Arbitrary Stereoscopic View Generation
Using Multiple Omnidirectional Image Sequences

Maiya Hori*, Masayuki Kanbara*, Naokazu Yokoya*
*Graduate School of Information Science, Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0192, Japan
Email: {maiya-h, kanbara, yokoya}@is.naist.jp

Abstract—This paper proposes a novel method for generating arbitrary stereoscopic view from multiple omnidirectional image sequences. Although conventional methods for arbitrary view generation with an image-based rendering approach can create binocular views, positions and directions of viewpoints for stereoscopic vision are limited to a small range. In this research, we attempt to generate arbitrary stereoscopic views from omnidirectional image sequences that are captured in various multiple paths. To generate a high-quality stereoscopic view from a number of images captured at various viewpoints, appropriate ray information needs to be selected. In this paper, appropriate ray information is selected from a number of omnidirectional images using a penalty function expressed as ray similarity. In experiments, we show the validity of this penalty function by generating stereoscopic view from multiple real image sequences.

Keywords-stereoscopic image; novel view generation; image-based rendering; omnidirectional image;

Figure 1. Flow diagram of proposed method.

I. INTRODUCTION

As represented by 3D TV, 3D applications and systems using stereoscopic views instead of flat 2D images have been investigated [1]. To have stereoscopic vision, the images seen from the points that have a binocular disparity are needed. Although using a stereo camera or a particular camera that can capture stereo images make stereoscopic views for users [2], [3], they can have stereoscopic vision only at captured points. These approach can not change viewpoint freely. Many algorithms for generating novel views from captured images have been developed to change the viewpoints freely. Some researchers reconstruct a 3D model with a model-based rendering (MBR) approach to generate images that have a binocular disparity. In these techniques, a high-quality 3D model without a missing region is needed to change viewpoints freely [4]. However, it is difficult to build such a high-quality 3D model automatically in an outdoor environment [5] . On the other hand, many works have been conducted for generating binocular views by applying an image-based rendering (IBR) approach [6] to binocular viewpoints [7], [8]. Although these works generate stereoscopic view at some points, positions and directions of viewpoints are limited to a small range depending on the captured points. To generate high-quality stereoscopic views in a large are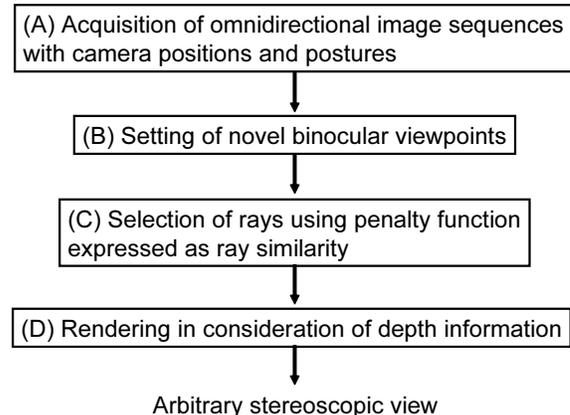a, much ray information is needed because the novel view is composed of parts of several rays captured at different points.

In this research, we attempt to generate arbitrary stereoscopic views in an outdoor environment with an IBR approach. To generate arbitrary stereoscopic views in a large area, omnidirectional image sequences are captured on multiple paths. Appropriate ray information for generating a high-quality stereoscopic view is selected from a number of omnidirectional images using a penalty function expressed as ray similarity.

The next section (Section II) of this paper explains a method for generating the arbitrary stereoscopic views using a penalty function based on ray similarity. In Section III, we demonstrate experimental results of stereoscopic images in an outdoor environment. Section IV describes our conclusion and proposals for future works.

II. STEREOSCOPIC VIEW GENERATION USING MULTIPLE OMNIDIRECTIONAL IMAGE SEQUENCES

This section describes a method for generating arbitrary stereoscopic views from multiple omnidirectional image sequences. Figure 1 shows the flow diagram of the proposed method. First, omnidirectional image sequences together with camera positions and postures are acquired in an outdoor environment in phase (A). Next, the binocular view-

points are set for generating stereoscopic view at arbitrary viewpoints in phase (B). Parts of images for synthesizing the novel stereoscopic view are selected from accumulated omnidirectional images using a penalty function expressed as ray similarity in phase (C). Finally, the stereoscopic view is rendered in consideration of depth information in phase (D).

A. Acquisition of omnidirectional images with camera positions and postures

Much ray information is required for generating a high-quality stereoscopic view with an IBR approach at arbitrary viewpoints. We use an omnidirectional multi-camera system to acquire much ray information. In addition, much ray information is acquired by capturing multiple omnidirectional image sequences on various paths repeatedly as shown in Figure 2. In our approach, not only omnidirectional images but also accurate camera positions and postures are required for generating the stereoscopic view with an IBR approach. We assume that the accurate camera positions and postures are acquired with various sensors or by using vision-based approaches such as structure from motion [9], synchronizing with image sequences.

B. Selection of rays using penalty function expressed as ray similarity

As illustrated in Figure 2, we set binocular viewpoints whose distance corresponds to the human eye distance and these moves freely in an area where a value of penalty described later is small. In this study, the stereoscopic view at novel viewpoints is generated from precaptured omnidirectional images using an approach based on light field rendering [10]. Novel view is synthesized using much ray information which is acquired at different points. The image at the novel viewpoint is divided into many slits of one pixel in width and the ray needed to generate the novel view is selected in every slit. In Figure 3, the novel view can be synthesized from three omnidirectional images captured at O_{A_2}, O_{B_2} and O_{A_3}. As a result, the novel view is vertically divided into three areas. If there is no ray information to generate stereoscopic view because omnidirectional images are captured at just discrete positions or some omnidirectional images may contain the same scene, appropriate omnidirectional images are selected for the novel view using a penalty function as follows.

$$P_{slit} = P_{ang} + wP_{pos}, \qquad (1)$$

where P_{ang} is a penalty that means a perpendicular distance from an image captured position to a ray required for generation. This penalty aims at improvement in quality of stereoscopic vision by selecting the appropriate ray which is required for generating the novel view. P_{pos} is a penalty which means the distance between the viewpoint that generates the novel image and the position where the image used
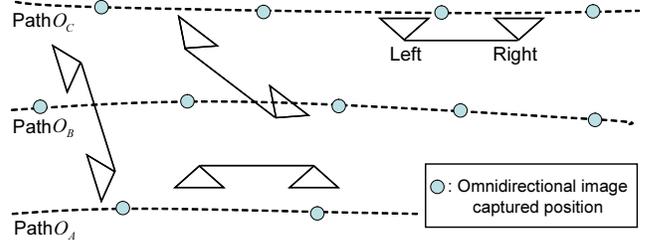

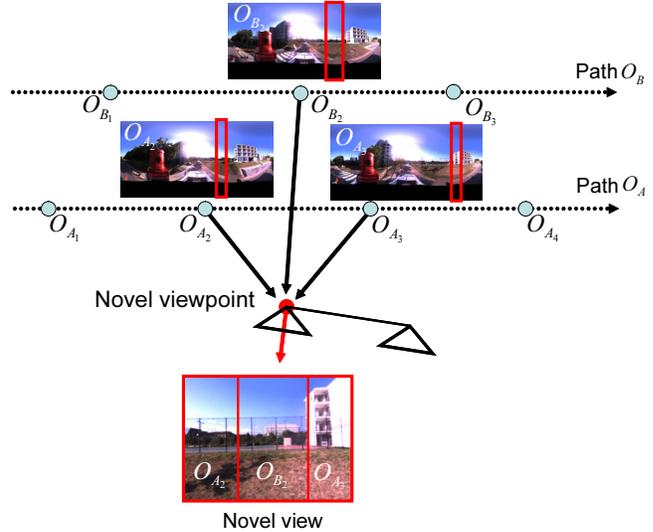Figure 2. Setting of novel binocular viewpoints around camera paths (top view).

Figure 3. Novel view synthesized from precaptured omnidirectional images (top view). Novel view can be synthesized from three omnidirectional images captured at O_{A_2}, O_{B_2} and O_{A_3}.

for generation is acquired. This penalty aims at equalization of resolution and connection of edges between the slits in generated images using omnidirectional images that are captured at a nearby area. w is a weighting coefficient. The penalty is calculated in each slit and an omnidirectional image in which penalty becomes the smallest is used for synthesizing the stereoscopic view.

In addition, the sum of penalties computed in all the slits is used as an indication of quality of the generated stereoscopic view as follows.

$$P_{sum} = \sum_{slit=1}^{width} P_{slit}, \qquad (2)$$

where P_{slit} is a penalty in each slit described in Eq. (1), $width$ represents the number of slits of a generated image. We generated stereoscopic view in the areas where the penalty P_{sum} is below a threshold value that makes high-quality stereoscopic vision to users.
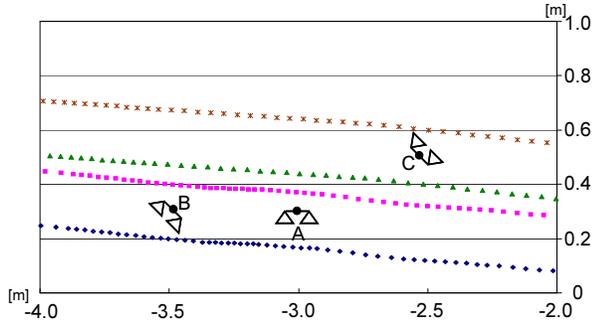
Figure 4. Positions where omnidirectional images are captured (top view). In the experiments, we show the result of stereoscopic views at point A, B and C.

Figure 5. Distribution of sum of penalties, P_{sum} (top view). The region surrounded by red lines has a small penalty.

C. Rendering in consideration of depth information

In this work, the novel view is synthesized from multiple omnidirectional images, but only one image is used in each slit. This is because it is easy to perform stereoscopic vision with the generated images that have little artifact noises. Here it should be note that an object scale may be different slit by slit because the novel view is synthesized from multiple images captured at different points. This problem with the object scale in a generated image is overcome by rendering in consideration of estimated depth information. The depth values are estimated with a standard stereo matching using images that are captured at different positions.

III. Experiments and results

To verify the validity of the proposed method, we generated arbitrary stereoscopic view in an outdoor environment. In experiments, omnidirectional image sequences were captured with a vehicle-mounted omnidirectional multi-camera system for which we used the Ladybug2 (Point Grey Research). The camera positions and postures are acquired with a hybrid sensor that consists of a real-time kinematic global positioning system (RTK-GPS) (LogPakII, Nikon-Trimble, horizontal accuracy, $\pm$ 3.0cm, vertical accuracy, $\pm$ 4.0cm) and an inertial navigation system (TISS–5–40, Tokimec). A part of horizontal trajectories of an omnidirectional multi-camera system in a wide area is shown in Figure 4. This figure shows that omnidirectional images are captured densely in an outdoor environment. Figure 5 shows a distribution of penalties P_{sum} in Eq. (2) when the stereoscopic view is generated at each point in the map. As shown in this figure, it turns out that a penalty P_{sum} is small at the positions where an omnidirectional image is captured densely. We generated stereoscopic views in region where penalty $P_{sum} < 150$. This region is expressed in the region surrounded by red lines in Figure 5.

Examples of stereoscopic views at point A in Figure 4 are given in Figure 6. These figures show the results of stereoscopic view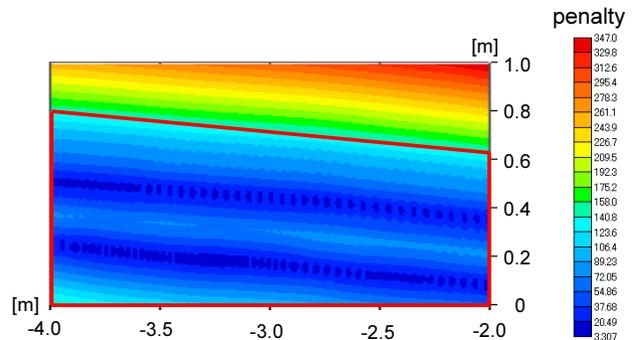 using weighting coefficient $w = 0$ and $w = 0.5$ in Eq. (1). As shown in these figures, generated images with $w = 0$ are difficult to have stereoscopic vision due to disconnection of edges between the slits. This results show that penalty P_{pos} is needed to generate stereoscopic view correctly. Examples of stereoscopic views when the users change the direction of view at point B and C in Figure 4 are given in Figure 7. These figures show the results of stereoscopic view using weighting coefficient $w = 0.5$. As shown in these figures, it turns out that the change of direction of view can be performed freely. It can be observed that the proposed penalty function can make stereoscopic views correctly. However, the sense of incongruity due to the difference of a color tone appeared between slits. This is because the stereoscopic view is synthesized from several images that are captured at different points and times.

IV. Conclusion

In this paper, we have proposed a method for generating arbitrary stereoscopic views from multiple omnidirectional image sequences in an outdoor environment. In our approach, much ray information is acquired at many points by capturing omnidirectional image sequences along arbitrary multiple paths and appropriate ray information is selected for generating stereoscopic views using the penalty function expressed as ray similarity. In experiments, we have confirmed that the proposed penalty function can make stereoscopic views correctly.

When omnidirectional images are captured at different points and times, the sense of incongruity occurred due to the difference of color tone appearing between slits. We have to investigate a method for eliminating the difference of color tone among input images. When omnidirectional image sequences are captured in an outdoor environment, moving objects such as humans and vehicles are often observed. In order to generate novel stereoscopic views in a dynamic environment, eliminating moving objects from the omnidirectional image sequences is required.
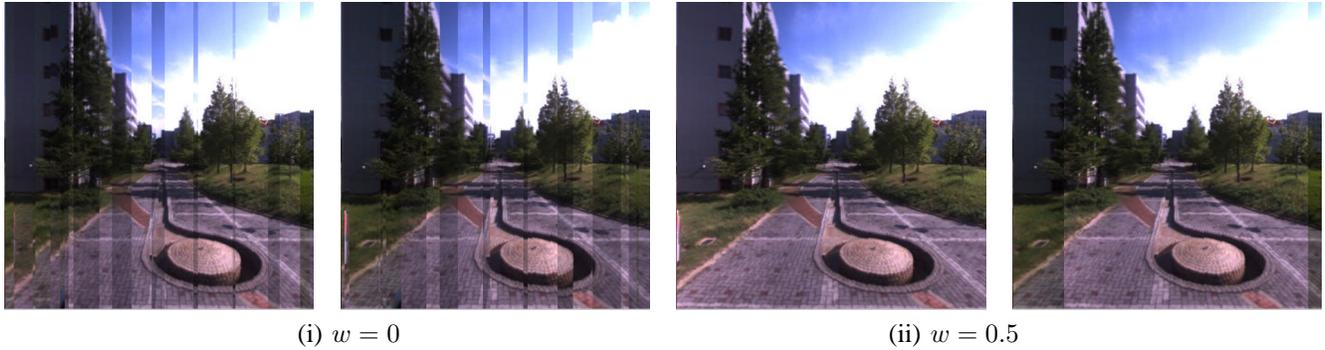
(i) $w = 0$ (ii) $w = 0.5$

Figure 6. Examples of stereoscopic view with various weighting coefficient w at point A (each left side is left view and right side is right view). Generated images with $w = 0$ are difficult to have stereoscopic vision due to disconnection of edges between the slits.

(i) at point B   (ii) at point C

Figure 7. Examples of stereoscopic view when the users change position and direction of viewpoint (each left side is left view and right side is right view). Change of direction of view can be performed freely.

REFERENCES

[1] W. Matusik and H. Pfister, "3D TV: A Scalable System for Real-Time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes," *Proc. of ACM SIGGRAPH 2004*, vol. 23, pp. 814–824, 2004.

[2] S. Peleg, M. Ben-Ezra, and Y. Pritch, "Omnistereo: Panoramic Stereo Imaging," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pp. 279–290, 2001.

[3] V. Vanijja and S. Horiguchi, "Omni-Directional Stereoscopic Images from One Omni-Directional Camera," *Jour. of VLSI Signal Processing Systems*, vol. 42, pp. 91–101, 2006.

[4] J. Becker, C. Stewart, and R. J. Radke, "LiDAR Inpainting from a Single Image," *Proc. of IEEE Int'l Workshop on 3-D Digital Imaging and Modeling*, pp. 1441–1448, 2009.

[5] A. Banno, T. Masuda, T. Oishi, and K. Ikeuchi, "Flying Laser Range Sensor for Large-Scale Site Modeling and Its Applications in Bayon Digital Archival Project," *Int'l Jour. of Computer Vision*, vol. 78, no. 2-3, pp. 207–222, 2008.

[6] J. I. Park and S. Inoue, "Image-Based Rendering from Multi-View Images," *jour. of the Institute of Image Information and Television Engineers*, vol. 52, no. 3, pp. 371–376, 1998.

[7] S. Knorr, M. Kunter, and T. Sikora, "Super-Resolution Stereo- and Multi-View Synthesis from Monocular Video Sequences," *Proc. of Int'l Conf. on 3-D Digital Imaging and Modeling*, pp. 55–64, 2007.

[8] M. Hori, M. Kanbara, and N. Yokoya, "Novel Stereoscopic View Generation by Image-Based Rendering Coordinated with Depth Information," *Proc. of Scandinavian Conf. on Image Analysis*, pp. 193–202, 2007.

[9] T. Sato, S. Ikeda, and N. Yokoya, "Extrinsic Camera Parameter Recovery from Multiple Image Sequences Captured by an Omni-directional Multi-camera System," *Proc. of European Conference on Computer Vision*, vol. 2, pp. 326–340, 2004.

[10] M. Levoy and P. Hanrahan, "Light Field Rendering," *Proc. of ACM SIGGRAPH'96*, pp. 31–42, 1996.