A Gaze-Reactive Display for Simulating Depth-of-Field of Eyes When Viewing Scenes with Multiple Depths

Tatsuro ORIKASA∗†, Nonmember and Takayuki OKATANI∗a, Member

SUMMARY The the depth-of-field limitation of our eyes causes out-of-focus blur in the retinal images. The blur dynamically changes whenever we change our gaze and accordingly the scene point we are looking at changes its depth. This paper proposes an image display that reproduces retinal out-of-focus blur by using a stereoscopic display and eye trackers. Its purpose is to provide the viewer with more realistic visual experiences than conventional (stereoscopic) displays. Unlike previous similar systems that track only one of the viewer’s eyes to estimate the gaze depth, the proposed system tracks both eyes individually using two eye trackers and estimates the gaze depth from the convergence angle calculated by triangulation. This provides several advantages over existing schemes, such as being able to deal with scenes having multiple depths. We describe detailed implementations of the proposed system and show the results of an experiment conducted to examine its effectiveness. In the experiment, creating a scene having two depths using two LCD displays together with a half mirror, we examined how difficult it is for viewers to distinguish between the real scene and its virtual reproduction created by the proposed display system. The results of the experiment show the effectiveness of the proposed approach.

key words: depth of field, eye tracker, stereoscopic display

1. Introduction

The limitation in the depth-of-field (DOF) of the human eyes causes out-of-focus blur in the retinal images. This type of blur, which we will call DOF blur in what follows, plays a number of roles in the human visual system. It is widely accepted that DOF blur is one of the pictorial depth cues [1]–[5]. It is also reported that the accommodation of the eyes is driven by DOF blur in retinal images [6], [7], and its motor control and sensory feedback could give depth sensation [8]. Blum et al. reported that visual discomfort of viewing stereoscopic displays can be mitigated by adding DOF blur [9]. Their study explains that the higher viewing comfort is considered to be mainly because adding DOF blur increases the fusional limits, i.e., the left and right images can be fused without double vision.

Considering these roles, we conjecture that an image display that can provide a viewer with the same visual experience as when viewing natural scenes should be able to precisely reproduce DOF blur on the viewer’s retinal images. In theory, this can be done if the light field of a scene can be reproduced so accurately that the viewer’s eyes can accommodate to focus. However, there is no such display technology that can do this for a sufficiently wide range of scene depths while maintaining sufficient image resolution.

A more practical approach would be to present DOF blur to a viewer by a conventional image display in such a way that uses two-dimensional image processing to replicate the blur effect. To simulate the dynamic nature of the DOF blur (namely, it will vary dynamically in response to the viewer’s gaze change), the viewer’s gaze is measured by an eye tracker. Then, the image with appropriate DOF blur is generated based on the measured gaze and shown on the display. By performing these steps in real time, it will be possible to provide the viewer with realistic visual experiences. To the best of our knowledge, [10] is the first study to propose this concept, and [11] is the first to develop a working system and reports experiments, where the effects of displaying DOF blur are investigated using a computer game in which a player moves in a virtual three-dimensional space. While their system employs a two-dimensional image display, [9] presents a system that uses a stereoscopic display; they perform experiments to verify the hypothesis that DOF blur reduces visual discomfort when viewing stereoscopic displays [12].

A key ingredient in this approach (of measuring the viewer’s gaze and controlling DOF blur displayed on a conventional image display in a reactive manner) is how to estimate the gaze depth of the viewer, i.e., the depth at which the viewer is looking. The above previous studies employ a simple method, which is to first assume a simplified scene model that the scene depth is a single-valued function of the image coordinates, and then regard the depth of the scene point at which the viewer is looking as the depth at which the viewer wants to look.

While it greatly simplifies estimation of the gaze depth, there are a number of problems in this simple approach. Firstly, it cannot deal with scenes having multiple depths, as shown in Fig. 1. Secondly, it is based on an overly simplified model of accommodation of the eyes, in which the eyes are assumed to always focus on the depth of the currently gazing point and their focus changes only in a passive manner, triggered by a gaze change. This is obviously inaccurate, which could make the visual experience unnatural. Thirdly, even for scenes having single-valued depths, it is technically difficult to choose the right depth particularly around occluding boundaries of objects in the displayed scene. As the scene depth is discontinuous across the occluding boundaries, a slight error of the gaze measurement will result in a wrong
estimation that differs from the viewer’s intention. This will be a serious limitation, as it is shown [5] that the DOF blur around occluding boundaries play an essential role in depth sensation.

In this paper, to resolve or alleviate these problems, we propose to measure the convergence of the viewer’s eyes and estimate gaze depth by triangulation. We show a working system that performs this to estimate the gaze depth and displays images with appropriate DOF blur based on the estimation. Our system consists of a stereoscopic display† and two eye trackers. The eye trackers measure independently the left and right eyes of the viewer, from which the convergence angle is calculated. In our system, we estimate the viewer’s gaze depth from not only the convergence angle but also knowledge of scene depths. This is based on the finding in our preliminary experiments that it is impossible to estimate the depth of the viewer’s intention solely from the measured convergence angle. The details will be described later.

We also show the results of an experiment conducted to investigate the effectiveness of the proposed system. The goal of our system is to provide the viewer with as realistic visual experiences as possible. The reality of the visual experiences can be evaluated by how difficult it is to distinguish a real scene and its virtual reproduction by the proposed system. Thus, in the experiment, we conducted a test of measuring this difficulty. There are a number of potential factors that make the viewer’s experiences provided by the system unrealistic or unnatural. The major one is depth cue inconsistency; accommodation is still incorrect in the proposed system, as in other image displays. As the proposed system provides DOF blur in addition to binocular disparity and convergence††, there could emerge a new, unknown visual discomfort when using the proposed system, in addition to the known ones with conventional stereoscopic displays.

In what follows, we first describe the detailed implementations of the proposed system. These are given in Sect. 2, which includes the implementation of the eye trackers, the algorithm for estimating the viewer’s gaze depth, and the method for synthesizing images with appropriate DOF blur given the depth to be focused on. We then show the experimental results in Sect. 3. Section 4 concludes this paper.

2. Proposed System

2.1 Hardware

Our system consists of a stereoscopic display with glasses and two eye trackers; see Fig. 2. For a stereoscopic display (shown as Display 1 in the figure), we use DELL Alienware OptX AW2310 along with nVidia 3D Vision 2 (stereoscopic glasses with active shutters and an IR emitter). This pair displays the image of 1920×1024 pixels in the viewable area of 509.8×286.7mm at refresh rate 60Hz for each of left and right image channels. The display is connected via two DVI cables to a graphics board on a PC; the board is equipped

†Note that estimating the gaze depth from the convergence angle necessitates the use of a stereoscopic display for displaying images.

††Motion parallax is not reproduced, as our system requires the viewer to fix its head.
with a GPU (Graphics Processing Unit; nVidia GTX 480), which is used for the creation of images with synthetic DOF blurs as well as for tracking the viewer’s eyes.

2.2 Eye Tracking

We have developed an eye tracking system from scratch to ensure accuracy and realtime performance. It consists of two high-speed cameras (DITECT HAS-L1), a PC to which the cameras are connected via Camera Link cables, and an infrared light source. The PC is the same as the one to which the stereoscopic display is connected. The two cameras capture the images of the viewer’s left and right eyes, respectively, through the shutters of the stereoscopic glasses. They capture the images of 640 × 480 pixels at frame rate 500Hz. We put a visible ray cut off filter (cut off wavelength 740nm) in front of each camera. We have confirmed that the shutters of the glasses do not interfere with the captured images, irrespective of opening/closing of the shutter. This is because the LCD shutter used in the glasses blocks only visible light and is transparent to infrared light. During the experiments, the viewer’s head is fixed on a chin rest; its forehead and jaw are lightly pressed to two plates attached to the rest so that the head will not move.

The algorithm of the eye tracker is based on the standard corneal reflex method. In every acquired image of each eye, the Purkinje reflection of the infrared source is extracted as a small pixel area by simple thresholding of image brightness, and the boundary of the pupil is extracted by fitting an ellipse to it. For the latter, we use a RANSAC-based method to increase robustness [13], in which hypotheses (i.e., fitted ellipses) are created for a set of randomly chosen five edge points around the pupil, and then they are verified using all the edge points. The generation and verification of each sample is efficiently performed in a parallel manner on the GPU, which makes the overall processing time of a single frame less than 2ms.

Assuming that the viewer’s head does not move, the orientation of the eye is estimated from the center of the fitted ellipse and that of the Purkinje reflection. To do this, calibration is performed for each viewer before the use of the system, in which a stationary cross mark is shown on the stereoscopic display and the viewer is asked to look at it; the mark is shown in turn at twelve different positions. The mapping from the positions of the mark and the centers of the ellipse and the Purkinje reflection are learned. It is then used for the estimation of the viewer’s gaze.

2.3 Estimation of Gaze Depth

The two eye trackers yield the positions \((x_l, y_l)\) and \((x_r, y_r)\) of the left and right gaze points on the display surface, respectively. The gaze depth of the viewer is estimated from these measurements in the following way.

Let \(p\) denote the distance between (the centers of) the left and right eye balls and \(d\) be the distance from them to the display surface. (The positions of the eyes are assumed to be parallel to the display surface.) Assuming the lines of sight for left and right eyes to cross at a single point in space, the distance \(x_l - x_r\) between the left and right gaze points can be represented using the gaze depth \(D\) of the viewer as

\[
x_l - x_r = \frac{p(D - d)}{D}.
\]

This equation enables determining \(D\) directly from \(x_l - x_r\). Then it is theoretically possible to use \(D\) thus determined for synthesizing DOF blur and presenting it to the viewer. However, our preliminary experiments revealed that this approach did not work; the viewers reported that they could not control the gaze depth as they intended. We conjecture that this difficulty with the gaze control could probably be because of slight inaccuracy of the eye tracking and/or small time lag from measuring gaze to redrawing images. These could probably have made unstable the closed-loop system consisting of the viewer and the proposed system. Verifying this conjecture will be a future work.

Thus, we determine \(D\) by additionally using the depth maps of the target scene. There is one depth map for each eye, and thus we have the scene depths \(D_l(x_l, y_l)\) at \((x_l, y_l)\) and \(D_r(x_r, y_r)\) at \((x_r, y_r)\) for the left and right eyes, respectively. Both \(D_l(x_l, y_l)\) and \(D_r(x_r, y_r)\) could be multi-valued, and we denote the set of all the depths by \(\mathcal{D} \equiv \{D_1, \ldots, D_n\}\). Here, \(n\) is the number of candidate depths, which is determined by the depth maps; for example, when \(D_l(x_l, y_l)\) and
$D_i(x_i, y_i)$ are both $m$-valued, then $n = 2m$. When the viewer gazes the depth $D_i$, the distance between the left and right gaze point will be given from Eq. (1) as $p(\hat{D}_i - d)/\hat{D}_i$. We choose from $\mathcal{D}$ the depth $\hat{D}_i$ that has the closest distance to the measured one $x_i$ as

$$\hat{D} = \min_{D_i} \left| \frac{p(\hat{D}_i - d)}{\hat{D}_i} - (x_i - x_i) \right|. \quad (2)$$

We use $\hat{D}$ thus determined for synthesizing the image with appropriate DOF blur. The depth maps are created as follows. When scene images are synthesized by CG, the depth maps are simply copied from the $z$-buffers of the scene 3D model. In the experiments shown in the next section, we measured the distances to the displays and used them to create images to be displayed as well as depth maps.

### 2.4 Image Synthesis

Given the gaze depth, we wish to display the image of a target scene with the appropriate DOF blur. Ideally, the image should be such that the retinal image of the viewer who sees the image shown on the display coincides with its real retinal image when seeing the scene. Because of the difficulties of modeling the optics involved, we approximate it with the image that will be captured for the scene by an ideal imaging system consisting of a thin lens and an image plane, as in the previous studies. Then, several methods can be used to create the images to be displayed. They can be captured for a real scene using some imaging device such as light field cameras. Alternatively, they can be synthesized by rendering from scratch or generating only DOF blur to a pan focus image of a scene with its depth map. In our implementation, we chose the latter because of its ease and sufficiency for the purpose of experimentally validating the proposed approach.

In general, DOF blur of a scene, which varies at each imaged point depending on the depth of the corresponding scene point, is essentially characterized by its size called the circle of confusion (CoC). As mentioned above, we assume an imaging system of a thin lens. Let $d$ be the aperture and $f$ be the focal length of the lens. Letting $\hat{D}$ be the current gaze depth, the distance between the lens and the image plane is given by the Gauss rule as $v_0(= f\hat{D} / (\hat{D} - f))$, and then the radius $c$ of CoC for a scene point having depth $u$ from the lens can be calculated as

$$c = \frac{(f - v_0)u + f v_0}{fu} d. \quad (3)$$

There are several methods for generating DOF blur in the image of a scene [14], [15]. Considering real-time performance, we employ the gathering approach [16] that is computationally efficient and also suitable for parallel computation. It generates blur at a pixel by gathering and summing the values of its nearby pixels in the inverse order from the real physical process of DOF blur generation. To make it possible to deal with multiple depths, we generate DOF blur independently for each layer and then calculate the sum of their brightnesses. Examples of the images generated for scenes with two depths are shown in Fig. 1.

When the human eyes accommodate to focus, their focal length (or its refractive power) changes according to their physical mechanics, dynamics, and neural control system [17]. Their temporal transition typically shows a smooth S-shaped curve. To simulate this, we calculate the following temporal mean over the latest $q$ estimates $[\hat{D}_T, ..., \hat{D}_{T-q+1}]$ that are obtained by the method described above as

$$\hat{D} = \frac{1}{q} \sum_{t=T-q+1}^{T} \hat{D}_t, \quad (4)$$

and use this as the input to Eq. (3). When the gaze depth changes from a depth to a different one, $\hat{D}$ varies continuously between the two depth values, and thus the resulting DOF blur changes smoothly. In the experiments, we set $q = 30$ (estimates, equivalently frames); the cameras in our system capture 500 frames per second, and thus $q = 30$ frames correspond to about 0.06 second.

### 3. Experiments

We conducted an experiment to examine the effectiveness of the proposed system. The goal of the proposed system is to provide the viewer with as realistic visual experiences as possible. Although it is in general difficult to evaluate such reality, we may consider it to be a success if the virtual reproduction of a real scene by the proposed system is indistinguishable from the real one. The details are shown below.

#### 3.1 Experimental Design

Besides the original display in the system, we use an additional two-dimensional (i.e., non-stereoscopic) display to create a real scene with two depths. We call the original one Display 1 and the added display Display 2. We insert a half mirror in between Display 1 and the viewer so that the two images displayed on Display 1 and 2 are seen overlaid from the viewer. The overview of the system and its schematic layout are shown in Figs. 2 and 3, respectively. The whole system is placed on a desk, and is lit by ceiling light; the illuminance of the desk is about 400lx.

We then consider three methods of displaying images using this setup. The first is the method that shows images on both Display 1 and 2 and uses the half mirror to physically make the two images overlaid when seen from the viewer; see the top row of Fig. 4. We call this method **Method 1**. Display 1 is stereoscopic but its left and right images are set to be identical so that the displayed image appear to lie on the surface of Display 1. In short, the near one and the far one of the overlaid images to be at the depth of Display 1 and 2, respectively.
Fig. 3  Dimensions of the experimental system.

Fig. 4  Three methods of displaying a scene with two depths. Method 1: The near image is shown on Display 1 and the far image is on Display 2. No synthetic DOF blur is added to the displayed images. (Natural DOF blur should emerge in the viewer’s retinal images.) Method 2: The near and far images are both shown on Display 1 with appropriate disparities so that the near one will be seen as if it is on Display 1 and the far one be seen as if it were on Display 2. No synthetic DOF blur is added to the displayed images. Method 3: The near and far images are both shown on Display 1; they have appropriate disparities as in Method 2 and additionally have appropriate DOF blur generated based on the measured viewer’s gaze. (The proposed system is in operation in this method.) The right two columns show the images captured by a camera from a typical position of viewers’ eyes to reproduce the viewers’ sights when focusing on the near and far images, respectively.

The second and third are the methods that simulate the scene (i.e., the two overlaid images displayed by Method 1) by using only Display 1. In both methods, the near and the far images are displayed by providing Display 1 with appropriate left and right stereoscopic images, which are such that the near and the far images appear to be at the depths of Display 1 and 2, respectively. The two (i.e., the second and the third) methods differ in whether or not DOF blur is synthesized by the proposed system. We call the one without DOF blur Method 2 and the one with DOF blur Method 3; see the middle and bottom rows of Fig. 4. Thus, Method 2 simply uses Display 1 as an ordinary stereoscopic display and presents both the near and far images as they are; it simulates Method 1 only in a geometric sense. Method 3
synthesizes the DOF blur based on the measurement of the viewer’s gaze depth by using the proposed system; it simulates Method 1 with respect to geometry as well as DOF blur. Its remaining difference from Method 1 is the absence of accommodation of the eyes; in the case of Method 1, the viewer’s eyes do accommodate to focus on between the surfaces of Display 1 and 2, whereas they do not in the case of Method 3.

Table 1 shows the summary of Methods 1, 2, and 3. Method 1 provides a real scene, whereas Method 2 and 3 give its virtual reproduction without and with DOF blur. We want to evaluate how difficult it is to distinguish Method 1 and 3. For the sake of comparison, we also evaluate how difficult it is to distinguish Method 1 and 2. Distinguishing Method 1 vs. 3 as well as Method 1 vs. 2 is equivalent to detect on which display the far image is being displayed. Therefore, we ask a viewer (a subject) the latter query in the experiments. The images shown in Fig. 5 are displayed as the near and far images. The near image is fixed to the one highlighted by a dotted line and the far image is chosen from all these images.

### Table 1

| Methods | Near image | Far image | DOF blur |
|---------|------------|-----------|----------|
| 1       | Display 1  | Display 2 | Real     |
| 2       | Display 1  | Display 1 | None     |
| 3       | Display 1  | Display 1 | Simulated|

3.2 Procedure

Eight subjects participated in the experiments; all males, aged 21 to 40. We obtained informed consent from all the subjects. This research was approved by the Ethics Committee for Human Research of Graduate School of Information Sciences, Tohoku University.

For each subject, we first instruct the configuration of the system. Notifying the existence of two displays, Display 1 and 2, we tell the subject that two images will be shown in two different visual depths, and the near one will be shown on Display 1 and the far one will be shown either on Display 1 or 2. We then tell the subject that we will ask him/her to detect on which display the far image is being shown.

After the instruction, we perform two tests. The first test is to distinguish between Method 1 and 2, and the second one is to distinguish between Method 1 and 3. These two tests are performed sequentially in this order with five minutes interval. Each test consists of a training phase and a test phase. In both phases, a fixed number of queries are in turn given to the subject, in each of which the displaying method is randomly chosen between the two, while keeping the counts of each method to 50% of the total queries. An example sequence of sixteen queries in the second test is

\[ q = [1, 3, 3, 1, 3, 3, 1, 1, 1, 1, 1, 1, 3, 3, 1, 1, 3], \]

where each number (1 or 3) indicates which Method (1 or 3) is used for the query. For each query, the subject is asked to answer within ten seconds on which display the far image is displayed. In the training phase, the subject is told the correct answer right after answering each query, which provides a chance of learning the difference between the two methods, if it can be learned. The subject is not told the answers in the test phase. The training and test phases are sequentially performed in this order with two minutes interval. Ten queries are given in the training phases and sixteen queries are given in the test phases. As mentioned above, the subjects are asked to answer on which between Display 1 and 2 the far image is being displayed. An example of subject’s answers to the above query sequence is

\[ a = [2, 1, 1, 1, 1, 2, 1, 2, 1, 1, 1, 1, 2, 1, 1, 1]. \]

In Method 1, the far image is displayed on Display 2, whereas in Method 3, it is displayed on Display 1. Thus, for \( i \)-th query, the subject’s answer is correct if either \( q_i = 1 \) and \( a_i = 2 \) or \( q_i = 3 \) and \( a_i = 1 \) and is incorrect otherwise. The number of correct answers in \( a \) for \( q \) in the above example is 12. The results from the test phases are analyzed below.

3.3 Results

Figure 6 shows the rate of correct answers for each of the two tests, namely, those of distinguishing Method 1 vs. 2 (left blue bars) and 1 vs. 3 (right green bars). It shows the averaged number over the eight subjects, which are indicated by A–H. The total number of correct answers over the eight subjects is down from 113 for Method 1 vs. 2 to 87 for Method 1 vs. 3, both out of 8 \( \times \) 16 = 128 queries. The \( p \)-value of the null hypothesis that it is more difficult to distinguish Method 1 vs. 2 than 1 vs. 3 is \( 4 \times 10^{-5} \). Thus, we can...
the dominant cause for the imperfect results is errors in eye tracking. In fact, some of the subjects report after the experiment that they sometimes experienced that the DOF blur shown in the displayed image “automatically” varied independently of their intention, which is mainly attributable to the errors of eye tracking. Eye tracking sometimes does fail because of the interference of eyelids and eyelashes; occasional eyeblinks can also make it fail though it is only in a short time. It is also observed that the frequency of the occurrence of the tracking errors differed considerably for each subject. This could explain the individual differences in the effectiveness seen in the results.

4. Conclusion

In this paper, we have described a stereoscopic image display that, together with two eye trackers, reproduces DOF blur. The mechanical difference from previous systems is that our system tracks the both eyes independently, calculates the convergence angle, and estimates the gaze depth.

We have also shown the results of the experiment conducted to evaluate the effectiveness of the proposed system. The experiment aims at assessing how natural the visual experience provided by the system is. Choosing a scene with two depths, we design a task of distinguishing a real scene physically having two depths from its virtual reproduction as created by the proposed system. Multiple subjects participated in the experiment, in which each subject solves the task for a number of times. The results show that reproducing DOF blur with our system did make it more difficult to distinguish the real scene and its virtual reproduction, which indicates the effectiveness of the proposed approach. It is noted, however, that the level of the difficulty was below the ideal one, that is, the virtual reproduction is not perfectly indistinguishable from the real scene. There are two possible reasons for this imperfect result. One is the errors of the eye tracking, and the other is the inconsistency among the depths cues. We conjecture that the former is a dominant factor in explaining the results. Further analysis will be the future work.

Acknowledgments

This work was supported by KAKENHI 23135501.

References

[1] A.P. Pentland, “A new sense for depth of field,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.PAMI-9, no.4, pp.523–531, 1987.
[2] G. Mather, “Image blur as a pictorial depth cue,” Proceedings of Biological Sciences, vol.263, no.1367, pp.169–172, 1996.
[3] G. Mather, “The use of image blur as a depth cue,” Perception, vol.26, no.9, pp.1147–1158, 1997.
[4] G. Mather and D.R.R. Smith, “Depth cue integration: stereopsis and image blur,” Vision Research, vol.40, no.25, pp.3501–3506, 2000.
[5] V.A. Nguyen, I.P. Howard, and R.S. Allison, “Detection of the depth order of defocused images,” Vision Research, vol.45, no.8, pp.1003–1011, 2005.
[6] K.J. Ciuffreda, “Accommodation and its anomalies,” in Visual Optics and Instrumentation, ed. W.N. Charman, pp.231–279, Macmillan Press, London, UK, 1991.

[7] K.J. Ciuffreda, “Accommodation, the pupil and presbyopia,” in Borish’s Clinical Refraction, ed. W. Benjamin, pp.77–120, W.B. Saunders Co., Philadelphia, PA, 1998.

[8] K. Fisher and K.J. Ciuffreda, “Accommodation and apparent distance,” Perception, vol.17, no.5, pp.609–621, 1988.

[9] T. Blum, M. Wieczorek, A. Aichert, R. Tibrewal, and N. Navab, “The effect of out-of-focus blur on visual discomfort when using stereo displays,” Proceedings of International Symposium on Mixed and Augmented Reality, pp.13–17, 2010.

[10] P. Rokita, “Generating depth-of-field effects in virtual reality applications,” Proceedings of IEEE Computer Graphics and Applications, vol.16, no.2, pp.18–21, 1996.

[11] S. Hillaire, A. Lécuyer, R. Cozot, and G. Casiez, “Using an eye-tracking system to improve depth-of-field blur effects and camera motions in virtual environments,” In Proceedings of IEEE Virtual Reality, pp.47–50, March 2008.

[12] M. Lambooij, M. Fortuin, I. Heynderickx, and W. IJsselsteijn, “Visual discomfort and visual fatigue of stereoscopic displays: a review,” Journal of Imaging Technology and Science, vol.53, no.3, pp.30201-1–30201-4, 2009.

[13] R.I. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, second ed., Cambridge University Press, ISBN: 0521540518, 2004.

[14] R.L. Cook, T. Porter, and L. Carpenter, “Distributed ray tracing,” Proceedings of SIGGRAPH, vol.18, no.3, pp.137–145, 1984.

[15] P. Haeberli and K. Akeley, “The accumulation buffer: Hardware support for high-quality rendering,” Computer Graphics, vol.24, no.4, pp.309–318, 1990.

[16] E. Hammon, GPU Gems 3, ch. 28 Practical Post-Process Depth of Field, Addison-Wesley Professional, 2007.

[17] C.M. Schor, L.A. Lott, D. Pope, and A.D. Graham, “Saccades reduce latency and increase velocity of ocular accommodation,” Vision Research, vol.39, no.22, pp.3769–3795, 1999.

Tatsuro Orikasa received his M.S. degree from Graduate School of Information Sciences, Tohoku University in 2012. He joined Koei Tecmo Games Co., Ltd. in 2012. He received MIRU Frontier Award in 2010.

Takayuki Okatani He is a Professor at the Graduate School of Information Sciences of Tohoku University, where he leads the Image Analysis Lab. He received his doctorate degree in Engineering from the University of Tokyo in 1999. He then worked as a Research Associate at Tohoku University and become a Professor since 2013. His areas of research are computer vision and image processing. He currently serves as an associate Editor of International Journal of Computer Vision.