Web-Based Telepresence System Using Omni-directional Video Streams

Kazumasa Yamazawa[1], Tomoya Ishikawa[1], Tomokazu Sato[1], Sei Ikeda[1], Yutaka Nakamura[2], Kazutoshi Fujikawa[2], Hideki Sunahara[2], and Naokazu Yokoya[1]

[1] Graduate School of Information Science, Nara Institute of Science and Technology,
8916-5 Takayama, Ikoma, Nara, Japan,
{tomoya-i,yamazawa,tomoka-s,sei-i,yokoya}@is.naist.jp,
WWW home page: http://yokoya.naist.jp/

[2] Information Technology Center, Nara Institute of Science and Technology,
8916-5 Takayama, Ikoma, Nara, Japan,
{yutaka-n,fujikawa,suna}@itc.naist.jp,
WWW home page: http://inet-lab.aist-nara.ac.jp/index_eng.html

Abstract. Recently, the telepresence which allows us to experience a remote site through a virtualized real world has been investigated. We have already proposed telepresence systems using omni-directional images which enable the user to look around a virtualized dynamic real scene. The system gives the feeling of high presence to the user. However, conventional systems are not networked but are implemented as stand-alone systems that use immersive displays. In this paper we propose a networked telepresence system which enables the user to see a virtualized real world easily in network environments. This paper describes the system implementation using real-time and stored omni-directional video streams as well as its experiments.

1 Introduction

Recently, there are many researches of telepresence that acquires a dynamic real world into a virtual world and enables the user to be immersed in the remote envi-ronment[1]. The telepresence can be applied to various fields such as entertainment, medical service, and education.

We have already proposed telepresence systems which use omni-directional cam-era and enable the user to look around the scene in order to increase the presence in telepresence[2, 3]. The systems acquire and transfer the remote omni-directional scene by an omni-directional camera, and show the user view-dependent images in real time. They have no significant delay from the rotation of user's head to the pres-entation of images. They also enable the user to look around the remote scene widely. However, the conventional systems are implemented as stand-alone systems with immersive displays and require special programs and equipments. It is difficult for multiple users at remote sites to look around the same scene with the conventional systems.

In this paper, we propose a new telepresence system which enables multiple users on network to look around remote environments captured by omni-directional cam-eras. The system uses web-browser and enables users to see omni-directional videos interactively.

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 Web-based telepresence system using omni-directional videos

The schema of proposed system is illustrated in Fig.1. Omni-directional videos are stored in a 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.

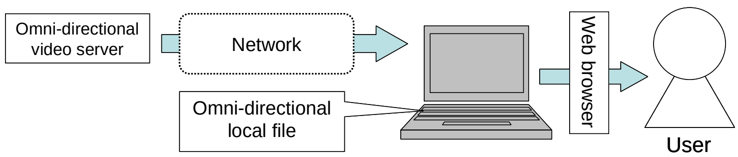

Fig. 1. Overview of web-based telepresence system.

2.1 Omni-directional video viewer on 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 distrib-uted by an automatic install function. Thus we implement an omni-directional video viewer for telepresence on a web browser in this study.

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. The omni-directional video viewer is started by the web browser, converts the omni-directional video to the common perspective video, and shows the user the common perspective video on the web page. The system uses a hardware texture mapping function for converting video in real-time by the method[2].

When the user access the web page which provides an omni-directional video con-tent, the omni-directional video viewer implemented by Active-X is installed auto-matically. The user can see an omni-directional video content without care of an omni-directional camera type, parameters of camera, file-path of the content, and so on, because a content provider embeds them in a HTML file.

The user can look around the omni-directional video by using a mouse-drag opera-tion. 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.

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 multi-ple users with the same contents simultaneously just like TV broadcasting. It is diffi-cult 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 con-tents 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 (see Fig.2)[3] and stored it in a PC for presentation (see Table1). The camera unit of Ladybug consists of six cameras (Fig.2(left)): Five configured in a horizontal ring and one pointing vertically. Fig.2(right) shows a storage unit which consists of an array of HDD. The camera system can acquire video covering 75% of the full spherical view. The acquired video has the size of 3840x1920 pixels and is captured at 15fps. In this experiment, we shrink the video to 1024x512 pixels because of the limit of HDD-access-speed of the PC. We used MPEG-1 video and MPEG-1 layer 2 sounds for the formats of the video.

Fig.3 shows a window 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.4) with a HMD and a gyro sensor for realizing more rich presence. The gyro sensor is Inter-Trax2 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. 2. Omni-directional multi-camera system; Ladybug.

Table 1. PC for presentation of stored video contents.

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

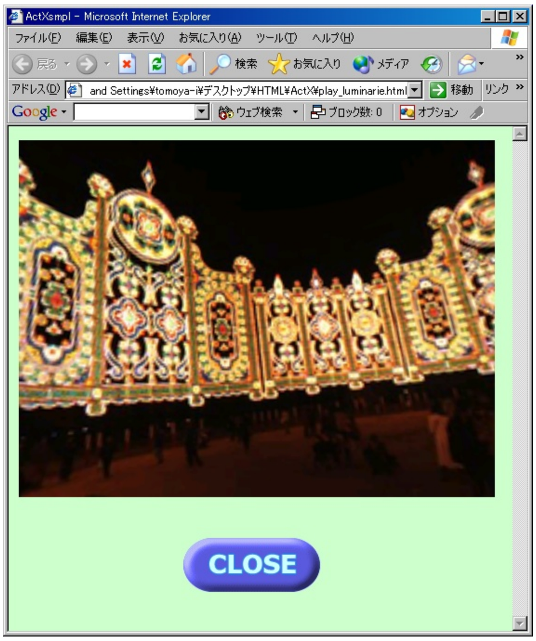

Fig. 3. Window shot of omni-directional video viewer with stored video content.

Fig. 4. 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], multicast relay server of video, omni-directional video viewer, and network (see Fig.5).

The car mounting the omni-directional camera acquires omni-directional progressive video surrounding the car, running in our campus. The acquired omni-directional video is transferred to a PC for encoding in the car through iLink. The PC encodes the omni-directional video (640x480 pixels, 30 fps) to Windows Media Format (1Mbps) by Windows Media Encoder[4]. 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.6 and Fig.7 show 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 multicast relay server of omni-directional video. The relay server distributes the omni-directional video by multicast 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 multicast.

In the experiment, actually four PC received the distributed omni-directional video. The four users could look around the scene in arbitrary directions. Fig.8 shows ex-amples of windows shots of the omni-directional video viewer. The video is dis-played 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 re-ceived omni-directional video viewer increased, the network load did not increase.

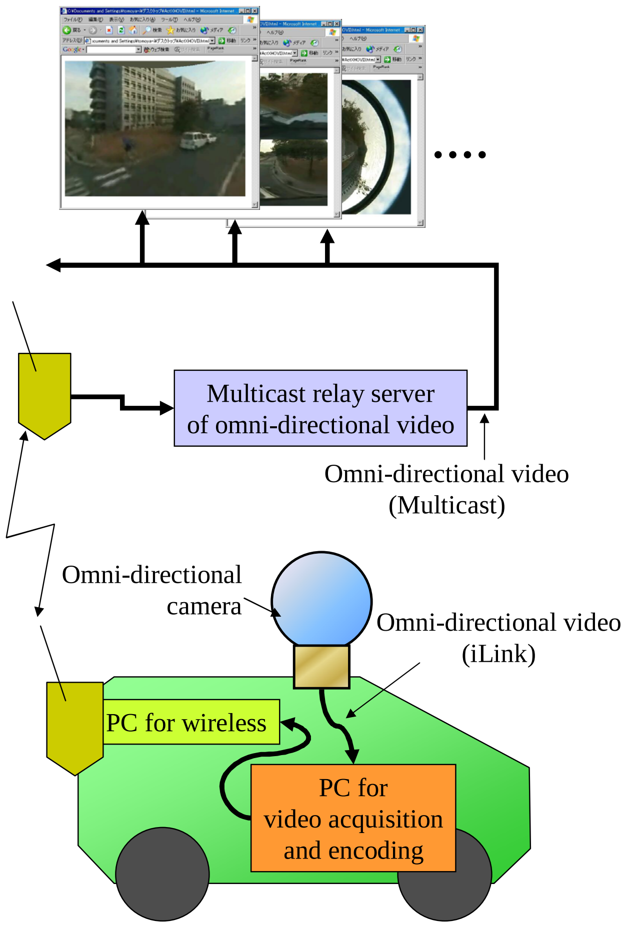

Fig. 5. Telepresence system with omni-directional camera mounted on car.

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 WindowsXP
Wireless network	IEEE802.11a and g
Car	Nissan ELGRAND (see Fig.6, Fig.7)

Fig. 6. Appearance of car for omni-directional video acquisition.

Fig. 7. Window shots of omni-directional video viewer.

Common perspective image

Panorama image
(Cylindrical image)

Omni-directional image
(Same as input image)

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

4 Conclusion

We have developed a new networked telepresence system which easily enables multi-ple 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. In the experiment of live video contents, the omni-directional video was distributed through wireless and wired network by multicast 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 e-nough. On the other hand, in the experiment of stored video contents, an 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. Moezzi, S., Ed.: Special issue on immersive telepresence. IEEE MultiMedia. **4** 1 (1997) 17–56
2. Onoe, Y., Yamazawa, K., Takemura, H., and Yokoya, N.: Telepresence by Real-time View-dependent Image Generation from Omnidirectional Video Streams. Computer Vision and Image Understanding. **71** 2 (1998) 154–165
3. Ikeda, S., Sato, T., Kanbara, M., and Yokoya, N.: Immersive telepresence system using high-resolution omnidirectional movies and a locomotion Interface. Proc. SPIE Electronic Imaging. **5291** (2004).
4. Yamazawa, K., Yagi, Y., and Yachida, M.,: Omnidirectional imaging with hyperboloidal projection. Proc. Int. Conf. on Intelligent Robots and Systems. **2** (1993) 1029–1034
5. Microsoft Corporation, Windows Media Encoder 9 Series. `http://www.microsoft.com/windows/windowsmedia/9series/encoder/`