Improved Displaying System for HMD with Focusing on Gazing Point Using Photographed Panorama Light Field

Fuko Takano*, Takafumi Koike**

Graduate School of Computer and Information Sciences, HOSEI University
Kajino 3-7-2, Koganei, Tokyo 184-8584, Japan

*fuko.takano.2c@stu.hosei.ac.jp, **takafumi@hosei.ac.jp

Abstract

We improve a displaying system for HMD which displays a photographed image focused on user's gazing point. The refocused image is generated from a trimmed panorama light field image. When we watch the real world, we perceive a clear image at a gazing point while we perceive a blurry image outside the gazing point. Based on these our eye's features, when we watch photographed images on an as-made HMD, displayed images are focused on some fixed depth position. Thus, there is some inconsistency to our eyes, and many people feel a sense of incongruity. By showing an image focused on user's gazing point, we've improved immersiveness of photographed images on the HMD. We adopt a depth estimation method using phase shift and cost volume filtering to get a more accurate depth map. We implemented the system on a commercial HMD which can detect user's gazing point. Our goal is developing visually realistic sensation toward a plenoptic display.

Keywords: Light field, refocus image, head-mounted display

1. Introduction

In recent years, Virtual Reality (VR) has been used in various places such as games and education. VR is one of the computer technologies that provides user experiences that they feel like they are on the spot by presenting the virtual environment created on a computer to the user's hearing and vision using some devices such as displays and speakers. The Head Mount Display (HMD) which can provide users with images with wide viewing angles and close to the human vision can be said to be the most widely used methods for presenting VRs to the user with a visual sense. Furthermore,

in recent years, due to the inexpensive HMDs such as Oculus Rift, FOVE, PSVR and their contents corresponding to these HMDs, visual VR spread rapidly in the world.

When we watch the real world, we perceive a clear image at a gazing point while we perceive a blurry image outside the gazing point. When we make images for HMDs from photographed images, it is common to shoot an omnidirectional image and paste them as texture on the inside the spherical model in the virtual space to create omnidirectional VR. The image displayed in the virtual space doesn't fit the focus to user but is focused on some fixed depth. Since this view does not match the characteristics of the human eyes, this causes some inconsistency to our eyes, and many people feel a sense of incongruity.

Thus, we have developed the display system which displays images focused on gazing point by Oculus Rift and handmaid eye-tracking system⁽¹⁾. But the eye-tracking system was less accurate so we displayed an image focused on a center point of the HMD and made users gaze at the center of the display and ran a user study whether our system improved immersiveness of the HMD⁽²⁾. We displayed a fully CG scene at that time. We had a questionnaire about how users feel, and got a result that we can improve aggressiveness and immersiveness by using our system. After that, we shoot panorama light field at many points and made a walk-through system⁽³⁾.

Our key contributions are: 1) We implement our displaying system on FOVE which is one of the HMD which can examine the user's gazing point. 2) We adopt the depth estimation of the light field to select more correct refocused panorama images. We made a system which can display images which always focused on user's gazing point. We show the screenshot displayed on the HMD in the section 6.

Our goal is developing visually realistic sensation toward a plenoptic display⁽⁴⁾.

2. Related Work

Human eyes have the following characteristics: high eyesight at gazing point and Depth-of-Field (DoF) cues. There are some researches that aims to present images with less eye fatigue and uncomfortable feeling by displaying images matching these characteristics of eyes in Virtual Environment.

Fujita et al. have proposed foveated rendering to perform ray tracing by increasing the number of sampling points of gazing point against the feature that gazing points have high visual acuity⁽⁵⁾. The DoF blur effects in real-time applications of VR has been suggested by Rokita⁽⁶⁾. After that, Hillaire et al. have implemented gaze-contingent DoF blur rendering and performed the perceptual experiments to evaluate humans' impressions. It was possible to improve fun and game-play without lowering performance in first-person shooting game implementation⁽⁷⁾. They found that better immersion can be obtained by DoF rendering with eye-tracking in first-person navigation which users can move freely in a virtual environment⁽⁸⁾. As a result of Hillaire's experiment to study the subjective preference of users regarding the DoF blur effect, the scene with DoF blur effect computed using a focus zone centered on the user's focus point on the screen is preferred rather than the scene with DoF blur effect computed using a focus zone constantly located at the center of the screen.

However, these technologies are model-based implementations and it is difficult to bring these technologies as them for displaying photographed images. This paper presents that displaying photographed images on the off-the-shelf HMD with DoF characteristics.

3. Light Field

The light field is defined as a 4D function which describe all rays flying through space. It's intended to reconstruct all rays on the virtual fields. Adelson et al. have introduced a Plenoptic function⁽¹⁰⁾ which can express the rays using 5-parameters; 3D arbitrary view point and 2D every possible angle. Levoy et al. have presented 4-parameters representation by reducing one parameter of the plenoptic function since radiance does not change along a line unless blocked⁽¹¹⁾. This 4D parameterization is called light field. To represent 4D light field coordinates, 2 plane

Fig. 1. Two plane parameterization of light field

Fig. 2. Overall flow of our displaying system

parameterization (2PP) is commonly used such as figure 1. A camera which records information acquired by the 2D image sensor as a 4D light field is called a light field camera. We can make refocused image and estimate depth and perspective changed image by rendering a light field image taken by the light field camera.

4. Our Approach

Figure 2 shows the process flowchart of our system. In the preprocessing, we calibrate the light field camera, shoot some light field images, estimate the depth map from light field images, synthesize the light field images, and render the panorama light field images. After these processes, we display refocused images depending on user's gazing point.

4.1 Light Field Camera Calibration

Our system decodes the light field images and calibrates the light field camera with Dansereau's method⁽¹²⁾. The inputs are white images of light field camera and checker board images taken by light field camera. The output is the calibration data of the light field camera.

4.2 Taking Light Field Photographs

Our system makes panorama light field images with Birklbauer's method⁽¹³⁾. We capture light field images by rotating horizontally the light field camera on the camera

Fig. 3. Panorama images mapped on cylinder used in this system. **Top:** panorama light field image **Middle:** panorama depth image using Jeon et al.'s depth estimation method⁽¹³⁾ **Bottom:** panorama depth image using Previous method⁽²⁾

platform with trying not to change the camera's zoom.

4.3 Synthesizing light field images and generating refocused panoramic images

Our system synthesizes the taken light field images with Birklbauer's method⁽¹³⁾ using angle data and RMS error. The synthesized light field images are the panorama light field. Our system uses Birklbauer's rendering method⁽¹⁴⁾ to render the panorama light field images and generate some panoramic images with focus on different depth point.

4.4 Estimating depth and synthesizing the depth map

We make the panorama depth map by synthesizing depth maps each Lytro image. We use a panorama depth map to get the distance from the virtual camera to an arbitrary object in the panorama image. In our previous research, we used a low-resolution depth map and it was not accurate. So we adopt the depth estimation method of Jeon et al.⁽¹⁵⁾ to each light field image. This method is an accurate depth estimation method using phase shift and gradient costs and feature correspondence. After that, we synthesize the obtained depth images to create panorama depth map.

4.5 Displaying an image focused on the user's gazing point

Our system maps the created panorama images and

Fig. 4. Left: A light field raw image taken by Lytro Right: enlarged part of light field raw image

Fig. 5. An infrared image taken by FOVE eye tracking sensor

panorama depth map on the cylindrical model as shown in Figure 3 and places the virtual camera, which is the viewpoint of the user, at the center of the cylindrical model. The panorama image which displayed on HMD is selected from various refocused panorama images by a depth value at the gazing point of the panorama depth map. The intensity of the depth map represents a distance between the eye position and a target object. When we experience the system, the system shows a photographed panorama scene assigned to a cylindrical model.

5. Implementation

We use a Lytro first generation to take light field images. Lytro has 331×382 hexagonal lens-array in front of a 3280×3280 CMOS sensor. An image taken by Lytro and its enlarged image is shown in the Figure 4. It shows that one small circle corresponds to each lens. We can get 9×9 multi-view images with a resolution of 662×662 by upsampling this light field with Birklbauer's method⁽¹³⁾.

FOVE HMD from FOVE, Inc. is one of the HMD which has stereo infrared eye tracking sensors. This display has a resolution of 2560×1440 pixels, a field of view of 100° , and the gaze tracker sensors' frame rate of 120 fps. Figure 5 presents an infrared image taken by FOVE eye tracking sensor when the user is looking at the front.

Fig. 6. A CG-based video focused at the gazing point. **Top**: focused on the back-wall **Bottom**: focused on the blue sofa in the front. A gazing point is green point.

6. Results

We show the displayed images on the HMD which focuses on the user's gazing point. Figure 6 presents the reproduced CG scene used in Hillaire's experiment which is rendered with DoF blur effect computed using a focus zone centered on the user's gazing point and displayed it on FOVE. The left images of figure 6 shows the displayed image for the left eye when this scene is displayed on the HMD and the right of them are expanded image of the left image. A green circle in each figure indicates the user's gazing point. When the user's gazing point on the checkerboard wallpaper on the back-wall, the blue sofa in the front is blurred while on the other hand, when the user's gazing point in on the blue sofa in the front, the checkerboard wallpaper on the back-wall is blurred.

Figure 7 presents the image-based scene with DoF blur effect which we described in this paper. Same as figure 6, the left image in figure 7 shows the displayed image for the left eye when this scene is displayed on the HMD and the right of them are expanded image of the left image. A green circle in each figure indicates the user's gazing point. When the user's gazing point on the checkerboard wallpaper on the back-wall, the pumpkin ornament in the front is blurred while on the other hand, when the user's gazing point in on the pumpkin ornament in the front, the checkerboard

Fig. 7. A photographed image-based video focused at the gazing point. **Top**: focused at the back-wall **Bottom**: focused at the front pumpkin ornament. A gazing point is green point.

wallpaper on the back-wall is blurred.

7. Discussions

The current system is far from our final goal; we think this work as the first step rather than the last word in gaze-contingent DOF blur for photographed VR. Our problems we have are divided into three types; the low angular and spatial resolution, narrow angular of view, and insufficient accuracy of eye-tracking.

We can improve the resolution of refocused image which is displayed on the HMD by using the light field image which has high spatial resolution. Further, we can reduce ghost and expand the aperture by improving the angular resolution. However, it is difficult to enlarge the sensors of the light field cameras easily. Therefore, it is effective that the other light field stitching method. Guo et al. shows the light field stitching method using 5×6 matrix⁽¹⁶⁾.

Improved stitching method for light field is also expand the angle of view. The omnidirectional light field image causes photographed virtual environment to improve its immersiveness since the scene is displayed in the entire field of view and it will be a step toward displaying more immersive images.

Eye tracking accuracy of FOVE is low to get the gazing point yet. Thus, our system displays a blurred image on the user's gazing point and clear image on another point. It

causes discomfort and VR sickness in users. The semantic weighting implemented by Hillaire et al. is a one way to reduce the effect of the eye tracker⁽⁷⁾.

8. Conclusions

We have described the improvement of the displaying system for the HMD which matches to human eyes' characteristics to display the more immersive VR. We improve our previous displaying system to display on FOVE and adopt the accurate depth estimation method. As a result, our system displays the virtual environment made from light field image with DOF blur effect computed using a focus zone centered on the user's gazing point on FOVE.

However, our system is not enough to display the very immersive image-based virtual environment. We will further improve our system by solving the problems; low-resolution, narrow angle of view, and insufficient accuracy of eye-tracking.

References

- (1) Fuko Takano and Takafumi Koike : "Photographed Image Presenting HMD System with Focusing on Gazing Point Using Panorama Light Field and Eye Tracking", Conference of the Virtual Reality Society of Japan, 2015. (in Japanese)
- (2) Fuko Takano and Takafumi Koike : "Image Presenting System for HMD with Focusing on Gazing Point Using Photographed Panorama Light Field", IMID, 2016
- (3) Fuko Takano and Takafumi Koike : "PlenoGap: Panorama Light Field Viewing for HMD with Focusing on Gazing Point", ACM SIGGRAPH 2016 Posters. ACM, 2016
- (4) Takafumi Koike : "Future of 4D Light Field Display", IMID 2015
- (5) Masahiro Fujita and Takahiro Harada : "Foveated real-time ray tracing for virtual reality headset", Tech. rep., Light Transport Entertainment Research, 2014
- (6) Przemyslaw Rokita : "Generating depth of-field effects in virtual reality applications", IEEE Computer Graphics and Applications, vol. 16, no. 2, pp. 18-21, 1996
- (7) Sébastien Hillaire, Anatole Lécuyer, Rémi Cozot, and Gry Casiez : "Depth-of-field blur effects for first-person navigation in virtual environments", Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology, pp. 203-206, 2007
- (8) Sbastien Hillaire, Anatole Lcuyer, Rmi Cozot, and Gry Casiez : "Using an eye-tracking system to improve camera motions and depth-of-field blur effects in virtual environments", IEEE Virtual Reality Conference, pp. 47-50, 2008
- (9) Douglas Lanman and David Luebke : "Near-eye light field displays", ACM Transactions on Graphics, 2013
- (10) Mark Levoy and Pat Hanrahan : "Light field rendering", Proceedings of the 23rd annual conference on Computer graphics and interactive techniques. ACM, 1996
- (11) Edward H. Adelson and James R. Bergen : "The plenoptic function and the elements of early vision", 1991
- (12) Donald G. Dansereau, Oscar Pizarro and Stefan B. Williams : "Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1027-1034, 2013
- (13) Clemens Birklbauer and Oliver Bimber : "Panorama light-field imaging", Computer Graphics Forum. Vol. 33, No. 2, 2014
- (14) Clemens Birklbauer, Simon Opelt, and Oliver Bimber : "Rendering gigaray light fields" Computer Graphics Forum, vol. 32, Wiley Online Library, 469-478, 2013
- (15) Hae-Gon Jeon, Jaesik Park, Gyeongmin Choe, Jinsun Park, Yunsu Bok, Yu-Wing Tai and In So Kweon : "Accurate depth map estimation from a lenslet light field camera." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015
- (16) Xinqing Guo, Zhan Yu, Sing Bing Kang, Haiting LIN, and Jingyi Yu: "Enhancing Light Fields through Ray-Space Stitching." Transactions on Visualization and Computer Graphics (TVCG), 2015