions as well as data regarding the frequency of communication between the responders and the incident commander. After the exercises were finished, further analysis could only be made from other data recorded during the exercise. The exercises incorporated a range of data collection methods in an attempt to capture the events for later viewing including: a) videotape of responders interacting with the interface; b) video of responders' activity within the virtual environment; and c) the heart rate of the first responders at 15 second intervals (measured via a chest heart monitor and sports watch).

Observational Requirements

The challenge that came up while determining how to evaluate the system was how to make these separate data sources meaningful. The issue stems from the fact that the data were recorded as a particular data type for a particular responder. An example would be responder "A" video footage from exercise 2. When looking at any one component in particular, one can only get an isolated view of what was occurring at the moment. Take for instance a videotape of a responder. If something unexpected occurs such as a visual indication of navigation loss, there would be no way to determine what had happened in the virtual environment. Likewise, in the case of a virtual environment viewpoint it would be impossible to determine if the responder was experiencing difficulties or was simply standing still. It was determined that a few key observations had to be made to evaluate the system effectively. These observations consist of how the responders' interactions correspond to the virtual environment as well as how the responders communicated with each other. To accomplish this task, both of the responders would have to be displayed simultaneously along with their virtual environments. More specifically, this led to two important requirements for a tool to facilitate the observational analysis: 1) Both responders should be viewable in and out of the virtual world simultaneously; and 2) The communication between the first responders should be viewable, synchronized with other activities.

These requirements related to all of the sources of data recorded with the exception of the heart rate data. The heart rate data was collected because a fundamental criterion for realistic training, specified at the beginning of development, was that the exercise should be physically taxing. In addition, real time collection of heart rate data could potentially provide information about the stress and fatigue associated with given activities and events within the environment. This would only be possible if the heart rate data could be viewed in conjunction with the video streams. Therefore, a third requirement for the observational tool was that it provide a viewable dynamic

representation of heart rate data synchronized with the combined video streams specified in the first two requirements, described above.

Existing Methods and Tools for Observational Analysis

One of the most common tools for observation of human interaction with information technology is the usability lab (Barnum, 2002; Dumas & Redish, 1999). The lab is set up with a participant room where the activity takes place and a separate observation room. Usability labs usually consist of one or more cameras to capture the user along with a scan converter to capture the computer screen. These video sources are then fed into a piece of hardware called a multiplexer or an advanced mixer. The mixer allows each source to be positioned and scaled arbitrarily. A usability lab was not feasible for this evaluation due to space constraints. Further, the traditional usability lab setup is based on a user seated in front of a computer monitor, whereas the user-interface interaction was much more complex in this research. Therefore, other methods were examined.

While the traditional usability lab is used for evaluation of non-immersive software interfaces, there have been a wide range of evaluation techniques used specifically for collaborative virtual environments. Unfortunately, none of those reviewed, addressed the observational needs of this project. In some instances, recordings were not made and qualitative post-analysis was impossible (Gerhard, Moore, & Hobbs, 2001; Mortensen et al., 2002). Instead, all of the observations were done during the live test (Goebbels, Lalioti, & Gobel, 2003; Jackson & Fagen, 2000). Most often, the user's voice was recorded along with screen capture of the environment using a VCR or video capture software (Hindmarsh, Fraser, Heath, Benford, & Greenhalgh, 2000; Hosseini & Georganas, 2002). Advanced techniques included both videotaping the users and recording the environment using functionality built into the 3D software or a custom tool (Greenhalgh et al., 2000). This functionality ranged from different camera angles or playback speeds of the user's interactions with the environment to visualizing the paths taken by users.

Using a typical usability lab or the most advanced VR evaluation techniques reviewed, a user and the environment are displayed simultaneously using a picture-in-picture display (Figure 1). This picture-in-picture method is able to solve the first challenge of being able to see both the user and the environment. However, it is limited to only one user. Therefore, we concluded that an integrated 4-video panel could be developed that displayed users both in and out of the virtual environment simultaneously as illustrated in Figure 2. In the next section, I will discuss the process and issues associated with the development of an observational tool based on this format.

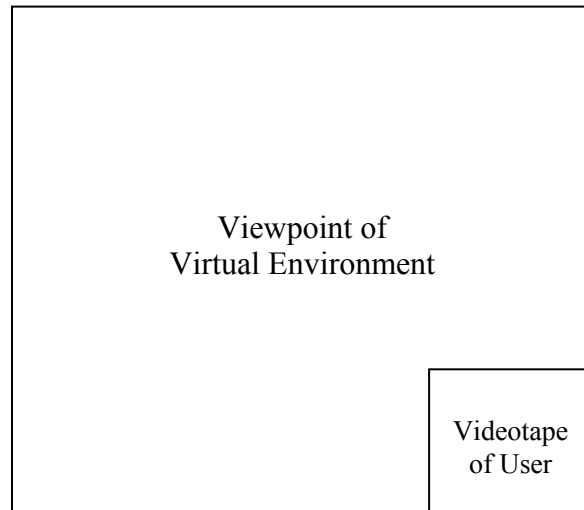


Figure 1: The format for a typical usability picture in picture video.

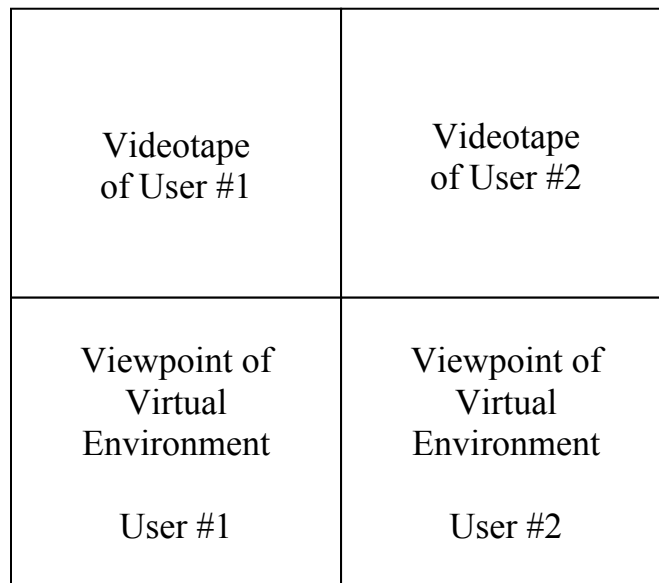


Figure 2: The format for our proposed 4-way video for observing collaborative Virtual Environments

Tool Development

Evaluation of Hardware and Software Solutions

The first challenge of developing this tool was finding hardware or software that could produce a four way video from four separate sources. It was previously mentioned that a hardware video mixer or multiplexer could be used. After an exhaustive search of consumer and professional level hardware, no suitable hardware could be found. Most video mixers on the market were primarily for video jockey or editing applications and could only display two sources at a time. The Laboratory for Information Technology Evaluation at UMR has a *Videonics MX Pro*[®] mixer, which can display up to four sources simultaneously in a special composition mode, but the functionality was meant for still

images only and thus resulted in sluggish video at roughly five frames per second. The other hardware solution examined was video multiplexers which are popular for multi-camera surveillance applications. The problem with these is that they are primarily made for the security industry and thus are designed to connect to analog cameras. Since the inputs were analog, a loss of quality would result from the digital to analog conversion. Also, each of the multiplexer's inputs typically had a resolution of QVGA (320x240) which was a significantly lower than the resolution of the original data sources (1280x1024 virtual environment, 720x480 video footage).

Since no hardware solution could be found, software was then examined. There are many types of video-editing software available, but all of the consumer grade packages focused primarily on video transitions and effects. Simultaneous multi-source functionality was not found in this level of software but was featured in more advanced software packages. The only consumer level product with the appropriate functionality that we reviewed is *Adobe Premiere Pro*[®], which allowed positioning and resizing of many video sources. An educational version was purchased for use in this evaluation.

File Conversion

The next challenge was to convert the data sources into a format that was compatible with *Adobe Premiere Pro*. Figure 3 illustrates the process that I followed, which I will describe below.

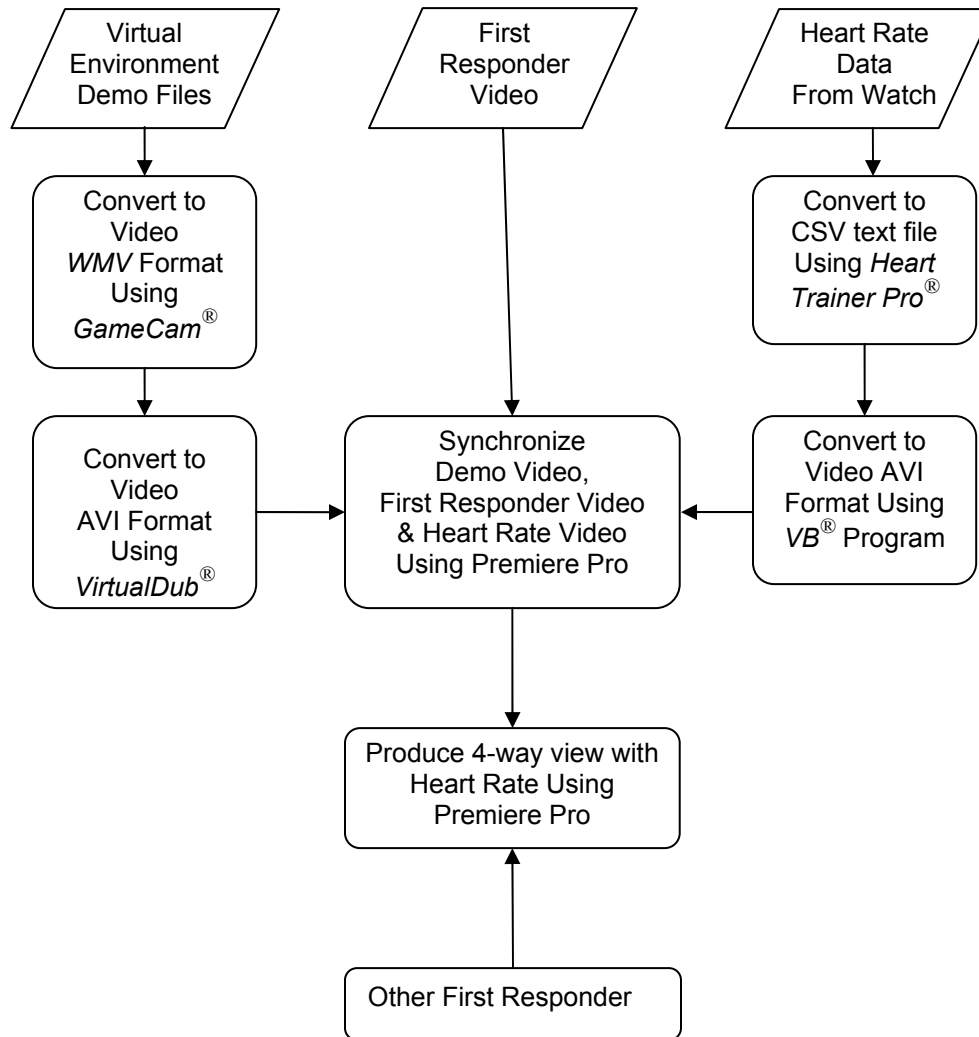


Figure 3: Illustration of conversion process

The videotaping of the first responders was done using digital video cameras so the footage was already in a compatible format. It was transferred to a hard drive via IEEE 1394 (Firewire[®]) using *Adobe Premiere Pro*[®]. The virtual environment was in an incompatible format and had to be converted, since it was developed using the game engine, *Half-Life*[®]. *Half-life*'s proprietary demo recording functionality was used to capture the virtual environment during the experiment. These demo files could later be played back using *Half-Life*. Fortunately there are programs designed to capture video from 3D games such as *Fraps* (www.fraps.com) or *Game Cam* (www.planetgamecam.com). *Game Cam Lite* was used to record the video from the demo files as they were being played back using *Half-Life*.

The last thing that had to be converted was the heart rate data obtained using *CardioSport*[®] heart rate monitors. Heart rates were downloaded from the watches into *CardioSport's Heart Trainer Pro*[®] software after the experiment using *CardioSport's PC Link*[®] cable. In order to be converted to video, the data had to be displayed at the same

rate as it was recorded. To display the heart rates, a program was created by a team member of the Laboratory for information technology evaluation (LITE) which would display comma separated value (CSV) text files at a certain frequency. The display of this program was then captured using screen recording software, *Camtasia Studio*[®] (www.techsmith.com).

File Synchronization

At this point all of the data had been converted to the appropriate video format, but it still had to be combined using *Premiere Pro*[®]. The first step was to synchronize the videos of the responders. This was done manually, by listening carefully to the audio of the first responders speaking with one-another. Since the responders were in the same room during the exercise, both videos contained the same dialog. Certain unique phrases served as reference points in each video for a particular exercise. *Premiere Pro*[®] was used to mark these points and then the videos were shifted in the timeline so that the markers overlapped. Overlapping the markers effectively made the responders videos in sync and playback was simultaneous.

The same technique was used for the virtual environment but instead of audio cues, video was used as reference. As long as one of the responders could see the other's actions, the videos of the environment could be synchronized based upon these actions. Unique points in the virtual environment were identified when the responders performed activities such as looking at each other, dropping tarps, and opening doors.

Now that the responder videos and environment videos were playing back correctly, the video of the responders could be synchronized with their environments. The primary reference between the environment and the responder was the responder's action and the environments response. There were times that the actions of the responders did not have expected results, but the treadmill input device did not experience any of these problems. The easiest way to determine a reference point was when the video of the responder started before the responder's first movement. If this was the case, it was simply a matter of finding the first instance where the responder moved the belt of the treadmill and the avatar moved in the virtual environment. Sometimes it was harder to tell when a particular action was causing a reaction so a cue had to be made from the dialog. One instance of this would be, for example, reporting a room number to the incident commander. Usually the responder had to find the room number when asked what room they were in and thus reported it as they read it from the door in the virtual environment.

The last set of videos to be included in the synchronized video panel was the heart rate data. In the third set of experiments, synchronization was based on when the recordings of the heart rate watches started. The responder videos had captured the exact instance the heart rate was started by starting earlier than the heart rate. In the second set of exercises, this was not the case and other methods had to be used. Heart rate was determined in this instance using physical exertion of the responders. It was determined that the physical activity of the first use of the input devices would raise heart rate

noticeably. Therefore, when the responders started using the devices the first increase in heart rate was used as a reference point, since it was consistently dramatic and noticeable. Due to the heart rate monitor's sampling rate of fifteen seconds, it was impossible to obtain accuracy greater than +/- 15 seconds in this case. Accordingly, the heart rate display was updated every 15 seconds. Figure 4 is a screen shot of the final video panel used for the observational analysis. Note the heart rates in the lower corner of the videos of the first responders interacting with the interface.



Figure 4: Screen Shot of the Video Panel Used for Observational Analysis

Observational Analysis of Videos: Results Summary

These 4-way videos were then used as an observation tool for qualitative observation as well as coding of behavior into categories. These analyses, which are described in detail elsewhere (Hilgers et al., 2005), indicated that the responders were performing realistically in the environment. They were able to complete tasks according to protocol but also made mistakes such as losing track of the building entrance or not keeping protocol. These realistic mistakes could happen in the real world and were not caused by hardware or software difficulties. It was noted that most of the time the responders were in character, immersed in the environment exhibiting a high degree of presence.

Communication was also observed in detail. The responders were in constant communication with each other and the incident commander. However, the responders sometimes lost sight of each other or did not know where to go. Usually they lost sight of each other due to lack of communication when they initially became separated.

The analyses indicated that the primary cause of unrealistic activity was due to technical problems with the hardware or software. Sometimes one of the input devices would not provide adequate locomotion in the environment or the wires would become tangled. When the wires became tangled, the experiment had to be paused for roughly twenty seconds while the cords were unwound. Also, the input device had to be recalibrated occasionally when the locomotion was incorrect. The few other technical problems were software based. In two of the ten exercises, the virtual environment software failed and had to be restarted. In one event, a responder had problems with doors pushing his avatar around or getting stuck behind them. Other times there was no way for a responder to perform particular activity such as drop a blanket. Intervention was required and a computer operator was required to press a key to complete the task.

The results of the observational analysis indicated that, overall, the experience was realistic. The majority of the time participants were able to communicate with each other and interact with the virtual environment. In these cases, the environment provided a realistic and effective training experience. However, in cases when problems such as hardware or software failure occurred, the effectiveness decreased notably.

References

- Barnum, C. M. (2002). *Usability Testing and Research*. New York, NY: Allyn & Bacon.
- Berry, A. J., & Hilgers, M. G. (2004). Modeling First Responder Emergency Sensors for Virtual Training. *Proceedings of the 15th IASTED Modeling and Simulation Conference.*, 234-239.
- Dumas, J. S., & Redish, J. C. (1999). *A Practical Guide to Usability Testing*. Exeter, England: Intellect.
- Gerhard, M., Moore, D. J., & Hobbs, D. J. (2001). Continuous Presence in Collaborative Virtual Environments: Towards the Evaluation of a Hybrid Avatar-Agent Model for User Representation. In Springer-Verlag (Ed.), *Lecture Notes in Computer Science* (Vol. 2190, pp. 137-155). London, UK.
- Goebels, G., Lalioti, V., & Gobel, M. (2003). Design and Evaluation of Team Work in Distributed Collaborative Virtual Environments. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, 231 - 238.
- Greenhalgh, C., Purbrick, J., Benford, S., Craven, M., Drozd, A., & Taylor, I. (2000). Temporal Links: Recording and Replaying Virtual Environments. *Proceedings of the ACM International Conference on Multimedia*, 67 - 74.
- Hall, R. H., Wilfred, L., Hilgers, M. G., Leu, M. C., Walker, C. P., & Hortenstine, J. M. (2004). Virtual Terrorist Attack on the Computer Science Building: Design and Evaluation of a Research Methodology. *Presence-Connect*, 4(4), <http://www.presence-connect.com>.
- Hilgers, M. G., Leu, M. C., Hall, R. H., Agarwal, S., Lambert, T., Albright, R., & Nebel, K. (2005). Virtual Environments for Training First Responders. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference*.

Hindmarsh, J., Fraser, M., Heath, C., Benford, S., & Greenhalgh, C. (2000). Object-Focused Interaction in Collaborative Virtual Environments. *ACM Transactions on Computer-Human Interaction*, 7(4), 477 - 509.

Hosseini, M., & Georganas, N. D. (2002). MPEG-4 Based Recording and Replay of Collaborative Virtual Reality Sessions. *Proceedings of the IEEE Virtual Reality Conference*, 271 - 272.

Jackson, R. L., & Fagen, E. (2000). Collaboration and Learning within Immersive Virtual Reality. *Proceedings of the International Conference on Collaborative Virtual Environments*, 83 - 92.

Leu, M. C., Hilgers, M. G., Agarwal, S., Hall, R. H., Lambert, T., Albright, R., & Nebel, K. (2003). Training in Virtual Environments for First Responders. *Proceedings of the ASEE Midwest Section Meeting*, http://campus.UMR.edu/firste/papers/leu_asee.pdf.

Misra, A., Decker, A., Barker, B., & Hilgers, M. G. (2004). Modeling Chemical Events within Complex Buildings for First Responder Virtual Training. *Proceedings of the 15th IASTED Modeling and Simulation Conference.*, 240-245.

Mortensen, J., Vinayamoorthy, V., Slater, M., Steed, A., Lok, B., & Whitton, M. C. (2002). Collaboration in Tele-Immersive Environments. *Proceedings of the Workshop on Virtual Environments*, 93 - 101.

Tichon, J., Hall, R. H., Hilgers, M. G., Leu, M. C., & Agarwal, A. (2003). Education and training in virtual environments for disaster management. *Proceedings of World Conference on Educational Multimedia, Hypermedia, & Telecommunications (EdMedia)*, 1191 - 1194.

Wilfred, L. M., Hall, R. H., Hilgers, M. G., Leu, M. C., Hortenstine, J. M., Walker, C. P., & Reddy, M. (2004). Training in Affectively Intense Learning Environments. *Proceedings of the ACE E-Learn Conference*, 2233 - 2240, <http://dl.ace.org/16761>.

Acknowledgements

This project was supported in part by the U.S. Army's Tank-automotive and Armaments Command (TACOM, grant # DAAE07-02-C-L068), and the Laboratory for Information Technology Evaluation (LITE) (<http://campus.UMR.edu/lite>) at the University of Missouri – Rolla.