Video Stereoscopic Avatars for the CAVE™ Virtual Environments

Lance Chong, Hank Kaczmarski, Camille Goudeseune
Integrated Systems Lab, Beckman Institute, University of Illinois at Urbana-Champaign
E-mail: lchong@uiuc.edu, hank@isl.uiuc.edu, camilleg@isl.uiuc.edu

Abstract

This proposed study is a design and development project based on the concept that by using stereoscopic video cameras together with a retro-reflective backdrop behind the user, it is possible to extract a CAVE™ or I-Desk™ user's stereoscopic image and integrate it into the 3D projection. This paper discusses various possible technical solutions that might make this concept work and their potential limitations. If proven workable, this concept could be an ideal first stage of bringing augmented reality into existing immersive virtual environments.

1. Introduction

Since the CAVE™ was first introduced by the University of Illinois at Chicago in 1992 [1] it has so far mostly been used to display computer synthesized 3D graphics. It could be potentially a huge usability expansion if stereoscopic videos can be integrated seamlessly with computer generated graphics in the CAVE™. A possible first step could be to use live stereoscopic video imagery of two remote CAVE™ or other VE users "to be their own avatars". Remote users can interact with each other as if they are in person, instead of using 3D avatar models. If this concept is proven workable, we will use the experiences gained in this test to conduct further experimentation in this direction. The CAVE™ can potentially be turned into an augmentation work-booth for human visual perception ideal for computer-aided remote collaboration.

There are past VR developments also considered intergrading 3D video into virtual environments such as the blue-c system developed by ETH in Zurich uses active visible lighting and "adaptive silhouettes" method to algorithmically subtract user stereo video image from dark background of a CAVE™-like environment [2]. Other previous work has successfully used motion histogram approach to track human motion through differentiating the silhouette of human figure using IR back projection onto a screen behind the user [3]. Because of their hardware design, both approaches put a lot of effort on developing software algorithm in differentiating user's video image outline from a VR hardware system background. We believe our approach of using retro-reflective material and digital infrared imagery will contribute to solving this problem more efficiently and reduce the space that will be required.

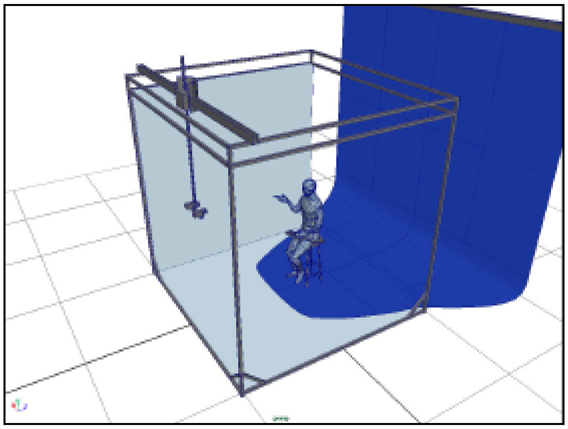

Figure 1. *Proposed layout of the Video Stereoscopic Avatar experimental setting for the CAVE™, with the retro-reflective backdrop extended to the CAVE™'s floor. A synchronized adjustable video cameras pair is in the user's front.*

2. Using "blue-screen" in the CAVE™

We propose to use a stereo video camera pair to capture stereoscopic live video of the users in the CAVE™. In order to do so, many factors need to be taken into consideration for achieving a visually comfortable final result that is usable for real applications. Above all, we need to find a way of extracting the user from the background in the video, in order to compose the user seamlessly into virtual environments. Setting up a normal blue-screen or green-screen next to the CAVE™ could be a solution. This motion picture and television special effects technique requires even and bright lighting that will very likely interfere with the CAVE™'s graphics display. We explore possible solutions next.

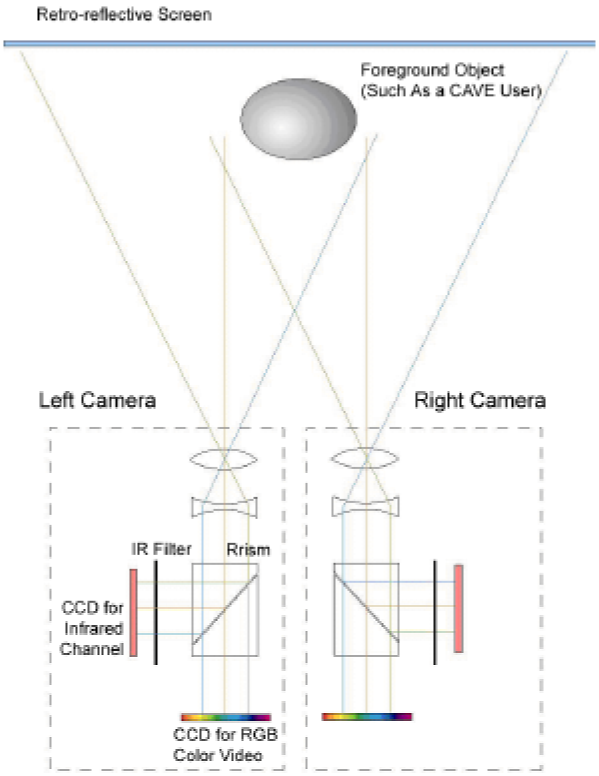

Figure 2. *Symbolic diagrams of the optical system design for using stereoscopic video inside the CAVE™.*

2.1. Retro-reflective Screen

Retro-reflective materials are designed to reflect light back to its incoming direction. If we replace normal blue-screen fabrics with a retro-reflective material, as shown in Figure 1., and place a blue or green light source close to a camera, from the camera's view, the retro-reflective screen will appear to be evenly lighted with blue or green as if in a traditional setting. Retro-reflective materials such as the "Scotchlite™" products made by 3M™ are widely available.

2.2. Infrared (IR) as invisible light source

With a retro-reflective screen, the need for intense, uniform lighting required by traditional blue-screen or green-screen can be eliminated. "White" (uncolored) Scotchlite™ fabrics are light gray, a color that matches well with the CAVE™'s projection wall and floor. But one light source of blue or green still needs to be placed near the video cameras and positioned between the user and front projection wall. In order to make the video cameras discreet, using infrared (IR) as an invisible light source is preferred. But to do so, a new set of issues must be considered.

Infrared video imagery with CCD camera and IR illuminator has been used in past VR and stereoscopic studies as ways to track user's head motion [4]. This gives us example of using IR imagery in VR settings, although our approach will only needs to track the outline of a user's image and could be algorithmically simpler.

2.3. Using IR as the alpha channel

Since we need an infrared channel as the alpha channel besides the normal visible light images in video, we will need to differentiate IR from visible light information. Normal CCD's convert full spectrum lights into red, green, and blue, three channels. They will also pick up infrared but how they interpret it into RGB channels still needs examination.

Our preliminary tests suggested that light emitted from several different kinds of infrared LED's are captured by normal CCD as "white" light. We have conducted several small-scale tests to evaluate the foundation of this design. In one of the setting, Four infrared LED's with wavelength of 880nm was positioned on to a Fujifilm® FinePix S1 Pro DDC still camera. Another setting uses night vision feature of a Sony DCR-TRV9 MiniDV hand-held video camera using the video camera's infrared illuminator as IR light source. Results from both settings suggested that the reflections from the retro-reflective backdrop are white in color and brighter than the IR reflections from the foreground objects. If the projected IR lights evenly covers the entire lens' viewing angle we should be able to get a gray-colored reflection from the backdrop, the brightness depends on the brightness of the camera-mounted light source. It is possible to generate a clean alpha channel from video images like this.

Our testing with both consumer CCD cameras and a Sony® XC HR50 progressive scan camera with a

Heliopan® RG830 infrared filter illuminated with a strobed array of 36 infrared LED's suggested that much better results could be achieved by using normal CCD video cameras to capture visible light imagery while using infrared video cameras to capture the IR reflections. We will consider hardware and software solutions according to further experiments. The goal is to find out if we can obtain video images of the retro-reflective backdrop to be relatively evenly lighted, constant in brightness, and easily differentiate from the foreground objects.

In order to make the system even simpler, we also have considered the possibility of finding special CCD's, if available, that can output "RGBI", four video channels. So far, we have concluded that a more practical approach, as shown in Figure 2., is, to use a prism to split the incoming light beam in two, one targeting a normal CCD camera that outputs RGB channels from visible light, the other targeting an infrared camera with an IR filter which outputs only IR imagery to be used for the alpha channel.

2.4. Differentiating multiple IR light sources

In theory, it is possible that extra IR lights will interfere with existing IR signals, such as those used to synchronize the shutter glasses with the stereo projection. So far, in our preliminary tests, the infrared sources we have tested do not interfere with the stereo shutter glasses in our CAVE™ system. In case interference occurs, possible solutions could be either to use IR lights of different wavelength, or to polarize different light sources using filters or prisms.

2.5. Possible alternative methods

We will continue to consider the possibilities of using alternative methods during our research and development. For example, one of the directions on our desk was to use a backdrop that evenly emits strong infrared light to the direction of the cameras. But technically, this could be even more challenging .We will need to build a backdrop that is either the IR light source or evenly transmits IR from other light sources. For safety reasons, this backdrop obviously cannot generate excessive heat.

3. Stereoscopic video imagery

Extensive research has been done in the area of stereoscopic photographic and stereoscopic computer imagery. Knowledge from these studies is extremely helpful for to customize a high quality stereoscopic video setting in the CAVE™.

3.1. Stereo camera calibration

For a standard size CAVE™, the distance from a user to the stereo camera is estimated to be between four and seven feet. Therefore, we will be working with near stereoscopic viewing, where stereoscopic depth distortions will likely be an important factor. The stereo video camera pair needs to be adjustable and allow precise control of its inter-viewpoint distance, convergence distance, and other system parameters [5]. Many other CAVE™-related factors, including variable ambient light levels and color temperature will need to be taken into consideration. Specific engineering design is very likely necessary.

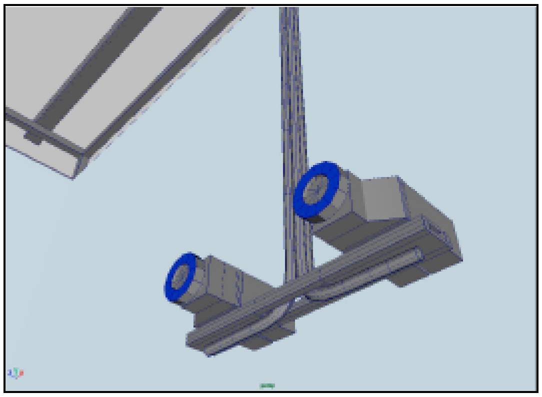

Figure 3. *A possible design of the fully adjustable stereoscopic video camera unit for the CAVE™ with infrared LED rings mounted in front of each lenses.*

3.2. Video resolution and lighting

Current CAVE™ wall projection resolution varies from 1280x942 pixels to 1024x764 pixels depending on the capabilities of the graphics engines and complexity of the applications. Normal NTSC video format is 648x486 pixels in resolution, which might need to be rotated by 90 degrees from "landscape" to "portrait" and stretched to compensate for the CAVE™ projectors' pixel distortion to fill a square screen.

A working CAVE™ is a low light environment with primary lighting from the imagery projections that are subject to change. In order to produce clear and stable video of the user, the cameras need to be very sensitive. We can arrange extra lighting for the user as long as it does not interfere with the image projections. A software filter that can dynamically compensate the stereo video based on graphics card outputs could also be helpful. However, the video avatar's brightness and color change resulted from local projection could be an interesting feature for certain applications.

3.3. Focal length and camera position

The camera-user distance will be mostly between four to seven feet. In order to cover the user from head to toe, wide-angle lenses will be needed. To allow a local user and an avatar to look at each other "in the eyes" without the cameras blocking their faces, the unit needs to be small enough and placed between the user and the avatar with a height slightly above local user's eye level.

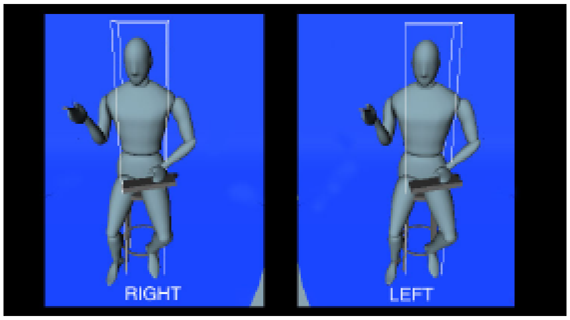

Figure 4, *Simulated left and right views form a stereoscopic video camera set with 28mm focal length. The light color box above the manikin's shoulder indicates a standing height of sit feet and two inches.*

4. Designing for Applications

With this project, we intend to find a way in virtual interaction that allow remote users virtually sit face to face and work as if they were locally in a studio setting. The primary function of the CAVE™ will be changed from content demonstration and evaluation to real time content creation. That means we need to design software and hardware interfaces to accommodate changes in the user's behavior. For example, we will consider ergonomic issues including designing seating and adding a wireless keyboard/mouse and other input devices so as to accommodate possible user demands during multi-hour "in-person" virtual collaborations. There will be plenty of new possibilities to design and develop new interfaces and functionality and challenges to understand the implications of users being immersed in a VE for longer times than in previous scenarios.

5. Possible future works

If proven workable, we believe this experiment will open the doors to many other possible developments including: mobile stereo panorama video camera units that can be dispatched to remote locations which allow user in the CAVE™ to observe from the camera unit's point of view, software and hardware interface design for control and navigation, composing computer synthesized 3D and 2D graphics with stereo video from multiple camera units, recording stereo video footage and editing in real-time, analyzing data from stereo video and use them to create synthesized graphics on top of stereo video. We believe studies in this direction could potentially be valuable to human-computer interaction and 3D imagery research.

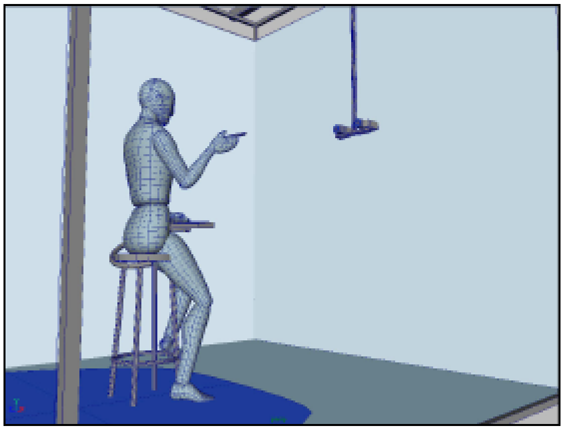

Figure 5. *A user will be working in front of a stereo video camera unit in the CAVE™, the live stereoscopic video footage of one or more remote users could be displayed on the wall to assist remote collaboration.*

6. References

[1] CAVE™ Overview. (n.d.). Retrieved July 18, 2004, http://www.evl.uic.edu/pape/CAVE/oldCAVE/CAVE.overview.html

[2] M. Gross, S. Würmlin, M. Naef, E. Lamboray, C. Spagno, A. Kunz, and E. Koller-Meier, "Blue-C: a Spatially Immersive Display and 3C Video Portal for Telepresence", *ACM Transactions on Graphics (TOG)*, Volume 22 issue 3, Association for Computing Machinery, New York, NY, July 2003.

[3] J. W. Davis and A. F. Bobick, "Sideshow: A silhouette-based interactive dualscreen environment." *Technical Report 457*, MIT Media Lab, August 1998.

[4] T. Hattori, T. Ishigaki, K. Shimamoto, A. Sawaki, T. Ishiguchi, and H. Kobayashi, "An Advanced Autostereoscopic Display for G-7 Pilot Project", *Stereoscopic Displays and Virtual Reality Systems VI, Proceedings of SPIE, Vol. 3639,* SPIE, Bellingham, Washington, January 1999.

[5] D.B. Diner, and D H. Fender, *Human Engineering in Stereoscopic Viewing Devices*, Plenum Press, New York, 1993, pp. 179.