Avatar Communication: Virtual Instructor in the Demonstration Exhibit

Tetsuro Ogi[1, 2, 3], *Toshio Yamada*[1], *Takuro Kayahara*[1, 2], *Yuji Kurita*[1]

[1] Telecommunications Advancement Organization of Japan
[2] IML, The University of Tokyo
[3] Mitsubishi Research Institute, Inc.
2-11-16 Yayoi, Bunkyo-ku, Tokyo 113-8656, Japan
tetsu@iml.u-tokyo.ac.jp

Abstract

Video avatar is a technique that represents a realistic human image in the virtual space by using a live video, and it has been used to realize a high presence communication in the shared virtual world. In this technique, various kinds of video avatar models, such as the 2-dimensional plane model, the 2.5-dimensional depth model, and the 3-dimensional voxel model, have been proposed and used. Particularly, in this study, a video avatar studio and a video avatar server were developed in order to generate various kinds of video avatar data and use them in the application systems. These technologies were applied to the psychological demonstration exhibits in the shared immersive virtual world and they were used effectively.

1 Introduction

In the recent virtual reality systems, high presence virtual worlds can be displayed, and therefore representation of a realistic human image is also required. In particular, when the immersive virtual reality environment such as the CAVE is used, a real size image of the whole body expression is desired in order to realize the interaction as a first person experience. In addition, since the infrastructure of the broadband network has been constructed, a high presence human image can also be used as a natural communication tool in the share virtual world.
In this research, in order to meet these demands, video avatar technology has been studied. Video avatar is a technique that generates a high presence human image in the virtual world using a live video. In this method, various kinds of video avatar models, such as 2-dimensional, 2.5-dimensional and 3-dimensional video avatars, have been proposed and they are used according to the purposes of the application systems.
This paper describes the several kinds of video avatar technologies developed in this study, and the video avatar studio and video avatar server that generate various kinds of video avatar data and transmit them to the virtual reality application systems. Moreover, the video avatar was applied to the psychological demonstration exhibits, in which the psychologist was represented as a video avatar and he instructed the psychological experiment in the shared virtual world.

2 Video Avatar Technology

In general, video avatar is generated in the following process. First, the person's figure is captured by the video camera, and then only the person's image is segmented from the background as well

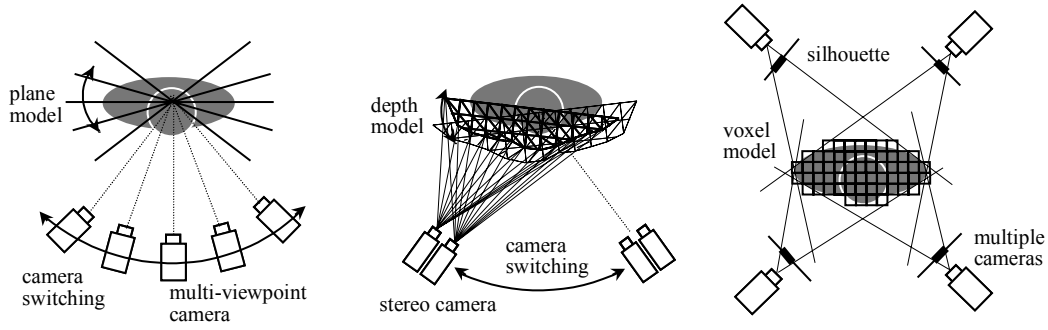
as the geometry model of the person is created. By texture mapping the segmented person's image onto the geometry model, the video avatar is created. In this process, various kinds of video avatars can be generated due to the created geometry models and the segmentation methods of the person's figure.

The simplest video avatar method uses a 2-dimensional plane model. In this method, since the segmented person's image is placed as a 2-dimensional board in the virtual space, the video avatar itself does not have depth information. However, when the multiple cameras are used to film the person's image from various directions, 3-dimensional expression such as a motion parallax can be performed, by changing the selected video images according to the movement of the user's viewpoint. Figure 1 (a) shows the generation method of the 2-dimensional video avatar using the multi-viewpoint camera system.

On the other hand, when a stereo video camera is used, depth information for each pixel of the filmed image can be acquired using stereo matching algorithm. By arranging each pixel of the captured image in the virtual space according to the depth distance, a geometric model of the filmed image is created. Then, a stereo video avatar that has depth information can be generated, by texture mapping the segmented person's image onto the geometric model in real-time. However, since this model has depth information only for the front surface that faces the stereo camera, it is called a 2.5-dimensional video avatar (Ogi et al., 2001). Moreover, this model can be also used to perform 3-dimensional expression by changing two or more stereo cameras according to the movement of the user's viewpoint as shown in Figure 1 (b).

A 3-dimensional video avatar that has a complete surface model for of all directions can be generated, by using multiple cameras placed surrounding a person (Moezzi et al., 1996). In this method, the silhouette of the person's figure is calculated for each image filmed by the multiple cameras. By judging the position of each voxel in the three-dimensional space whether it is placed inside or outside the silhouette calculated by each camera, a voxel model of the person's figure is generated. Then, the 3-dimensional video avatar can be generated by texture mapping the segmented person's image onto the surface polygons created from the voxel data. Figure 1 (c) illustrates the generation method of the 3-dimensional video avatar.

Since these video avatar models have individual features about the calculation time and the image quality, the users should select the appropriate method according to the purposes of the application systems.



(a) 2-dimensional video avatar (b) 2.5-dimensional video avatar (c) 3-dimensional video avatar

Figure 1: Generation method of various kinds of video avatar

3 Video Avatar Studio and Video Avatar Server

The video avatar technology can be used not only to represent a realistic human in the virtual world but also to realize a high presence communication, by transmitting the video avatar data

mutually. In addition, the video avatar can be reproduced using the recorded data as well as it is transmitted in real-time. In this study, a video avatar studio and a video avatar server were developed in order to generate various kinds of video avatar data and use them in various methods. The video avatar studio is a cylindrical room with a diameter of 4000 mm and a height of 2570 mm. Figure 2 and Figure 3 show the system construction of the video avatar studio. In this system, eighteen CCD cameras (SONY DFW-X700) are mounted on the blue back wall at intervals of 20 degrees. The heights of the camera positions can be changed among 600 mm, 1200 mm, and 1575 mm from the floor. These cameras are attached on the outside of the wall and only the camera heads are stuck out into the inner side of the room through the holes. Since the background filmed by every camera is blue, the person's image can easily be segmented from the background using the chroma-key method.

Figure 2: Cylindrical video avatar studio

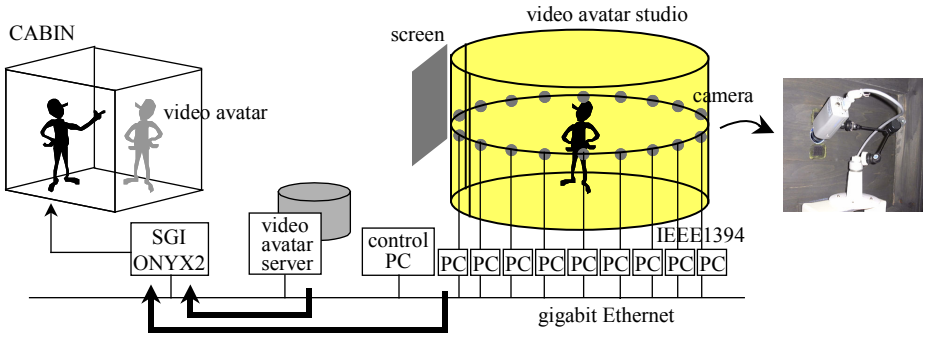


Figure 3: System construction of video avatar studio and video avatar server

The filmed images are transmitted to the PCs through IEEE 1394 connections, and they are used to generate various kinds of video avatar data. For example, the 2-dimensional video avatar using a plane model is created simply switching the images captured by eighteen cameras. When the 2.5-dimensional video avatar is used, two cameras placed in front of the user are used as a stereo camera to calculate the depth value. And the 3-dimensional video avatar is generated using all cameras to create a voxel model.

The PCs are connected to the graphics workstation through the gigabit Ethernet and the video avatar data are transmitted to the virtual reality system. These data can be stored in the database as well as they are superimposed on the virtual world in real-time. The stored data is transmitted to the virtual reality applications by the video avatar streaming server and they can be used repeatedly. The video avatar server communicates with the virtual reality applications and

transmits the avatar data retrieved from the database according to the requests from the applications. In this case, the streaming server transmits the video avatar data at the same frame rate as the recording frame rate. Therefore, the application system can receive the video avatar data and integrate it in the virtual world without considering whether it is a real-time data or a recorded data.

4 Psychological Demonstration Exhibit

In this study, the video avatar technology was applied to the psychological demonstration in the immersive projection display CABIN (Hirose et al., 1999). CABIN is a CAVE-like multi-screen display that has five screens at the front on the left, right, ceiling and floor. In this system, the video avatar of the psychologist was superimposed on the demonstration of the virtual psychology laboratory, and the audiences experienced the psychological demonstrations such as the scene recognition and the pointing gesture recognition. In these demonstrations, the figure of the psychologist appeared in the virtual laboratory as a video avatar and he explained the psychological phenomenon and how the audiences can experience the demonstration in the shared virtual world.

In the demonstration of the scene recognition experiment, the audiences and the remote psychologist shared the virtual room through the network, and several objects were put on the table placed in it. In this experiment, the arrangement of the objects was changed and the table or the room was rotated. And then, the subjects were asked whether they could recognize the change of the positions of the objects.

Although the psychologist and the subjects are in the distant places, they were able to communicate with each other using the video avatar technology. In this demonstration, the psychologist's figure was represented as a 2.5-dimensional video avatar that has the front surface model, and he explained the contents of the experiment using his gesture in the virtual world as shown in Figure 4. The figure of the psychologist was also recorded as video avatar data and it was transmitted to the demonstration system using the video avatar server, so that it could be used in the demonstration exhibit repeatedly even when the psychologist was absent. Although the transmission of the psychologist’s image was stopped during the experiment so that the subject carried out the experimental task alone, the psychologist was able to observe the subject’s behaviour by his side without being noticed, because the subject’s image was transmitted to the psychologist site.

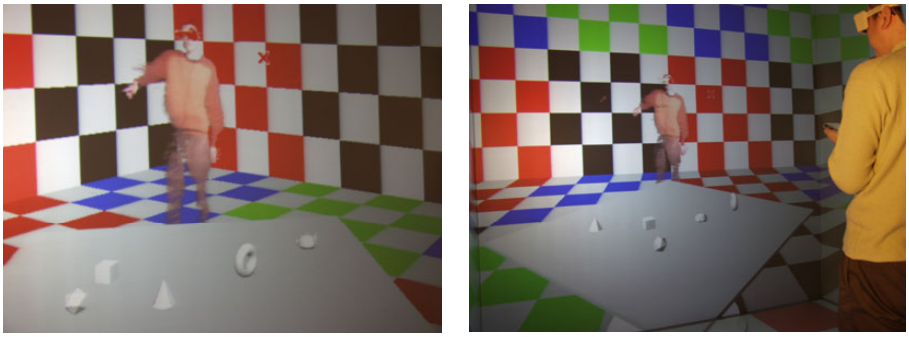
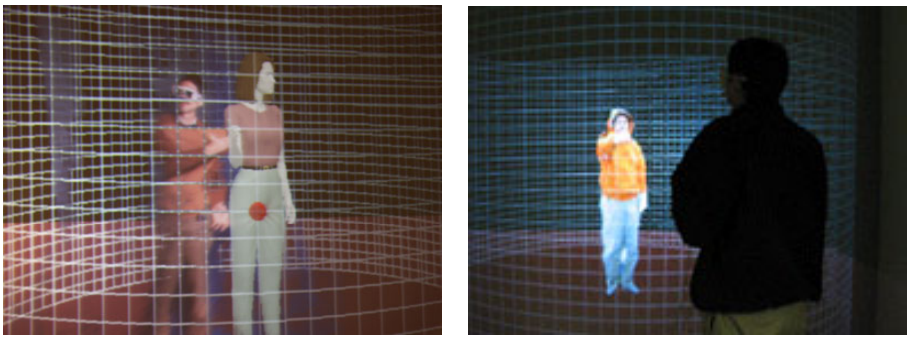


Figure 4: Demonstration of the scene recognition experiment

The demonstration of the pointing gesture recognition is a psychological experiment that shows how accurately the user can recognize the human’s pointing gesture and which element of the

person's body, such as the direction of the arm or the direction of the face, affects the user's recognition. Also in this demonstration, the psychologist explained the contents of the experiment to the remote subjects using the 2.5-dimensional video avatar. In addition, the figure of the person who was actually pointing at the objects in the real world was represented in the virtual world using the 2.5-dimensional or 3-dimensional video avatar. These video avatar data were recorded beforehand and were used in the experiment as well as they were transmitted from the remote site in real-time. Figure 5 (a) shows the video avatar of the psychologist who is explaining the contents of the experiment, and Figure 5 (b) shows the video avatar of the pointer who is actually pointing at the object in the real world.
In these demonstrations, several users experienced the psychological experiments in the shared virtual world, and the video avatar technology was effectively used for the instruction and the explanation of the demonstration exhibits.



(a) Video avatar of psychologist (b) Video avatar of pointer

Figure 5: Demonstration of the pointing gesture recognition experiment

5 Conclusions

In this study, various kinds of video avatar methods, such as using the 2-dimensional plane model, the 2.5-dimensional depth model and the 3-dimensional voxel model, were developed. These models can be easily generated and used using the video avatar generation studio and the video avatar streaming server technologies. In this study, these technologies were applied to the psychological demonstration exhibits in the shared virtual world and they were used effectively. Future work will include applying the video avatar communication technology to several practical exhibits and evaluating the effectiveness of this technology.

References

Ogi, T., Yamada, T., Kano, M., Hirose, M. (2001). Immersive Telecommunication Using Stereo Video Avatar, *Proceedings of Virtual Reality 2001 Conference*, 45-51.

Moezzi, S., Katkere, A., Kuramura, D.Y., Jain, R. (1996). Immersive Video, *Proceedings of VRAIS'96*, 17-24.

Hirose, M., Ogi, T., Ishiwata, S., Yamada, T. (1999). Development and Evaluation of Immersive Multiscreen Display "CABIN". *Systems and Computers in Japan*, 30 (1), 13-22.