Spacetime Freeview Generation Using Image-based Rendering, Relighting, and Augmented Telepresence

Fumio Okura
Nara Institute of Science and Technology (NAIST)
8916-5, Takayama-cho, Ikoma, Nara, Japan
fumio-o@is.naist.jp

ABSTRACT

This paper proposes an freeview generation technique providing the users to change their viewpoints beyond time and space. The study consists of three technical elements: image-based rendering, relighting, and augmented telepresence. Before now, we have developed two systems relating this study: an augmented telepresence system and a full spherical HDR aerial imaging system.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual reality

Keywords

Freeview generation, image-based rendering, relighting, augmented telepresence, aerial omnidirectional image

1. INTRODUCTION

Telepresence is a technology providing a user with a view of a remote site immersiveness and is one of the major research fields in virtual reality. In particular, offline telepresence uses preliminarily recorded omnidirectional images to provide users with views in any direction from a location and has become very important with the popularization of walk-through applications such as Google Street View [1]. Recently, there have been studies dealing with augmented telepresence [11], which provides the user with both the view of a remote site and related information by using augmented reality (AR) techniques. Typically, the offline telepresence systems including augmented telepresence systems can enable users to change the viewpoints only to the positions that the omnidirectional images have been captured.

To realize offline telepresence system enables users to change position of the viewpoints freely, there are some studies dealing with freeview generation technique which virtually generates scenes of novel viewpoints from preliminarily captured images. In particular, image-based rendering (IBR) [10] approaches which generate novel viewpoints by compositing large amount of images (e.g. reconstructing light-fields of the scene) is very popular research field in computer vision to realize freeview generation. Recently, IBR techniques are used for map applications like Microsoft Street Slide [7], broadcasting, and so on.

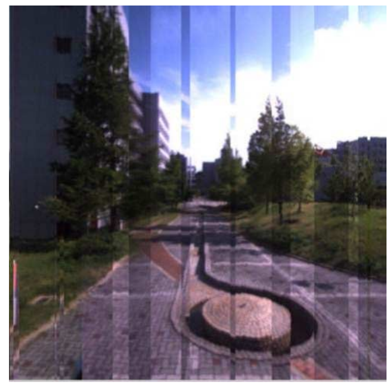


Figure 1: Artifacts in IBR freeview generation [4].

Image-based rendering usually deals with the change of viewpoint in three-dimensionally space. However, in practical use especially in large outdoor environment, artifacts in generated images can be appear due to change of lighting environment while capturing input images as shown in Fig. 1 [4].

This study resolves the problem in 3D freeview generation using IBR using relighting technique. In addition, we realize 4D freeview generation by combining augmented telepresence technique and 3D freeview generation.

2. FREEVIEW GENERATION BEYOND TIME AND SPACE

This study realizes 4D freeview generation beyond time and space in wide outdoor environment using aerial omnidirectional images by three technical element: a) IBR, b) relighting, c) augmented telepresence as shown in Fig. 2. Freeview generation using IBR is used to provide users 3D movement of the viewpoint. Images used for IBR are captured from omnidirectional cameras mounted on an unmanned airship. The change of lighting environment while capturing the images is a problem of IBR. Relighting technique which generates the scene under different lighting environment is used to remove the effect of change of lighting environment. In addition, the relighting approach can be employed on freeview generation beyond time, in particular, changing user's viewpoint to different time of day or weather by changing

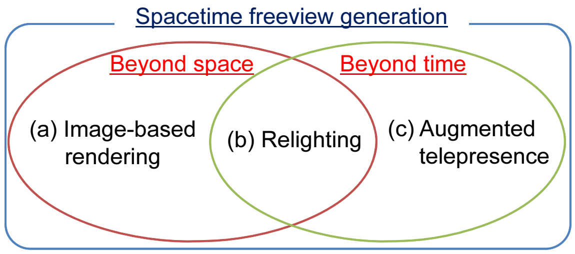


Figure 2: Approaches to realize spacetime freeview generation.

lighting environment. When you want to go to the time when the scene is different not only lighting environment, for example, the future and ancient times, augmented telepresence technique can be employed. That is, CG models simulating the future and ancient are superimposed using augmented telepresence technique.

Approaches for realizing 3D freeview generation and freeview generation beyond time are detailed below.

2.1 Freeview generation beyond space

3D freeview generation using IBR.

A view-dependent texture mapping (VDTM) [3] approach which is a IBR method using 3D models as supplementary information is used to realize 3D freeview generation. The 3D models are automatically reconstructed from the input omnidirectional images using multi-view stereo (MVS) [6]. Typically, 3D shapes reconstructed using MVS have large errors. The VDTM approach is used to generate novel viewpoints using 3D shapes with large errors.

Integration of lighting environment using relighting.

To integrate lighting environment while capturing images and remove artifacts in images under novel viewpoints, we are going to employ relighting approach which generate scenes under various lighting conditions. Relighting approaches are mainly proposed for indoor use (e.g. furniture layout simulation). Model-based approaches use accurate 3D shapes as previous knowledge [8], on the other hand, image-based approaches uses large amount of images [9]. To realize relighting in large environment, a hybrid approach may be effective. In this study, 3D shapes including large amount of errors and omnidirectional images captured under various lighting environments are used for relighting.

2.2 Freeview generation beyond time

Beyond time of day or weather using relighting.

By capturing from same area under various time of day and the weather, the relighting approaches also realize change of viewpoints among captured lighting environments. In addition, seamless transfer between two lighting environments can be realized by interpolating light sources between two environments.

Beyond age using augmented telepresence.

Augmented telepresence technique, which provides the user with both the immersive telepresence and AR, realizes change of viewpoint beyond the age: transferring to the future and ancient times. CG models simulating the future and ancient can be superimposed with higher quality than usual real-time AR systems [11].

Figure 3: 3D model reconstructed using multi-view stereo.

Before now, we have developed two systems related to spacetime freeview generation:

- augmented telepresence system without using freeview generation technique [11],
- full-spherical HDR image generation [12].

In the rest of paper, overview of each progress and results are described.

3. AUGMENTED TELEPRESENCE

3.1 Introduction

The developed system superimposes CG models of an old palace built 1,300 years ago on recorded video captured from an unmanned airship using an augmented telepresence technique [11]. The year 2010 was the 1,300th anniversary of Nara Heijo-kyo, an ancient capital in Japan. Today, almost all of the ground in Heijo-kyo is a field of grass excepting some buildings and bases that were physically restored for the commemorative events of the anniversary. The proposed system generates augmented videos by superimposing CG models of buildings that existed 1,300 years ago on the "plain grassland" omnidirectional images using an augmented telepresence technique. Omnidirectional aerial videos are used to provide users with a wider view than on the ground.

In this study, we have developed an offline augmented telepresence system using an unmanned airship and an omnidirectional multi-camera system (OMS). The proposed system resolves the geometric and photometric registration problems in order to generate high-quality augmented videos using omnidirectional videos. The omnidirectional videos captured by the OMS attached to the airship increase the immersiveness of the telepresence by providing the user with the sense of being able to look around the location. To render augmented scenes, we use image-based-lighting (IBL) and global illumination (GI) techniques with an omnidirectional environmental map.

3.2 Augmented omni video generation

To overcome the geometric registration problem, the proposed system estimates the camera position and orientation. The proposed system employs offline process which is a high-accuracy camera position and orientation estimation method [14] with a video and position data measured by GPS. Each frame of the omnidirectional video is aligned using the position and orientation of the OMS (e.g. Figure 4(a)).

The omnidirectional video has missing area including sections that is out of field of view of OMS and background scenery occluded by the airship as shown in Figure 4(a). To

(a) A captured omnidirectional image. A missing area appear at top of the image.

(b) An augmented omnidirectional image. The missing area is filled using sky model.

Figure 4: Omnidirectional images generated by the proposed augmented telepresence system.

Figure 5: Appearance of virtual tourism system.

Figure 6: A view-dependent perspective image generated from an augmented omnidirectional image.

use the omnidirectional video as an environmental map for IBL, the missing area must be filled-in and complete environmental maps must be generated. To fill-in missing areas and generate a complete environmental map, other frames in the video are searched for an area similar to the missing area, and the pixel intensities belonging to this area is copied to the pixels of the missing area. The all-sky-model [5] is used to fill-in the areas having no corresponding areas, where the background scene is occluded in all the frames. The all-sky-model is a statistical model representing luminance of various skies with some parameters. In this study, the parameters are estimated from intensities of background scene except missing areas.

Augmented video is rendered using a commercial GI rendering software 3ds Max (Autodesk, Inc.) and Mentalray (Mental Images GmbH.), as shown in Figure 4(b). The virtual objects are registered geometrically using the estimated location and position of the OMS. An omnidirectional augmented video is generated through IBL and GI rendering using a complete environmental map.

3.3 Configuration of demonstration system

Figure 5 shows appearance of the system for the demonstration using HMD and motion tracker. The augmented omnidirectional video is converted to view-dependent perspective video (see Figure 6) in real-time [13] using view direction of the user captured by an electromagnetic motion tracker FASTRAK (Polhemus, Inc.) attached to the HMD. The view-dependent perspective video is presented to the user via a HMD.

The user can freely look around the augmented scene moving viewpoint on the captured path of the omnidirectional image sequence. The system also switches sequences with and without virtual objects. A similar system was demonstrated at the Commemorative Events for the 1,300th Anniversary of Nara Heijo-kyo Capital, and over a thousand people experienced the system.

4. FULL SPHERICAL HDR IMAGING

4.1 Introduction

We developed an aerial imaging system [12] that can capture full spherical images without missing areas and with high dynamic range (HDR) for telepresence, image-based lighting (IBL) [2], and augmented telepresence [11]. We have proposed an aerial imaging system to resolve two problems when spherical images are observed by OMS for telepresence systems: 1) occurrence of missing areas, and 2) deficiency of dynamic range.

1) Missing areas sometimes appear in a spherical image as shown in Figs. 7(a) and 7(b). These areas are caused by the restriction of the field of view of the OMS and (in the case of Fig. 7) by occlusions from the airship. Such areas decrease the immersiveness in telepresence and make it impossible to use the spherical images for IBL.

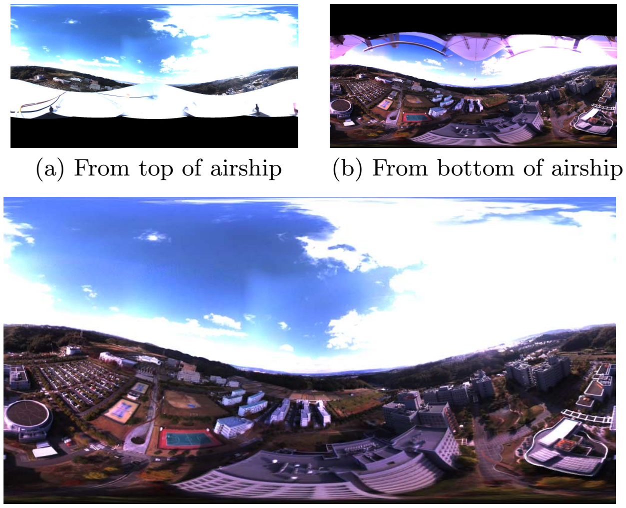
2) Spherical images without saturation of intensity are recommended for IBL. However, the dynamic range of outdoor scenes is too high to capture using standard cameras that have normal (8-bit) dynamic range, as the sun can be 2^{17} times brighter than the dark areas in the clouds or the sky.

4.2 Full spherical HDR imaging system

The proposed system generates full spherical images using multi-exposure images captured with two OMS mounted on the top and bottom of an unmanned airship. The process of the proposed system is divided into three parts as follows.

Multi-exposure imaging using two OMS.

Multi-exposure images are captured with multiple shutter speeds using two OMS mounted on the top and bottom faces of an unmanned airship. Since the camera atop the vehicle mainly captures the sun and sky, ND filters are attached to that camera. In addition, shutter speeds are controlled automatically so as to capture multi-exposure images effectively.



(a) From top of airship (b) From bottom of airship

(c) Full spherical image

Figure 7: Full spherical HDR image generated from multi-exposure images captured by two omnidirectional cameras.

HDR image generation from multi-exposure images.

HDR spherical images from two OMS are respectively generated using multi-exposure images. There are misalignments among the multi-exposure images caused by the changing position and orientation of the camera. We deal with misalignments due to change in the orientation of the camera, which dominantly affects the amount of misalignment. These are corrected by estimating the change in camera orientation from the captured image sequence.

Combination of HDR images from two cameras.

Full spherical HDR images are generated from HDR images from two different OMS. In this study, it is difficult to fix the relative geometric relationship between the cameras, because we use an unmanned airship that has a deformable body. The two HDR images are aligned by estimating the relative rotation between the cameras in every frame. In addition, we correct a chromatic change that occurs in the images captured with ND filters.

4.3 Experiments

We have conducted experiments to generate a full spherical HDR image using both still images and video sequences captured using an unmanned airship above our campus. Full spherical HDR still images were successfully generated as shown in Fig. 7(c). It has also been confirmed that a full spherical HDR video can be generated with the position and orientation information of the camera for applying to augmented telepresence.

5. CONCLUSION

This paper focused on an freeview generation technique enabling users to change viewpoints beyond time and space. The study consists of three technical elements: image-based rendering, relighting, and augmented telepresence. Although we have developed two systems relating this study: augmented telepresence system and full spherical HDR aerial imaging system, the study is in progress. In future, IBR and relighting technique are intended to be developed, and the working system of spacetime freeview generation system in large outdoor environment should be realized in the end.

6. ACKNOWLEDGMENTS

This research is partially supported by Grant-in-Aid for Scientific Research (A), No. 23240024 from MEXT and by the "Ambient Intelligence" project funded by MEXT.

7. REFERENCES

[1] D. Anguelov, C. Dulong, D. Filip, C. Frueh, S. Lafon, R. Lyon, A. Ogale, L. Vincent, and J. Weaver. Google street view: Capturing the world at street level. IEEE Computer Magazine, 43(6):32–38, 2010.

[2] P. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In Proc. ACM SIGGRAPH'98, pages 189–198, 1998.

[3] P. Debevec, C. Taylor, and J. Malik. Modeling and rendering architecture from photographs: A hybrid geometry-and image-based approach. In Proc. ACM SIGGRAPH'96, pages 11–20, 1996.

[4] M. Hori, M. Kanbara, and N. Yokoya. Arbitrary stereoscopic view generation using multiple omnidirectional image sequences. In Proc. 20th IAPR Int'l Conf. on Pattern Recognition (ICPR'10), pages 286–289, 2010.

[5] N. Igawa, Y. Koga, T. Matsuzawa, and H. Nakamura. Models of sky radiance distribution and sky luminance distribution. Solar Energy, 77:137–157, 2004.

[6] M. Jancosek and T. Pajdla. Multi-view reconstruction preserving weakly-supported surfaces. In Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR'11), pages 3121–3128, 2011.

[7] J. Kopf, B. Chen, R. Szeliski, and M. Cohen. Street slide: browsing street level imagery. In Proc. ACM SIGGRAPH'10, pages 96:1–96:8, 2010.

[8] C. Loscos, G. Drettakis, and L. Robert. Interactive virtual relighting of real scenes. IEEE Trans. on Visualization and Computer Graphics, 6(4):289–305, 2000.

[9] V. Masselus, P. Peers, P. Dutré, and Y. Willems. Relighting with 4d incident light fields. In Proc. ACM SIGGRAPH'03, pages 613–620, 2003.

[10] L. McMillan and G. Bishop. Plenoptic modeling: An image-based rendering system. In Proc. ACM SIGGRAPH'95, pages 39–46, 1995.

[11] F. Okura, M. Kanbara, and N. Yokoya. Fly-through Heijo palace site: Augmented telepresence using aerial omnidirectional videos. In Proc. ACM SIGGRAPH'11 Posters, page 78, 2011.

[12] F. Okura, M. Kanbara, and N. Yokoya. Full spherical high dynamic range imaging from the sky. In Proc. IEEE Int'l Conf. on Multimedia and Expo (ICME'12), pages 325–332, 2012.

[13] 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, 71(2):154–165, 1998.

[14] Y. Yokochi, S. Ikeda, T. Sato, and N. Yokoya. Extrinsic camera parameter estimation based-on feature tracking and GPS data. In Proc. Seventh Asian Conf. on Computer Vision (ACCV'06), volume 1, pages 369–378, 2006.