Proceedings of the ASME 2009 International Design Engineering Technical Conferences &
Computers and Information in Engineering Conference
IDETC/CIE 2009
August 30 - September 2, 2009, San Diego, California, USA

DETC2009-86762

HIGH PRESENCE COLLABORATION USING PLUG-IN VIDEO AVATAR

Tetsuro Ogi
Keio University
4-1-1 Hiyoshi, Kohoku-ku,
Yokohama, Kanagawa, Japan

Yoshisuke Tateyama
Keio University
4-1-1 Hiyoshi, Kohoku-ku,
Yokohama, Kanagawa, Japan

Satoshi Oonuki
University of Tsukuba
1-1-1 Tennodai, Tsukuba,
Ibaraki, Japan

ABSTRACT

In order to realize effective collaboration between remote places, it is necessary to achieve high presence communication in the shared space. In this study, plug-in video avatar was developed. Video avatar is a technology that supports the communication between remote users by capturing the user's video image and integrating it in the shared virtual world. In this method, stand-alone application can be expanded to the shared space application easily by using the plug-in function. The video avatar technology was applied to the remote collaboration by synthesizing it to various applications such as visualization, collaborative design, and remote presentation. The efficiency of the communication in the remote presentation and the system performance among three sites were investigated to evaluate the effectiveness of this system.

INTRODUCTION

In recent years, since large scale systems have been developed in various fields, a large number of people often collaborate to develop them. In the cases when collaborative work is conducted, the collaboration through the network between remote places is necessary because collaborative workers are not usually in the same place. As for the technology to support the collaborative work between remote places, video conferencing system [1][2] or application sharing tools [3] have been used. However, in the conventional video conferencing system, the users cannot feel the sensation that they are in the same place, because they communicate with each other through the two dimensional screen. In the application sharing tool, since remote users only share the application window, it is difficult for them to inform where they are paying attention to the other user.

In order to realize high presence and effective collaboration between remote places, it is desired to achieve emotional communication in which the user can see the other user's figure and recognize where he or she is paying attention as well as transmit the necessary information between each other. In this case, it is preferred that the user's figure is represented using not the computer graphics model [4] but the live image [5][6]. This study aims at constructing tele-immersive communication using a video avatar technology to achieve such a demand. Video avatar is a technology to enhance the communication in the shared virtual space, by transmitting the user's live video image and integrating it in the shared virtual world [7]. Particularly, the tele-immersion application should be constructed easily by extending the stand-alone application.

This paper describes the construction method of the video avatar and the plug-in video avatar function that was developed in this study to use the video avatar technology easily. In addition, the applications to the remote collaboration and the performance experiment of the video avatar are discussed.

NETWORKED IMMERSIVE ENVIRONMENT

It is desired to achieve high presence communication in the shared space in order to realize effective communication between remote places. In this study, tele-immersion environment was constructed as a platform of the shared space by connecting several immersive projection displays through the high-speed network, and it was used for the collaboration between remote places. As immersive projection displays, CAVE-type multi-screen displays were used [8].

Figure 1 shows the K-Cave system that was developed at Keio University. This display system has three screens at the front, right and on the floor, and generates three-dimensional virtual space by projecting the synchronized passive stereo images through the polarizing filters. The user in this display can experience an interactive virtual world by tracking his view position with the electro-magnetic sensor.

In this study, several CAVE-type immersive projection displays equipped at several sites were connected through the network and video avatar technology was applied to perform an immersive communication.

Figure 1. K-Cave system

VIDEO AVATAR TECHNOLOGY

Video avatar is a technology that supports the communication between remote places by transmitting the user's video image to the remote site. This technology can be used for the communication in the three-dimensional shared space, by integrating the user's video image at the three-dimensional position in the virtual world.

Figure 2 shows the method of generating the video avatar. In this method, the user's figure is captured by video camera and it is transmitted to the remote place with the position data. Since the view position of the user is tracked by the sensor in the immersive projection display, it can be used as the positional information of the video avatar. At the remote site, the transmitted user's image is synthesized in the virtual world by texture mapping the video image onto the rectangular plate. In this case, only the user's figure is segmented from the background in every frame, by subtracting the background image that was captured beforehand from the transmitted user’s image. Then, the video avatar image can be placed at the position where the remote user exists in the virtual world.

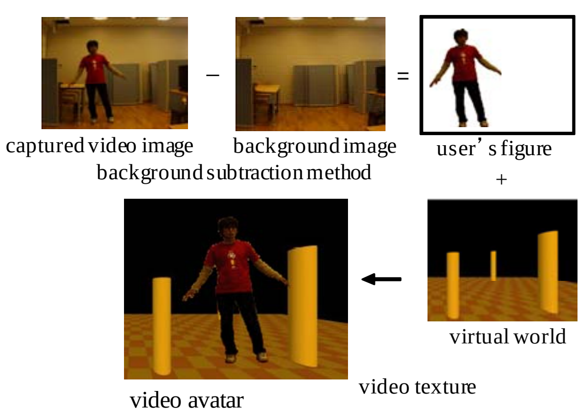

Figure 2. Method of generating video avatar

By performing these processes between remote places mutually in real time, the video avatar can be used for the communication in the shared virtual world. In this method, the user can communicate with the remote user in the first-person point of view, and they can feel the sensation of sharing the three-dimensional space in the immersive virtual environment. In addition, it is possible to express where the remote user is paying his attention by using the gesture and the facial expression, since the video avatar is represented using the video image of his whole body. Though the image of the video avatar is two-dimensional, the spatial relationship between remote users in the shared virtual world can be represented by placing it at the three-dimensional correct position. Thus, it would be possible that the remote users have high presence and emotional communication using the video avatar technology.

OPEN CABIN LIBRARY

In this study, plug-in function of the video avatar was developed using OpenCABIN library [9], in order to extend the stand-alone application to the shared space application easily. The OpenCABIN library is platform software for the virtual reality application in the immersive projection display such as the CAVE or CABIN [10]. Figure 3 shows the software construction of the application program using the OpenCABIN library. This library has features of master-renderer programming paradigm and plug-in mechanism, so that it can effectively be used to develop a network application.

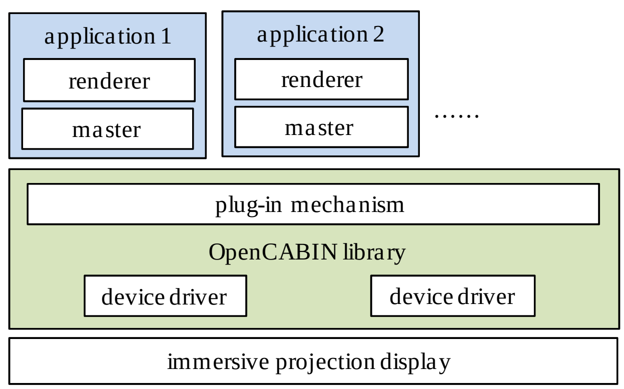

Figure 3. Software construction of OpenCABIN

The master-renderer programming paradigm is a style that defines an application program consisting of master part that calculates the changes of the virtual world and renderer part that renders the image of the virtual world. In the multi-screen display such as the CAVE, though the plural processes of the renderer should be implemented according to the number of the screens, one process of the master is executed to calculate the state of the virtual world. This software configuration is effective to construct the network application that needs to access to the database or needs to transmit data to other processes, because only one master process communicate with the remote processes.

On the other hand, the plug-in mechanism is a method to construct an application program in the form of dynamically loaded library. When API of the usual library is used, it is necessary to rewrite the source code of the program to add a certain function to improve it. For example, in order to construct a shared space application, it must be developed specially besides the stand-alone application. However, the plug-in mechanism can add a necessary function to the application even when other program is being executed. In this method, it is not necessary to develop an application program of the contents that contains all models of the objects that exist in the virtual world. The virtual world can be constructed by developing plural programs that render the model of each object and executing the selected programs according to the necessary objects. By using the video avatar application using the plug-in mechanism, a shared space application could easily be developed by executing it together with other applications.

PLUG-IN VIDEO AVATAR

In this study, basic program of the plug-in video avatar that transmits a video avatar image between two sites was developed, and it was used to construct shared space applications. The system and functional configuration of the plug-in video avatar are shown in Figure 4. At the sending site, the user's image is captured by IEEE1394 video camera (Sony, DFWX-710), and it is sent to the other site frame by frame. At the receiving site, the user’s image is segmented from the background using the background subtraction method, and the video avatar is generated by rendering it at the three-dimensional position in the virtual space. Since the program at the receiving site is constructed using the OpenCABIN library, the master program runs on the communication PC and some renderer programs run on the rendering PCs corresponding to the multiple screens.

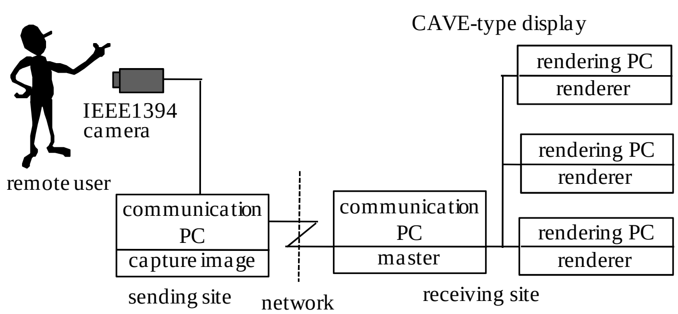

Figure 4. Functional configuration of plug-in video avatar

By using the plug-in video avatar program at both sites between remote places to transmit the video avatar image mutually, remote users are able to communicate with each other in the shared space. Since this program is constructed using the plug-in mechanism, it can be used to generate the collaboration environment in which remote users can perform the collaborative work while sharing the visualization data, by executing it together with the other applications. Moreover, though this program was developed basically for the usage between two sites, it can also be used to generate the collaboration environment among multiple sites, by executing plural processes of the plug-in video avatar between every two sites.

APPLICATIONS

In this study, the plug-in video avatar was integrated with several applications, and high presence collaborations in which remote users communicate with each other while looking at the other user’s figure and sharing data in the networked virtual world were realized.

Figure 5 shows an example in which remote users are conducting collaborative work to analyze seismic data by integrating the video avatar with the visualization program of the earthquake data [11]. In this example, several kinds of data such as the map, hypocenter, basement depth and plate structure data were visualized together with the video avatar image, and the seismic phenomenon was analyzed in the collaboration environment by representing the relationship among several data. Then, the video avatar could be used effectively to perform the discussion while sharing the visualization data between remote places.

Figure 5. Application to collaborative visualization

Next, Figure 6 shows another example of integrating the video avatar with the visualization program of the design model. In this example, remote users conducted the collaborative design in the three-dimensional shared virtual space, by synthesizing the video avatar image with the three-dimensional model that was designed using the CAD system. The remote users were able to conduct a discussion while indicating the design-model and paying attention to it, by using the video avatar in the shared virtual world.

Moreover, Figure 7 is an example of the video avatar application to the remote presentation [12]. In this example, presentation was achieved between remote places, by synthesizing the video avatar to the visualization program of the PowerPoint slides in the virtual world. In this case, though the video avatar of the speaker was represented as one person, the video avatars of the listeners were represented as plural persons. However, the question and answer between the speaker and the listeners were naturally performed using the asymmetrical condition of the video avatar.

Figure 6. Application to collaborative design

Figure 7. Application to remote presentation

EFFICIENCY OF COMMUNICATION

In this study, the efficiency of the remote communication was compared between when the conventional video conferencing tool was used and when the video avatar technology was used, in order to examine the effectiveness of the video avatar technology.

In this experiment, remote presentation in which the research contents were explained using the PowerPoint slides was conducted from another room in the university campus. When the video conferencing method was used, the image of the presenter and the PowerPoint slide was captured by the video camera, and it was projected onto the front screen of the CAVE system as a two-dimensional image. On the other hand, when the video avatar technology was used, the images of the video avatar and the PowerPoint slide were represented in the three-dimensional virtual space displayed in the CAVE system. Therefore, the presenter was able to move the standing position against the PowerPoint slide in the three-dimensional space. Figure 8 shows the appearances of this experiment.

(a) Video conferencing method

(b) Video avatar method

Figure 8. Experiment on efficiency of communication

Table1. Results of questionnaire

question	answer
Which method is easy to hear the presenter’s explanation?	4.7
Which method generates the feeling of being looked at by the presenter?	3.3
Which method is easy to talk to the presenter?	5.0
Which method is easy to understand the presenter’s gesture?	3.0
Which method is easy to concentrate on the presentation?	3.3

(1: video conference <- -> 5: video avatar)

In the experiment, the subjects of three university students were asked to hear the presentation in both methods and answer the questionnaires shown in Table 1. The subjects answered "1" when they felt the video conference was more effective, and they answered "5" when they felt the video avatar was better. In this table, mean value of the answers for each question was shown. From the result, the answer for the video avatar method was better than the video conferencing method, except for the question about the understandability of the presenter's gesture.

In particular, since the effectiveness of the video avatar was clearly shown for the questions about the easiness of talking to the presenter and hearing the presenter's explanation, we can understand that the video avatar technology generates the sensation of sharing the space between remote users.

SYSTEM PERFORMANCE

Next, we conducted a communication experiment among three sites; University of Tsukuba, the University of Tokyo, and Kyoto University, in order to evaluate the performance of the plug-in video avatar function. University of Tsukuba is located about 60km and 530km away from the University of Tokyo and Kyoto University respectively. The communication performance depends on the display system at each site and the network environment among them. In these sites, CAVE-type immersive projection displays are equipped and they are connected to the JGN2 (Japan Gigabit Network 2) network. JGN2 is a testbed network for research and development operated by NICT (National Institute of Information and Communications Technology) and it has 20Gbps backbone. Then, the tele-immersion environment was constructed through the high speed network.

In the experiment, the plug-in video avatar was integrated with the seismic data visualization program, and the execution performance was measured. Figure 9 shows the appearance of this experiment. Since the user at each site can walk through the virtual space freely, the positional relationship among three users in the shared virtual world can be changed arbitrarily. In this application, since the visualization programs for the hypocenter data, basement depth data and plate structure data were also developed using the OpenCABIN library, each data can be added or removed freely by using the plug-in function. In addition, the visualization program for each data can easily be expanded to the shared space application, by integrating the video avatar program with it. In this case, tele-immersion environment among three sites was realized by executing two processes of the video avatar programs that construct the communication with other two sites at each site as shown in Figure 10.

Figure 9 Experiment on video avatar communication

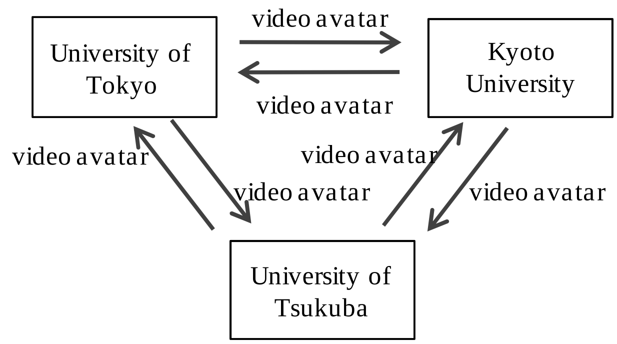

Figure 10. Video avatar transmission among three sites

Table 2 shows the result of the performance when the communication experiment was conducted between Tsukuba and Tokyo sites. In this experiment, the frame rate of rendering the video avatar was measured, when the user's image was captured at 15Hz and the resolution of the captured image was changed among XGA (1024x768), SVGA (800x600), and VGA (640x480). From the result, it was possible to transmit and render the video avatar image at the same frame rate as capturing the image when VGA resolution image was used, though it was not possible when XGA and SVGA images were used.

Table2. Performance of communication between two sites

resolution	XGA	SVGA	VGA
frame rate (Hz)	6.01	9.96	14.9

Next, Table 3 shows the result of the performance measured at Tsukuba site, when the communication experiment was conducted among three sites using the video avatar of VGA resolution image. Though the video avatar image was sent from Tsukuba site to Tokyo and Kyoto sites at more than 15 Hz, it was received from both sites at 6 or 7 Hz. It seems that this is because the concentrated load was applied to the communication PC that was used to send and receive the video avatar data as well as to execute the master process of the application. As for the performance of the renderer for each screen of the immersive projection display, though the frame rate was 27 Hz when only the video avatar was rendered, it was decreased to 13 Hz when the hypocenter data was visualized with the video avatar. However, when the light load data such as the basement depth or the plate structure data was visualized together with the video avatar, the performance did not reduce compared with the case when only the video avatar was rendered. This result shows that when the video avatar communication is used among multiple sites, the rendering performance of the virtual world would not decrease so much, even if the update of the video avatar image is reduced.

In this system configuration, since the several processes were executed on the communication PC, the calculation load was concentrated on it in each site. Therefore, it was necessary to consider the tuning of performance for the whole system, because when the performance at one site was improved, it would have an influence on the other sites. In the communication among multiple sites, the design of a flexible system configuration in which the low performance at one site does not have an influence on the whole system would be necessary.

Table3. Performance of communication among three sites

	avatar	avatar basement	avatar basement plate	avatar basement plate hypocenter
capture image	15.001	15.001	15.002	15.000
send to Tokyo	22.357	22.319	22.390	22.067
send to Kyoto	22.467	22.435	22.353	22.488
receive from Tokyo	6.463	6.114	6.118	6.116
receive from Kyoto	7.404	7.354	7.682	7.548
master	13.562	13.574	13.554	13.573
renderer	27.122	27.172	27.102	13.568

(Hz)

CONCLUSIONS

In this study, plug-in video avatar function was developed as a method of realizing high presence collaboration among remote places. By using this method, stand-alone application could be expanded to the shared space application easily. Though the efficiency of transmitting the information of the video avatar technology was suggested in the remote presentation, we are planning to conduct a further experiment by increasing the number of subjects. In addition, the video avatar was used for the communication among three sites and the system performance was measured to evaluate the effectiveness and extract problems of this system. Future work will include applying the video avatar technology to various applications to realize the effective collaboration as well as developing some tools of accurate pointing or controlling operation in the shared virtual world.

ACKNOWLEDGMENTS

This work was supported by Keio University Global COE (Center of education and research of symbiotic, safe and secure system design) Program and NICT (National Institute of Information and Communications Technology). And we thank Prof. Yasuo Ebara of Osaka University, Dr. Hideo Miyachi and Dr. Nobuyuki Kukimoto of KGT Inc., for their supports.

REFERENCES

[1] Jian Sun, Holger Regenbrecht: J.W. Brackett: Implementing three-party desktop videoconferencing, Proc. of 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces, pp.95-102, 2007.

[2] Slovak, P., Hladka, E., Troubi, P.: Videoconferencing design for remote group, CollaborateCom 2007, pp.437-441, 2007.

[3] Omer Boyaci, Henning Schulzrinne: BASS Application Sharing System, Proc. of 2008 Tenth IEEE International Symposium on Multimedia, pp.432-439, 2008.

[4] Leigh J., Johnson A.E., Vasilakis C.A., DeFanti T.A., Multi-Perspective Collaborative Design in Persistent Networked Virtual Environments, Proceedings of the IEEE VRAIS, pp.253-260, 1996

[5] Insley J., Sandin D.J., DeFanti T.A., Using Video to Create Avatars in Virtual Reality, Visual Proceedings of 1997 SIGGRAPH, pp.128, 1997.

[6] Mojtaba Hosseini, Nicolas D. Georganas: Design of Multi-sender 3D Videoconferencing Application over an End System Multicast Protcol, Multimedia ’03, pp.480-489, 2003.

[7] Tetsuro Ogi, Toshio Yamada, Yuji Kurita, Yoichi Hattori, Michitaka Hirose: Usage of Video Avatar Technology for Immersive Communication, Proceedings of First International Workshop on Language Understanding and Agents for Real World Interaction, pp.24-31, 2003.

[8] C. Cruz-Neira, D.J.Sandin, T.A. DeFanti, "Surround-Screen Projection-Based Virtual Reality: The Design and Implementation of the CAVE," Computer Graphics (Proc. of SIGGRAPH '93), pp. 135-142, 1993.

[9] Yoshisuke Tateyama, Satoshi Oonuki, So Sato, Tetsuro Ogi: K-Cave demonstration: Seismic information visualization system using the OpenCABIN library, ICAT 2008 (Proceedings of 18th International Conference on Artificial Reality and Telexistence), pp.363-364, 2008.

[10] Michitaka Hirose, Tetsuro Ogi, Shohei Ishiwata, Toshio Yamada: Development and Evaluation of Immersive Multiscreen Display "CABIN", Systems and Computers in Japan, Scripta Technica, Vol.30, No.1, pp.13-22, 1999

[11] Satoshi Oonuki, Yoshisuke Tateyama, Tetsuro Ogi: Seismic Data Visualization in Tele-immersion Environment, ASIAGRAPH 2008 in Tokyo Proceedings, Vol.2, No.2, pp.186-189, 2008.

[12] Tetsuro Ogi, Hiroaki Sakon: Distance Learning in Tele-immersion Environment, 21th International Conference on Advanced Information Networking and Applications Workshops/Symposia Volume II (AINA2007), pp.947-952, 2007.

