Motion Capture Integration with Tele-immersion Virtual Environment

Mariheida Córdova-Sánchez
Univ. of Puerto Rico, Mayagüez
mariheida.cordova@upr.edu

Ruzena Bajcsy
Univ. of California, Berkeley
bajcsy@eecs.berkeley.edu

Gregorij Kurillo
Univ. of California, Berkeley
gregorij@eecs.berkeley.edu

This paper is in the field of tele-immersion, which pertains to the use of technology to virtually represent three dimensional bodies from remote locations. The primary goal of the project is to integrate real time motion capture data into the existing virtual environment and subsequently minimize the time required to render the bodies when facial movements and other detailed characteristics are not essential. The data from the motion capture system would be used to either control a three dimensional human model avatar or be represented as a set of markers for interaction in virtual space.

1. Introduction

Tele-immersion is a technology that allows local users to physically interact with others that are remotely located. Their bodies are represented in a virtual reality in a way that it seems like the person is really present.

Using avatars to represent users is effective in reducing the rendering time and the size of data sent through the network [1]. Although this is desired, it is only useful when the visualization of detailed characteristics such as: movements of clothes, face, hair, or chest when breathing is not essential.

2. Background

2.1. Tele-immersion Environment

The tele-immersive environment, as shown in Figure 1, consists of 48 cameras arranged in clusters of 4 cameras. Each camera cluster is connected to a computer server which process images to obtain depth information on the scene. The system allows real time as well as prerecorded 3D rendering of the data.

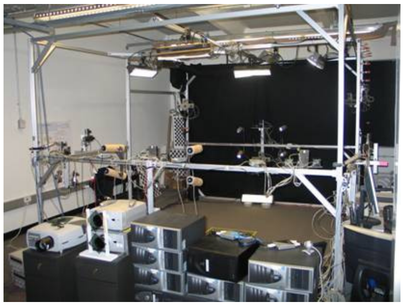

Figure 1: Tele-immersive Environment

2.1.1. Applications of Tele-immersion

The tele-immersion technology can be used for numerous applications such as 3D video conferencing, training, medical therapy, education, art, and entertainment.

Figure 2a: Tai Chi lessons

Figure 2b: Dancers

In Figure 2a we can see a local Tai Chi student learning in real time from a prerecorded video of the teacher. In figure 2b, a local dancer from the University of California, Berkeley, is interacting with two remote dancers located at University of Illinois. Learning from a 3D video has proven to be more effective than a 2D video [2].

2.2. Motion Capture System

The motion capture system [3] is composed of 8 cameras, a software tool to visualize the data, and a full body suit with 38 active infra-red markers included. The cameras are positioned in a circle surrounding an open space to provide a 360 degree view. The cameras provide accurate 3D position of each marker with the frequency of up to 480 Hz. If a marker cannot be seen by at least 2 cameras, it is not rendered.

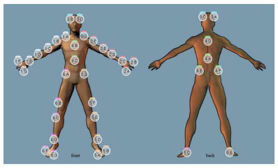
Figure 3a: Motion capture suit [3]

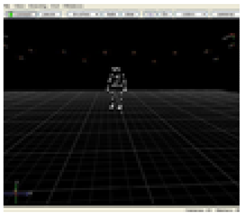
Figure 3b: Motion capture software

Figure 3a shows the position of the set of 38 markers located on the motion capture body suit. Each one of them has a unique identifier. Figure 3b is an illustration that shows how a user is displayed by the motion capture software tool.

3. Methodology

In order to fulfill the goal it was required to development a client program for network connection with the motion capture server to download data in real time. Once the program was able to connect with the motion capture data, each marker location would be represented in the virtual reality environment using OpenGL functions. The captured coordinates were matched with the coordinate system of the multi-camera space. Since the environments have different coordinate systems, in order to get the image in the correct orientation, the y and z coordinates had to be swapped. Later on, the markers were connected forming a human stick figure.

Future use of this technology would permit remotely located users to interact with virtually represented users rendered in real time using a multi-camera system. The motion capture data was connected to the tele-immersion virtual environment [4] so that the interaction between the two systems could be evaluated. Experiments could be performed to compare interactive experience with the real-time captured human participants and real-time controlled avatars inside this environment.

4. Results

Figure 4a shows the representation of the markers in the virtual reality environment. This was done by obtaining the x, y, and z coordinates of the markers through the network at a frequency of 480 Hz. Then, a sphere was created to represent each marker. The spheres were created using OpenGL functions.

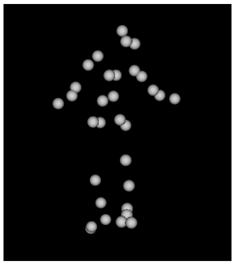
Figure 4a: Set of markers representing a human body

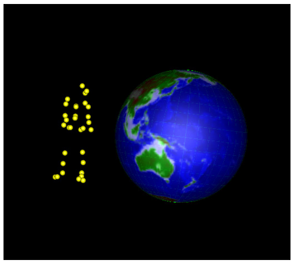
Figure 4b: User interacting with another object

In Figure 4b we can see the same representation of the user, but in this case, interacting with another object. Objects in the virtual reality interface can be rotated to maintain a 360 degree view. In addition, different backgrounds could be added to the virtual space that the users share.

3.1. Connecting the markers

The markers were connected by rendering a cylinder between two consecutive markers, representing body segments. Since the markers are constantly changing position, so are the cylinders between them.

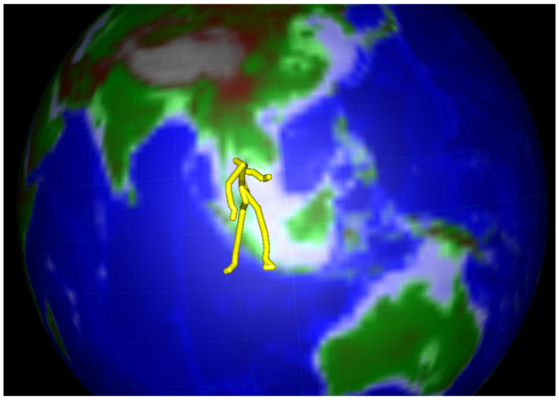
Figure 5: Stick figure

Figure 5 shows the user after connecting the markers with cylinders. The image created by cylinder fitting is smoother and appears to be more human-like than representing only markers.

4.1 Interaction between the systems

A difficulty encountered was that once the data was represented as a set of markers on the VRUI toolkit, the markers followed the same movements they were supposed to, but at a slower pace. This problem occurred because the network could not send the data as fast. It would still send all the data since the buffer sends the next data in the queue, but at a different frequency. At this point, another approach had to be used. To avoid the problem, a thread was created to manage the data. The function called by the thread was reading data from the network independently of the rendering loop. The rendering loop then copied the data to display the stick figure inside the virtual space.

5. Future work

The project could be extended by mapping the markers to a human model avatar with rigid segments. After this is done, experiments should be performed in order to evaluate the interaction of the avatar with other human users reconstructed from the multiple cameras.

Another interesting experiment is to render more than one avatar at once. This interaction should also be analyzed and compared to the former.

6. References

[1] Y. Yang, X. Wang, and J. X. Chen. Rendering avatars in virtual reality: integrating a 3d model with 2d images. Computing in Science and Engineering, 4(1):86–91, 2002.
[2] G. Kurillo, R. Bajcsy, K. Nahrsted, O. Kreylos. Immersive 3D Environment for Remote Collaboration and Training of Physical Activities. Virtual Reality Conference, March 269-270, 2008.
[3] PhaseSpace Motion Capture system.
Web: http://www.phasespace.com
[4] O. Kreylos, Vrui VR Toolkit, Web: http://graphics.cs.ucdavis.edu/~okreylos/ResDev/Vrui/index.html