Using Video to Create Avatars in Virtual Reality

The VideoAvatar Library is a collection of functions that works in conjunction with the CAVE Library and can be used to add static, photo-realistic, three-dimensional representations of remote users, as well as other objects or agents, to virtual reality applications. The process involves obtaining views from 360 degrees around the person, then selecting two of these images, one for each eye, to represent the user in 3D space.

To acquire these images, the person stands on a turntable placed in front of a blue screen at a distance of 10 feet from a video camera, which is at eye level. We preview the images and set the levels of the various parameters that are to be used in calculating the chroma key to drop out the background. Once these values are set, we record one revolution of the person turning on the turntable and capture the chroma key settings. Recording at 30 frames per second for the 15 seconds required for one revolution generates a movie file consisting of 450 frames. A configuration file can later be used to specify the maximum number of images that should be used, as well as other pertinent information. This allows individual users to adjust parameters based on the specific hardware and memory capacities of their machines.

From this series of images, a different image is selected to represent the person to each eye. Which two images are to be used is calculated based on the positions of both the local and remote users in the space, and the direction that the remote user is facing. These images are then texture mapped onto two separate polygons, each of which is rotated around the vertical axis toward the eye which is meant to see it.

As we are not actually creating a three-dimensional model, but using two two-dimensional images to represent the avatars, not all depth cues are supported. Among those supported are convergence, binocular disparity, horizontal motion parallax, and occlusion. The proper perspective of the VideoAvatar is maintained in relation to the other objects in the modeled environment.

Despite the limitations of using image-based versus model-based rendering techniques for the VideoAvatars, the results are stunningly realistic. The goal is to supply a range of avatars with varying capabilities. Future versions of the library will incorporate multiple positions within the avatar model and real-time live video in order to fulfill this goal.

Joseph A. Insley, Daniel J. Sandin, Thomas A. DeFanti
Electronic Visualization Laboratory
University of Illinois at Chicago
851 South Morgan Street, Room 1120
Chicago, Illinois 60607 USA
insley@evl.uic.edu

The authors wish to acknowledge **Dave Pape, Javier I. Girado, and Terry Franguiadakis** for their valuable input to this research.