Networked Telepresence System Using Web Browsers and Omni-directional Video Streams

Tomoya Ishikawa[a], Kazumasa Yamazawa[a], Tomokazu Sato[a], Sei Ikeda[a], Yutaka Nakamura[a], Kazutoshi Fujikawa[a], Hideki Sunahara[a], Naokazu Yokoya[a]

[a] Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara, 630-0192 Japan

ABSTRACT

In this paper, we describe a new Telepresence system which enables a user to look around a virtualized real world easily in network environments. The proposed system includes omni-directional video viewers on web browsers and allows the user to look around the omni-directional video contents on the web browsers. The omni-directional video viewer is implemented as an Active-X program so that the user can install the viewer automatically only by opening the web site which contains the omni-directional video contents. The system allows many users at different sites to look around the scene just like an interactive TV using a multi-cast protocol without increasing the network traffic. We describe the implemented system and the experiments using live and stored video streams. In the experiment with stored video streams, the system users an omni-directional multi-camera system for video capturing. We can look around high resolution and high quality video contents. We have also implemented a view-dependent presentation with a head-mounted display (HMD) and a gyro sensor for realizing more rich presence. In the experiment with live video streams, a car-mounted omni-directional camera acquires omni-directional video streams surrounding the car, running in an outdoor environment. The acquired video streams are transferred to the remote site through the wireless and wired network using multi-cast protocol. We can see the live video contents freely in arbitrary direction.

Keywords: Network, Telepresence, Web Browser, Omni-directional Video Stream, Multi-cast

1. INTRODUCTION

Recently, telepresence which allows us to experience a remote site through a virtualized real world has been investigated[1]. The technology can be widely applied to a number of fields such as entertainment, medical equipment and education. In previous work, we have proposed telepresence systems using omni-directional images which enable a user to see the virtualized dynamic scene freely in arbitrary direction[2-3]. The users can look around a wide area in a virtualized remote site without significant delay from the rotation of user's head to the presentation of images. The conventional systems were realized without network and with a stand-alone system that uses an immersive display. Therefore, it was difficult for many users to use the system simultaneously at different places.

In this paper, we describe a new telepresence system which enables the user to look around a virtualized real world easily in network environments. The new system includes omni-directional video viewers on web browsers and allows the user to look around the omni-directional video contents on the web browsers. The omni-directional video viewer is implemented as an Active-X program. The user can install the viewer automatically only by opening the web site which contains the omni-directional video contents.

Moreover, we describe the implemented system and the experiments using live and stored video streams. In the experiment with stored video streams, the system uses an omni-directional multi-camera system Ladybug for video capturing. We can look around high resolution and high quality video contents.

Further author information:
T.I, K.Y, T.S, S.I, N.Y: E-mail: {tomoya-i, yamazawa, tomoka-s, sei-i, yokoya}@is.naist.jp
Y.N, K.F, H.S: E-mail: {yutaka-n, Fujikawa, suna}@itc.naist.jp

We have also implemented a view-dependent presentation with an HMD and a gyro sensor for realizing more rich presence. In the experiment with live video streams, a car-mounted omni-directional camera acquires omni-directional video streams surrounding the car, running in an outdoor environment. The acquired video streams are transferred to the remote site through the wireless and wired network. Then, we can see the live video contents freely in arbitrary direction. The system allows many users at different site to look around the scene just like an interactive TV using a multi-cast protocol without increasing the network traffic.

Section 2 describes the proposed system. Section 3 describes experiments of the system with live video and stored video. Finally, Section 4 summarizes the present work.

2. NETWORKED TELEPRESENCE SYSTEM USING WEB BROWSERS AND OMNI-DIRECTIONAL VIDEO STREAMS

The schema of proposed system is illustrated in Fig. 1. Omni-directional videos are stored in remote server and are acquired by the viewer which is started by a web browser. The user looks around the omni-directional video contents on the web browser.

2.1. Omni-directional video viewer using web browser

A web browser is one of the most popular network applications. Especially, Internet Explorer installed on Windows machines can execute various application programs by a JAVA applet or Active-X for providing users with interactive contents on a web page. Moreover, the JAVA applet and the Active-X programs can be easily distributed by an automatic install function. Thus we implement an omni-directional video viewer for telepresence on a web browser in this study.

Fig. 2 shows the structure of the omni-directional video viewer. The omni-directional video viewer which shows the user omni-directional video contents needs the functions of GPU (graphic processor unit) and is implemented by Active-X. When the user access the web page which provides an omni-directional video contest, the omni-directional video viewer implemented by ActiveX is installed automatically. The omni-directional video viewer is started by the web browser and is provided the parameters such as an omni-directional camera type, parameters of camera, file-path of content, and so on. The parameters are embedded in a HTML file by a content provider and so the user does not need the care of them. Firstly, the omni-directional video viewer accesses the omni-directional video stream or file using DirectShow functions. DirectShow functions can handle the various video formats such as AVI, MPEG, WMV, and so on, and acquire the image of video. Secondly, the acquired image is transferred to the OpenGL function. We use OpenGL functions with GPU functions for converting the omni-directional image to the common perspective image. OpenGL function use a hardware texture mapping of GPU function to a sphere object for converting image in real-time by the method[2]. Finally, the omni-directional video viewer draws the converted image to the window of the web browser. These functions are synchronized with the frame rate of inputted video.

The user can look around the omni-directional video by using a mouse-drag operation. The omni-directional video viewer is installed only by opening the web page, and the omni-directional video can be seen easily. In the other implementation, the user can look around the omni-directional video through a HMD (Head Mounted Display) with a gyro sensor. When the user uses a HMD, the omni-directional video viewer draws the video on full screen by setting the "full screen" parameter.

2.2. Omni-directional video contents

There are two kinds of omni-directional video contents in present implementation: stored video contents encoded in advance and live video contents encoded in real-time. The stored video contents mainly consist of high-resolution omni-directional videos heavy for network. Note that stored video contents can be provided as an on-demand-service. The live video contents are used for the purpose of providing multiple users with the same contents simultaneously just like TV

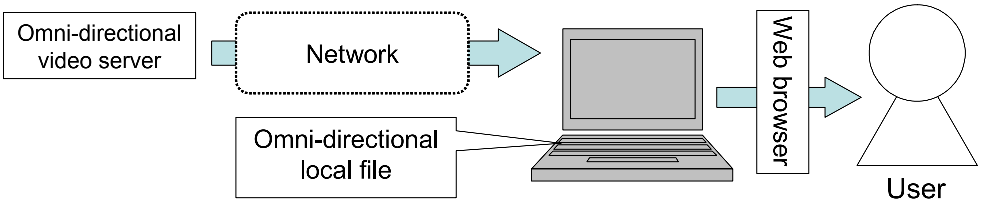

Fig. 1. Overview of proposed system

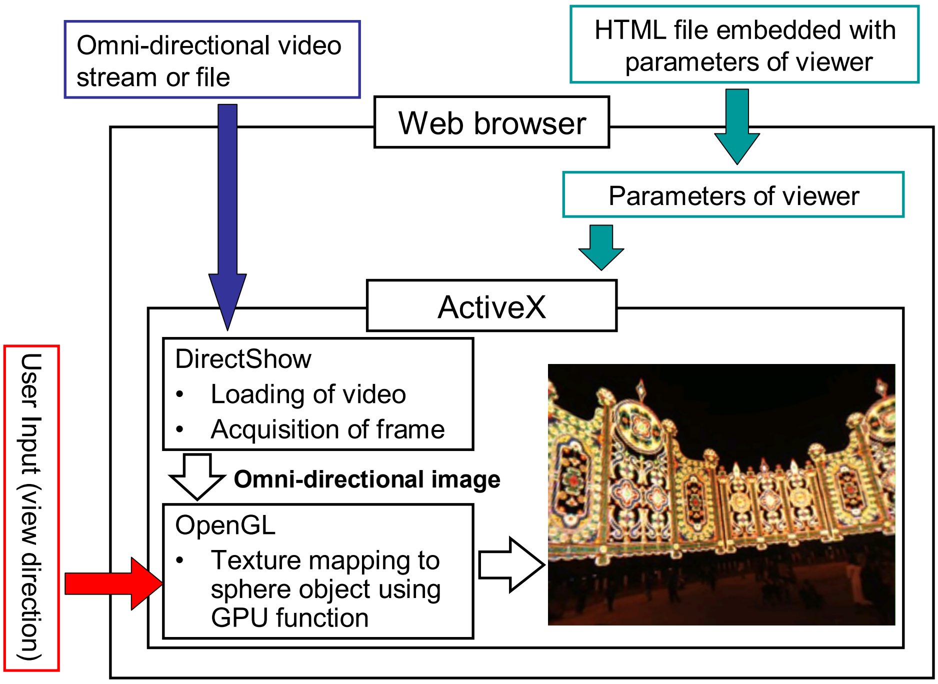

Fig. 2. Structure of omni-directional video viewer

broadcasting. It is difficult to transfer a high-resolution video because of the limit of standard network bandwidth. The user can see the live video contents acquired by an omni-directional camera and transferred immediately. It can be transferred to many sites by multi-cast without increasing the network load.

3. EXPERIMENTS

We have implemented the proposed system and experiments with both stored video contents and live video contents. Stored video contents consist of high-resolution videos obtained by using an omni-directional multi-camera system. Live video contents are acquired by using an omni-directional camera mounted on a car and that are transferred via wireless and wired network in real time.

A) Stored video contents

We acquired the omni-directional video by an omni-directional multi-camera system; Ladybug[3] (see Fig .3) and stored it in a PC for presentation (see Table 1). The camera unit of Ladybug consists of six cameras (Fig. 3(left)): Five configured in a horizontal ring and one pointing vertically. Fig. 3(right) shows a storage unit which consists of an array of HDD. The camera system can acquire video covering 75% of full spherical view (see Fig. 4). The acquired video has the size of 3840x1920 pixels and is captured at 15fps. In this experiment, we shrink the video to 1024x512 pixels

Fig. 3. Omni-directional multi-camera system; Ladybug

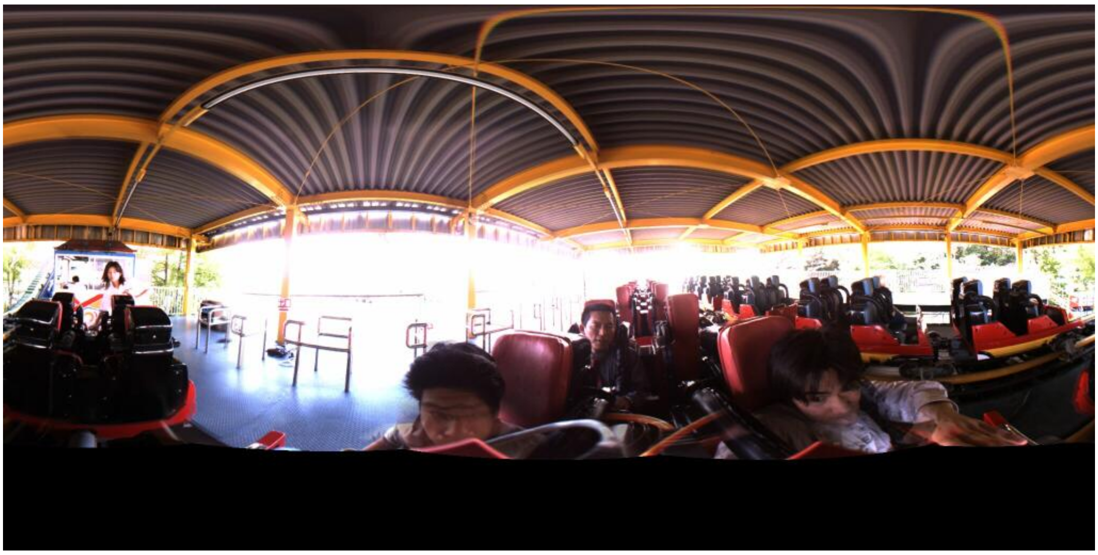

Fig. 4. High-resolution omni-directional image from Ladybug

Table 1. PC for presentation of stored video contents

CPU	Pentium4 2GHz
Memory	512MB
Graphics card	ATI RADEON9700pro
OS	Windows XP

because of the limit of HDD-access-speed of the PC. We used MPEG-1 video and MPEG-1 layer2 sounds for the formats of the video.

Fig. 5 shows a widow shot of the web browser. The user can look around the scene on the web browser. The PC can playback the video at 30fps. The user can pause and fast-forward the video with the stored video contents.

We have also implemented a view-dependent presentation system (see Fig. 6) with a HMD and a gyro sensor for realizing more rich presence. The gyro sensor is InterTrax2 made by INTERSENSE. It can acquire the user view-direction at 256Hz. The user can look around the omni-directional scene without significant delay.

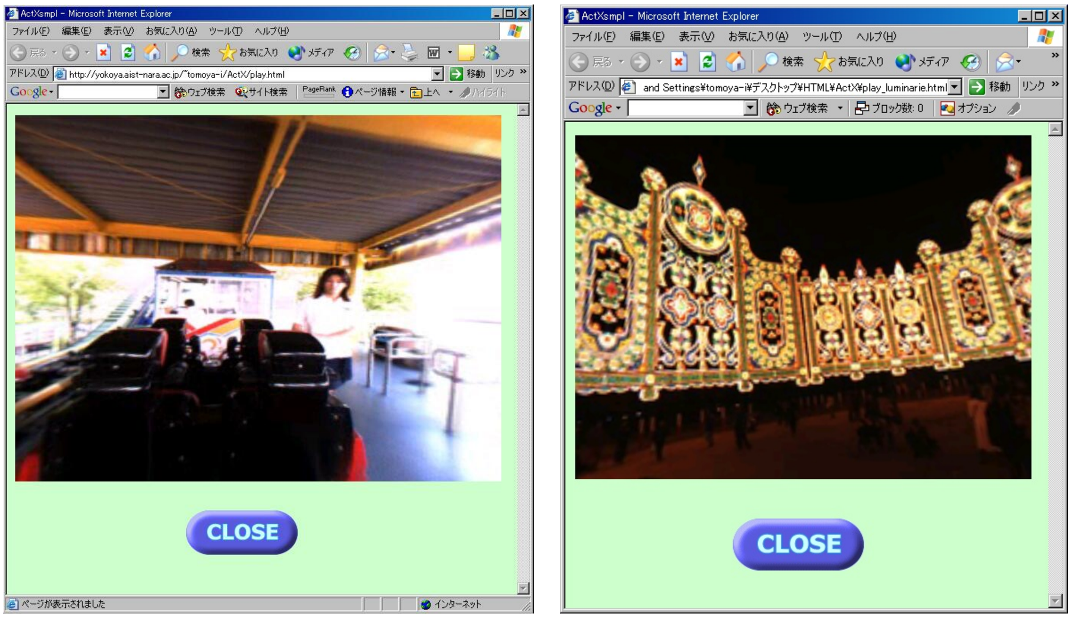

Fig. 5. Window shots of omni-directional video viewer with stored video content

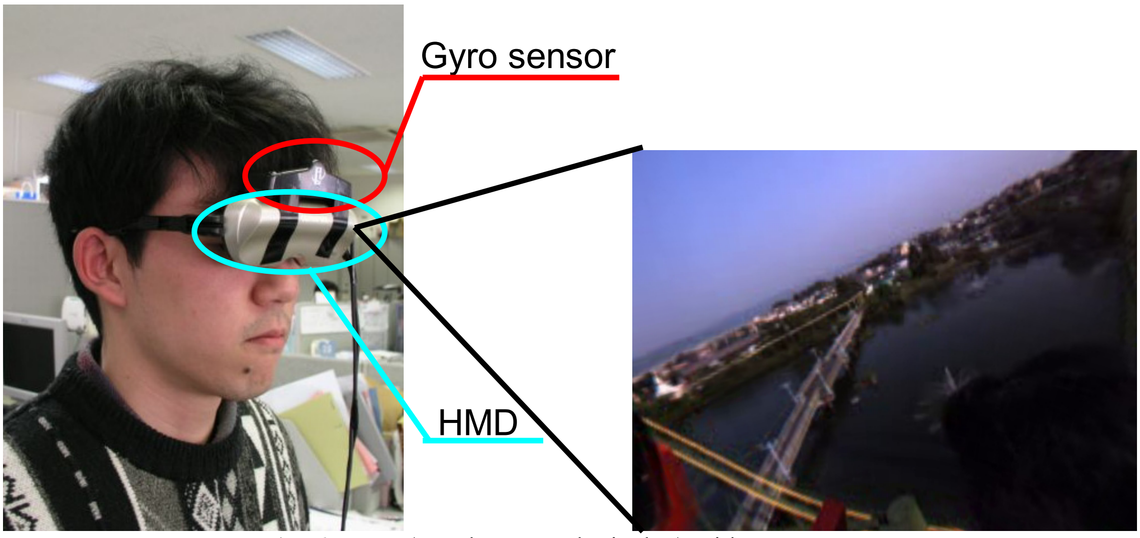

Fig. 6. HMD(Head Mounted Display) with gyro sensor

B) Live video contents

In the experiment of live video contents, the system consists of a car which mounts omni-directional camera HyperOmni Vision[4], multi-cast relay server of video, omni-directional video viewer, and network (see Fig. 7).

The car mounting the omni-directional camera acquires omni-directional progressive video surrounding the car, running in our campus. The acquired omni-directional video transferred to a PC for encoding in the through i.Link. The PC encodes the omni-directional video (640x480 pixels, 30fps) to Windows Media Format (1Mbps) by Windows Media Encoder[5]. The encoded omni-directional video is transferred to the indoor relay server through IEEE802.11a or g network via wireless network. Table 2 describes the configuration of the system for acquiring omni-directional video streams. Fig. 8 shows the car which mounts the omni-directional camera and the system which mounted on the car.

The transferred omni-directional video is received by the multi-cast relay server of omni-directional video. The relay server distributes the omni-directional video by multi-cast such as RTSP protocol. The distributed omni-directional video is seen by using the same omni-directional video viewers as for stored video contents on the web browsers. When many viewers receive the omni-directional video, the load of network does not increase because of using not unicast but multi-cast.

In the experiment, actually four PC received the distributed omni-directional video. The four users could look around the scene in arbitrary directions. Fig. 9 and Fig. 10 show examples of windows shots of the omni-directional video viewer. The video is displayed on the web browser at 30fps. The time delay between the acquisition and the presentation omni-directional video is 10 seconds. In this time, the both transmitting and receiving network loads of the relay server are 1Mbps. When the number of received omni-directional video viewer increased, the network load did not increase.

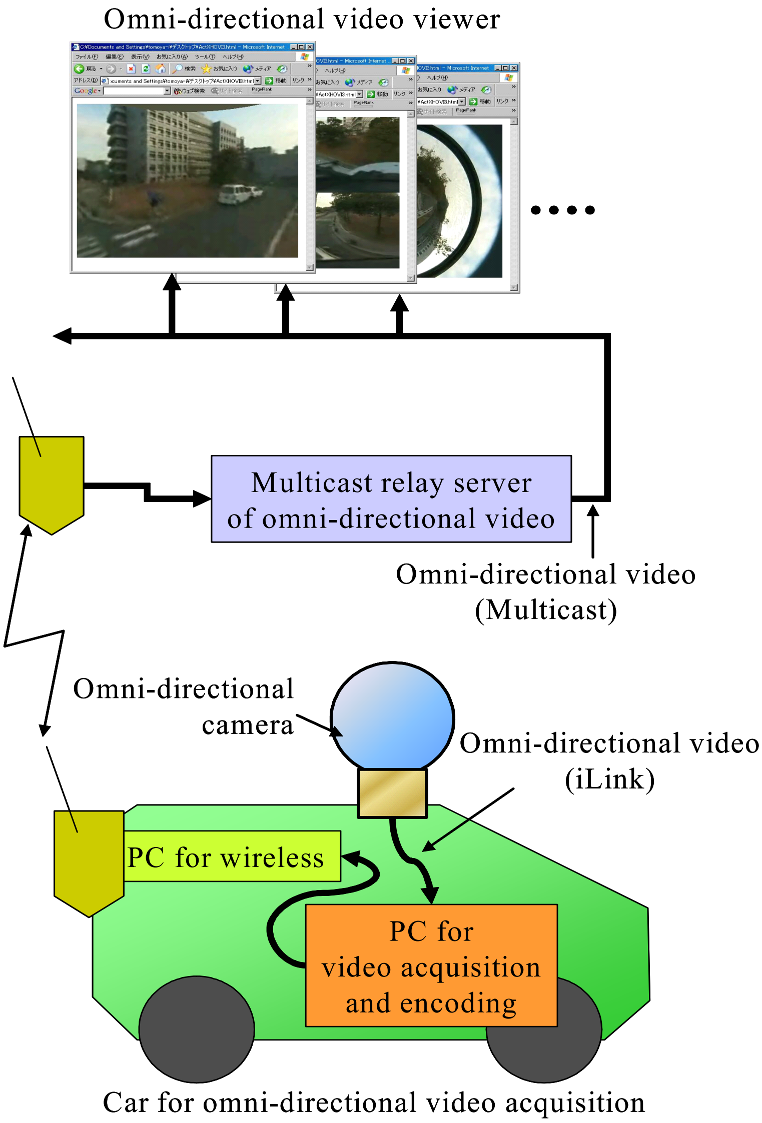

Fig. 7. Telepresence system with omni-directional camera mounted on car

Table 2. Facilities for omni-directional video acquisition in outdoor environments

Omni-directional camera	SONY DCR-TRV900 + Hyperboloidal mirror (field of view : 30 degrees upper)
PC for video acquisition and encoding	Pentium4 2.53GHz Memory 1GB Windows XP
Wireless network	IEEE802.11a and g
Car	Nissan ELGRAND (see Fig.8)

Fig. 8. Appearance of car for omni-directional video acquisition

common perspective image Panorama image omni-directional image

Fig. 9. Window shots of omni-directional video viewer (user 1)

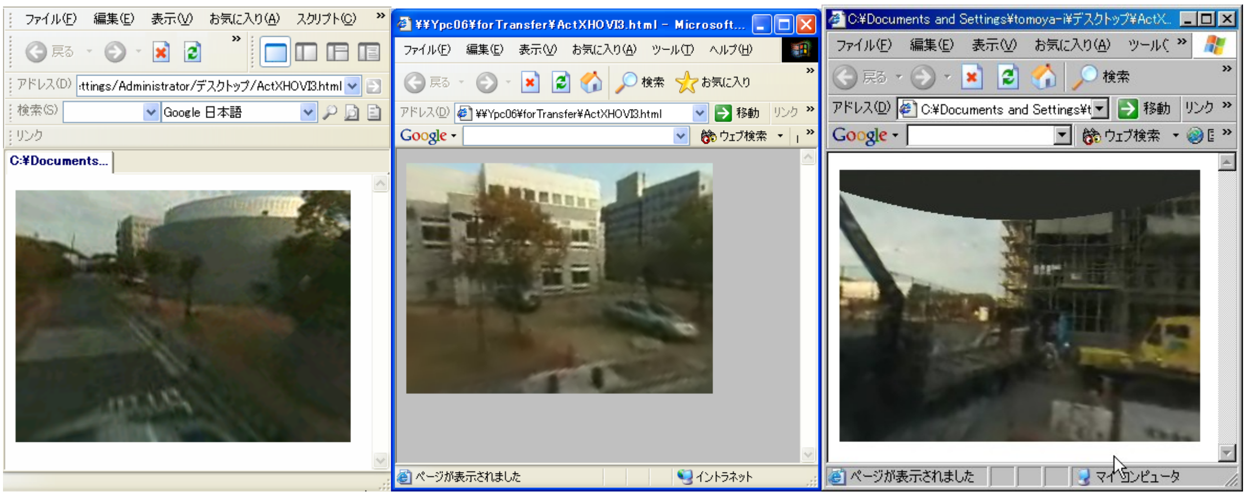

Fig. 10. Window shots of omni-directional video viewer (other users)

4. CONCLUSION

We have developed a new networked telepresence system which easily enables multiple users to look around a remote scene with omni-directional camera. The system uses a web browser, and enables users to see omni-directional video such as common video. In the experiment of stored video contents, the user could see the high-resolution omni-directional video was distributed through wireless and wired network by multi-cast protocol, and multiple users could look around the scene in arbitrary directions in real time. In the experiment of live video contents, the omni-directional camera was NTSC. The resolution of the omni-directional video was low. Thus the presence was not rich enough. On the other hand, in the experiment of stored video contents, and omni-directional multi-camera system was employed for acquiring a high-resolution video. It was not suitable for distributing an omni-directional live video. Therefore omni-directional camera should be high resolution and should not be multi-camera system. The omni-directional HD camera can acquire omni-directional high-resolution live video. In the experiment of live video contents, the delay between the acquisition and the presentation omni-directional video is 10 seconds. It is difficult to use the system for communication with a remote user. In future work, we should reduce the delay in transmitting an omni-directional video stream.

REFERENCES

1. S.Moezzi, "Special issue on immersive telepresence", *IEEE Multimedia*, vol.4, pp.17-56, 1997.
2. Y.Onoe, K.Yamazawa, H.Takemura, and N.Yokoya, "Telepresence by Real-time View-dependent Image Generation from Omnidirectional Video Streams", *Computer Vision and Image Understanding*, vol.71, pp154-165, 1998.
3. S.Ikeda, T.Sato, M.Kanbara, and N.Yokoya, "Immersive Telepresence System Using High-resolution Omnidirectional Movies and a Locomotion Interface", *Proc. SPIE Electronic Imaging*, vol.5291, 2004.
4. K.Yamazawa, Y.Yagi, and M.Yachida, "Omnidirectional imaging with hyperboloidal projection", *Proc. Int. Conf. on Intelligent Robots and Systems*, vol.2, 1993.
5. Microsoft Corporation, Windows Media Encoder 9 Series, http://www.microsoft.com/windowsmedia/9series/encoder/default.aspx