VR-CAVE AS A MULTIMODAL INTERACTION ENVIRONMENT

WINGLY SHIH
Graduate Institute of Architecture, National Chiao Tung University
1001 Ta Hsueh Road, Hsinchu 300, Taiwan.
swingly@arch.nctu.edu.tw

Abstract. This research focused on building a software system on VR-CAVE for multiple user interaction with various media, alongside simultaneous communication amongst users. While all users interact with virtual objects in VR-CAVE concurrently, by means of several kinds of input device, including conventional mouse and keyboard, PDAs, mobile phones and laptops, they can communicate with each other through the most natural way at the same time. The system prototype was implemented and is still in development for extension on the strength of supporting more input device and media formats. Further research would make VR-CAVE not only a representing tool, but a platform for design communication and evaluation, or a collaborative design environment.

1. Introduction

Due to the continuous expansion of digital media adoption, more and more designers nowadays are familiar with employing various digital tools. Among all the digital tools, VE (Virtual Environment) has the ability to simulate 3D virtual environment, thus becoming an increasingly important tool for architectural design (Tang, 2003). VR-CAVE, a kind of VE, is proficient in providing stereoscopic visualization, immersive experience and multi-user navigation (Cruz-Neira et al., 1993), making it a widely-used visualization tool to deliver virtual space simulation for design representation and communication. However, its distributed complex infrastructure constrains interaction on VR-CAVE to walk through or navigation in space, and is only available for a single-user control, leaving other participants no opportunity to have direct interaction with the media concurrently.

There are some kinds of meeting, including brainstorming sessions and design reviews, where designers communicating with each other and a computer display slides, images or movies. When a computer is used as part of the presentation, a most common scenario that occurs is that different attendee would like to take turns controlling the mouse or keyboard connected to it. They might want to manipulate the media under consideration for more investigation, make their annotations or flip slides for specific information.

Most computer software built today is for single-user application; therefore a distinct software framework needs to be developed. A software system for multiple users interacting concurrently within VE was implemented on the VR-CAVE situated in National Chiao Tung University, Taiwan. This paper describes the system framework for making VR-CAVE a multimodal interaction environment and discusses the implementation and further development of the software system.

2. Related Work

Multiple access to the shared medium and communication among users was treated as the key subject in Computer-Supported Cooperative Work.

Achten (2002) outlined the requirements of a collaborative design environment: give participants the ability to tackle all design problems, provide awareness of the presence of other participants, allow participants to present themselves in various manners, allow flexible and open access to the design data, provide representations of information and participants that are specific for collaborative design, make participants aware of the shared goals of the project, make the design environment such that it resembles more a medium than a tool. Several communication issues were also mentioned for further research in collaborative design, like communication language, communication environment and communication behaviour.

The Distributed Interactive Virtual Environment (DIVE) is an internet-based multi-user VR system where participants navigate in 3D space and see, meet and interact with other users and applications. DIVE supports the development of virtual environments, user interfaces and applications based on shared 3D synthetic environments. DIVE is especially tuned to multi-user applications, where several participants interact over a network (Hagsand, 1996).

The MUG (Multiuser Groups for Conceptual Understanding and Prototyping) lets a team of designers—collaborating over the Internet in a shared, multimodal, 3D workspace—model a 3D layout and semantically grounded behavioural description of a product or device. The MUG prototype provides several modalities for interaction. Users can control and model individual shapes in the workspace. One example is of multiple users, operating at a distance, manipulating the same non-uniform rational B-splines (NURBS)-based surface. Designers can work and converse in the digital design space (Cera et al., 2002).

Despite the fact that the networked VE can provide the environment for multi-user interaction with digital media, users cannot interact with each other through conventional communication methods, such as facial expression, body gesture, and behavioural observation.

Rekimoto and Saitoh (1999) developed a collaborative system, Augmented Surface, built with an interactive desktop and a shared projection display. The system allows users to smoothly interchange digital information among their portable computers, a computerized table and wall.

The Ubi Table examined the design space of tabletops as scrap displays. Scrap displays support kiosk-style walk-up interaction for impromptu face-to-face collaboration. This design offered the affordability of a physical table. It provided flexibility by allowing users to layout shared documents with desired orientation and position; at the same time it augmented traditional paper-based interactions by providing a flexible gradient or shades of sharing semantics (Shen et al., 2003).

SDG (Single Display Groupware) is defined as computer programs that enable co-present users to collaborate via a shared computer with a single shared display and simultaneous use of multiple input devices. It enables the co-located users to communicate with each other in the most natural way, and gives each user the ability to interact with digital media in the computer (Stewart et al., 1999).

The Pebbles project held in Human–Computer Interaction Institute in the School of Computer Science of Carnegie Mellon University is exploring the many ways that small handheld PDA (Personal Digital Assistant) can serve as a useful adjunct to the “fixed” computers. One set of applications supports meetings where the participants are co-located. All participants’ PDAs are in continuous two-way communication with each other, and with the main computer which is often projected on a screen to serve as the focal point of the discussion (Myers, 2001).

3. The Multimodal Interaction Virtual Environment

VR-CAVE is a distributed system comprising several computers. All computers in the system communicate via LAN (Local Area Network). A server is necessary for synchronization purposes, or the disordered frame from different client computers

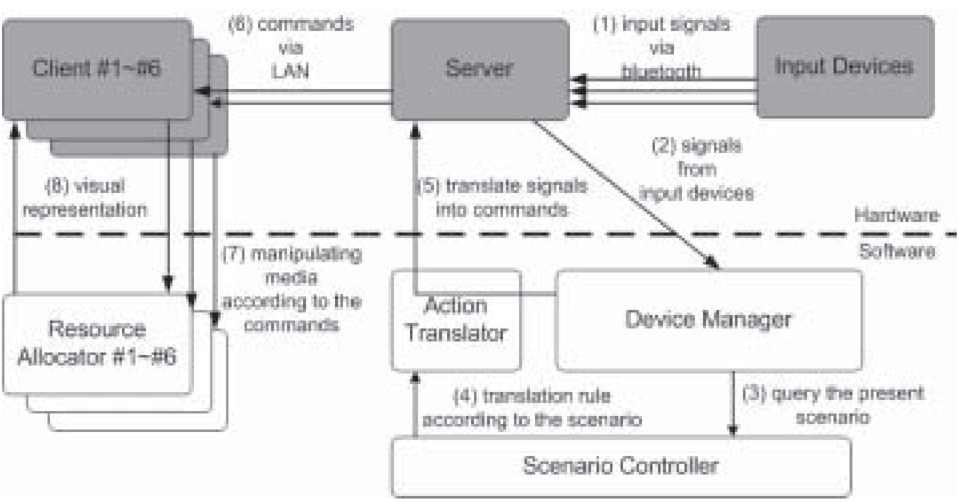

Figure 1. The framework of the system.

would not form the stereoscopic visualization correctly. Consequently, the software framework needs to be a client-server infrastructure, and all input devices have to be connected with the server. The hardware part of the system comprehends several input devices, one server and six clients (for stereoscopic visualization on 3 walls), and the software part is dividing into four modules: device manager, action translator, scenario controller and resource allocator. Figure 1 shows the scheme of the system framework and the work flow among hardware and software.

3.1. HARDWARE

3.1.1. Input devices

For the sake of avoiding wire annoyance, a state-of-the-art wireless technology, Bluetooth is picked as the cable replacement solution. There are many kinds of Bluetooth-enabled devices, such as laptops, PDAs, cellular phones and headsets, and most of them have capability of computing. Besides, these Bluetooth-enabled devices are coming into wide use as personal communication appliances and are easily available. From the above, diversified input methods including text, voice, pen, trigger, coordinates-oriented signals and file inter-transfer are still probable options. Without special-purpose hardware, users simply use their own handy communication appliances to interact within VR-CAVE and freely connected/ disconnected with the system.

In addition to all of the above, conventional wired input devices like webcams, joysticks and mice are still usable to connect with this system without any conflict. Up to now, some kinds of input devices are supported and the variety is enlarging with the progressive development of the system. Devices and input methods planned/ implemented to be supported are:

(Supported) cellular phone: keypad press

(Supported) PDA: click, mouse move, file transfer

(Ongoing) headset: voice recognition

(Ongoing) laptop: typewriting, screen share and all above

(Planned) USB cameras: image/gesture recognition.

3.1.2. Server

A server plays the main signal-processing role in the system. It receives signals from input devices, passes the signals in accordance with information from device manager module and then translates the signals into common commands according to command mapping rules from the scenario controller module. For example, keypad ‘*’ pressed signals from a cellular phone and click signals from a PDA might be translated into the same command as “mouse click”, or click signal from one PDA is translated into “move(0,0,1)” and the same signal from the other PDA is translated into “click”. After the translation, the server sends commands to a specific client computer via LAN. The target client the commands are to be sent to is also defined in scenario controller module.

The other onerous job of the server is the synchronization. The server sends out synchronous packets containing time information of the server every 50 milliseconds to keeping all clients synchronized.

3.1.3. Client

Each client computer is connected with a projector for projecting on to the display wall, and a waiting commands from server. While a command arrives, the client will execute the command and then rasterize the visualization result.

3.2. SOFTWARE

3.2.1. Device Manager

This module maintains identity and capability information of every input device, and keeps a list of all devices involved in the system. Input devices dynamically joined or removed from the system would notify this module and refresh the list. When an input signal is triggered, the server will identify the signal and the signal source device from this module for further processing with the scenario controller and action translator.

3.2.2. Action Translator

This module receives signals from all devices maintained in device manager module and translates signals into action command according to the capability information of each input device.

3.2.3. Resource Allocator

All media resources supported by this system, including Power-Point files, images, movies and Real-time OpenGL computational visualization are handled and rendered by this module. It receives action commands from action translator module and reacts accordingly for visual representation. The co-operation of resource allocator and action translator module makes up an abundant interaction scheme. Until now, media formats supported by the system are:

- Microsoft PowerPoint slide file
- Image: bmp, jpeg
- Movie: avi, mpg, and all file formats supported by Microsoft Media Player
- Apple QuickTimeVR panorama
- Real-time OpenGL simulating.

3.2.4. Scenario Controller

This module is the master module that coordinates the aforementional three modules.

It defines the action command as well as the response to each command.

The definition is recorded in a text file composed beforehand by the main presenter who will host the meeting/review. While the system is starting up, it will parse the file and set up itself. A sample fragment of the file looks like the following:

```
[s0001]
1.Phone(Z600)  = s0002                  //switch to scenario #2 while '1' is pressed
2.Phone(Z600)  = open(test.ppt, left)   //open media file on left wall
[s0002]
open(opengl_test.exe,all)               //open media file using whole display
1.Phone()      = select_model(1)        //specific function in opengl_test.exe
2.Phone()      = select_model(2)        //same as above but different argument
u.Phone()      = move(0,1,0)            //move selected model
```

During the meeting/review period, the main presenter might switch between different scenario and various interaction scheme providing by the system will come into being.

4. Implementation

With the implementation of the system described above, a conventional VR-CAVE only capable of navigation is turned into an environment for users to interact within.

Figure 2 shows that a user controls playback of the pre-rendered movie. The user controls the movie playback using his own cellular phone connected to the system via Bluetooth. While keypad '#' is pressed, the movie will play/pause. Keypad '2' will load the second movie clip, '3' for third and so on. This is the simplest scenario of the system.

Figure 2. The user controls playback of the movie.

Figure 3 shows the most interesting characteristic of the system. The two users holds different sort of input device, one is a cellular phone and the other is a PDA. The three display walls of VR-CAVE show different media, the left one is a PowerPoint slide, the middle one is a real-time OpenGL simulation program, and a movie is playing on the right wall. Except the PowerPoint slide, the other two are stereoscopic visualized. User (right) holding the cellular phone acts as the main presenter. She flips the PowerPoint slide to next or previous one controls the movie playback, and real-time manipulates the 3d model in the middle screen. The left user uses a PDA to manipulate the model in the middle screen, and has nothing to do with the PowerPoint file and movie.

Figure 3. Multiple users interact within VR-CAVE.

5. Conclusion and Future Work

The system revealed in this paper turned VR-CAVE into a design thinking and communication tool, instead of a representation tool. It made VR-CAVE not only a space simulator, but also a platform for designers to communicate with each other, and interact with virtual objects collaboratively. With more input and media expansions implemented, new interactions could be proposed and discussed on this system. Through integration with other researches in designing within VR-CAVE, future developments of this system will make VR-CAVE a collaborative CAAD environment.

Acknowledgement

The author would like to express deep thank to Prof. Aleppo for his comprehensive help and guidance.

References

Achten, H.H. 2002, Requirements for Collaborative Design in Architecture, Timmermans, H. (ed), *Proceedings of the 6th Design & Decision Support Systems in Architecture & Urban Planning Conference*, Ellecom, pp. 1–13.

Cera, C.D., Regli, W.C., Braude, I., Shapirstein, Y. & Foster, C.V. 2002, A Collaborative 3D Environment for Authoring Design Semantics. *IEEE Computer Graphics and Applications*, pp. 43–55.

Cruz-Neira, C., Sandin, D.J. & DeFanti, T.A. 1993, Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE. *Proceedings of SIGGRAPH'93 Conference*, ACM SIGGRAPH, pp. 135–142.

Hagsand, O. 1996, Interactive Multi-User VEs in the DIVE System, *IEEE Multimedia Magazine*, vol. 3, no. 1, 1996.

Myers, B.A. 2001, Using Hand-Held Devices and PCs Together. *Communications of the ACM*, vol. 44, no. 11, ACM Press, pp. 34–41.

Rekimoto, J., & M. Saitoh. 1999, Augmented Surfaces: A Spatial Continuous Workspace for Hybrid Computing Environments. *Proceedings of CHI'99 Conference on Human Factors in Computing Systems*, ACM Press, pp. 378–385.

Shen, C., Everitt, K. & Ryall, K. 2003, UbiTable: Impromptu Face-to-Face Collaboration on Horizontal Interactive Surfaces. *Proceedings of the Fifth International Conference on Ubiquitous Computing*, pp. 281–288.

Stewart, J., Bederson, B.B. & Druin, A. 1999, Single display groupware: A model for co-present collaboration. *Proceeding of the CHI'99 Conference on Human Factors in Computing Systems: the CHI is the limit*, pp. 286–293.

Tang, S.K. & Liu, Y.T. 2003, Space, Place and Digital Media:Towars a Better Simulation of a City that has now Disappeared. *International Journal of Architectural Computing*, vol. 1, no. 1, 2003, pp. 112–129.