Reducing Motion Blur Artifact of Foveal Projection for Dynamic Focus-Plus-Context Display

Daisuke Iwai, Kei Kodama, and Kosuke Sato, *Member, IEEE*

***Abstract*—This paper presents a novel technique to reduce the motion blur artifacts of foveal projection in a dynamic focus-plus-context (DF+C) display. The DF+C display is generally configured with multiple projectors and provides a non-uniform spatial resolution that consists of high-resolution regions (foveal projection) and low-resolution regions (peripheral projection). A serious problem of the DF+C display is motion blur, which inevitably occurs when a foveal projection is moved by a pan-tilt mirror or gantry. We propose a solution that reduces the motion blur artifacts, and evaluate how this solution improves the image quality by using both qualitative and quantitative experiments. Our proposed method defines an error function to assess the displayed image quality as the difference between an original high-resolution image and the displayed image by taking the non-uniform spatial property of human visual acuity into account. Then, it decides the set of positions and moving techniques of foveal projections so that the sum of errors during a video sequence is minimized. Through experiments, we confirmed that the proposed method can provide a better image quality and significantly improve the motion blur artifacts when compared to a conventional DF+C display.**

***Index Terms*—Focus-plus-context display, projection display.**

I. Introduction

THE number of pixels used to make a digital image is constantly increasing. For example, a photography of more than one gigapixel can be obtained either by capturing multiple photographs and stitching them together [1] or by using a single image detector [2]. However, there is no single display that can show such gigapixel images without downsampling them. The number of pixels used in next-generation display formats, such as 4K ultra-high definition (8 megapixels) or Super Hi-Vision (33 megapixels), is still much less than that of gigapixel digital imagery. A potential solution is a tiled display approach, which connects multiple displays both vertically and horizontally, and regards them as a single display [3]. However, such systems require a huge number of display devices and consequently become unreasonably expensive.

Manuscript received February 16, 2014; revised July 3, 2014; accepted August 21, 2014. Date of publication XXXXXXX XX, 201X; date of current version XXXXXXX XX, 201X. This work was supported by a Grant-in-Aid for Scientific Research on Innovative Areas “Shitsukan” (No. 22135003) from MEXT, Japan. This paper was recommended by Associate Editor A. Kaup.

D. Iwai, K. Kodama and K. Sato are with Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, JAPAN, email: (see http://www.sens.sys.es.osaka-u.ac.jp/).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier XX.XXXX/TCSVT.201X.XXXXXXX

Another solution is to use a focus-plus-context (F+C) display, which offers a non-uniform spatial resolution consisting of high resolution (hi-res) and low resolution (low-res) regions [4]. The visual acuity of the human eye is not spatially uniform, but the highest acuity is at the fovea rather than other parts of the retina. At a certain moment, an observer focuses on a small region of a whole display area, and the fovea sees only the central two degrees of the visual field. By taking this spatial non-uniform property of the human visual system (HVS) into account, the F+C approach displays the original resolution of a gigapixel image only within the small regions of the display on which the observer is focused, and displays lower resolution images within the other regions. Because a small number of display devices are required for an F+C display system, the cost is much smaller than that of a tiled display system.

A popular configuration of an F+C display uses a multi-projector system. For example, one projector displays a small region of an image with high resolution on which a user focuses (foveal projection), while another projector displays lower resolution images (peripheral projection) in the other regions [5]. One of the advantages of this approach is that the hi-res region can be moved either when a pan-tilt mirror is placed in front of the hi-res projector, or when the hi-res projector is mounted on a pan-tilt gantry. We call such a dynamic system a dynamic focus-plus-context (DF+C) display. The characteristics of this system are important in several different user scenarios because the region on which a user is focused normally changes according to the changes of his or her interest even when watching a static image. In addition, when dynamic image content, such as a gigapixel video [6], is displayed, the region on which the user is focused moves as his/her eyes track a moving object in the video footage. However, the image quality of foveal projection is inevitably degraded because motion blur occurs when the projected image is moved by the pan-tilt mirror or gantry.

The motion blur we focus on in this paper, which is caused by the mechanical movement of the foveal projection, is not a common issue. Suppose the peripheral projection displays a uniform black image and the foveal projection displays a white dot as shown in Fig. 1(a). When the white dot moves horizontally in the image space within a single frame, the motion blur is computed and displayed as shown in Fig. 1(c). We do not deal with such software-based motion blur in this paper. Instead, we focus on the motion blurs caused by a mechanical movement of the foveal projection, which are illustrated in Figs. 1(b) and (d). When the foveal projection displays a white dot and is vertically moved by a pan-tilt mechanism for a single frame, the motion blur occurs as shown in Fig. 1(b). When a vertically moving foveal projection

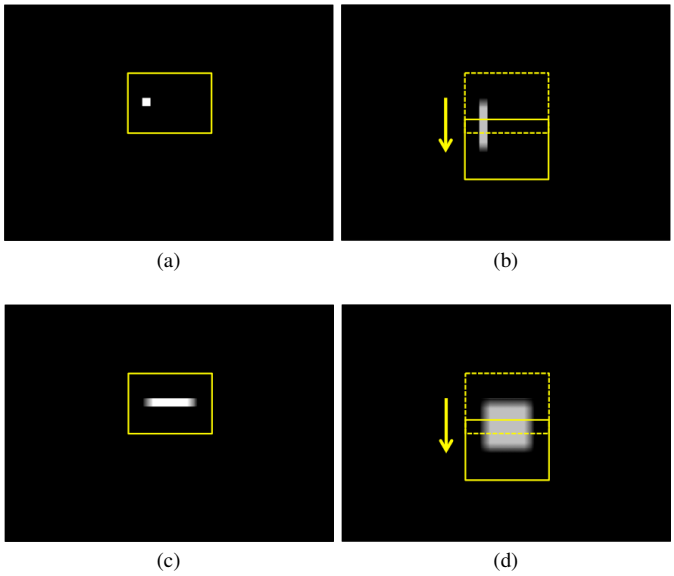

Fig. 1. Illustration of motion blur by foveal projection movement with uniform black peripheral projection: foveal projection displaying a white dot without mechanical movement (a) and with vertical mechanical movement (b); foveal projection displaying a horizontally moving white dot without mechanical movement (c) and with vertical mechanical movement (d). The yellow rectangle with the solid line and that with the dashed line indicate the foveal projection at the current frame and at the previous frame, respectively. The yellow arrow indicates the direction of the mechanical movement of the foveal projection.

displays a blurred white dot as in Fig. 1(c), the resulting image has a larger blur, because the already blurred image is blurred by the tilt movement, as shown in Fig. 1(d). This paper focuses on such mechanically caused motion blurs.

In this paper, we present an attempt to reduce the motion blur artifacts of foveal projection in a DF+C display. Our proposed method defines an error function to assess the displayed image quality as the difference between the original high resolution image and the displayed image by taking the non-uniform spatial property of human visual acuity into account. Given a video sequence, we decide a technique to move a dynamic foveal projection from one frame to another so that the sum of errors during the sequence is minimized. Figure 2 illustrates the comparison between the normal and proposed DF+C displays. If we acquire the optimal positions for the foveal projection in every frame, then one of the simplest DF+C display methods is to move the foveal projection to the optimal positions in these frames. However, such a naïve method results in undesirable motion blur artifacts in the displayed result, as shown in Fig. 2(b). On the other hand, for this particular example, our proposed method does not move the foveal projection to minimize the error (Fig. 2(c)). In the paper, we explain the principle of the proposed method as well as present the experimental results and evaluate how the proposed method can improve the displayed image quality of a DF+C display.

II. Related Work

Much research has been carried out on F+C displays. Tsai et al. and Godin et al. proposed the use of two static

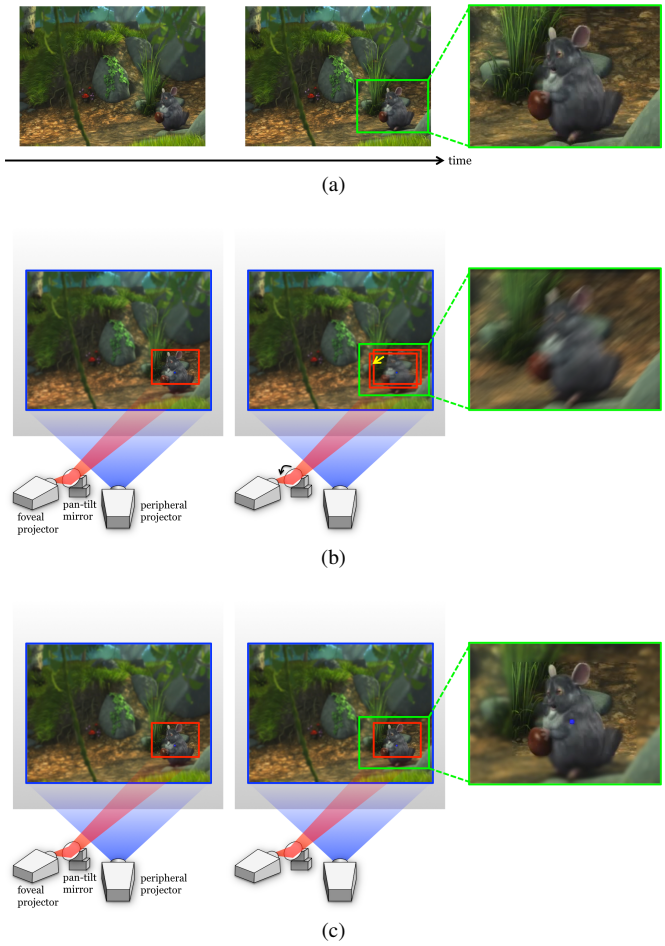

Fig. 2. Comparison between normal and proposed DF+C displays: (a) two successive frames from original hi-res video sequence, (b) projected results in normal DF+C display, and (c) projected result in proposed DF+C display. The green rectangle is used for the magnified view. The blue and red rectangles represent the projection areas of peripheral and foveal projections, respectively. The yellow arrow indicates the movement of the foveal projection.

projectors to implement an F+C display system in which one projector displayed a small portion of the original image at a high resolution and another projected the rest at a lower resolution [7], [8]. Ashdown and Robinson proposed a tabletop F+C display on which a user could interact with a working digital document in front of him/her at a high resolution, while other documents displayed around it were shown at a lower resolution [9]. Although these works applied two projectors to display hi-res and low-res images, they did not allow the hi-res or foveal projection to move within the whole display area.

On the other hand, other researchers have implemented DF+C displays by using a pan-tilt mirror or pan-tilt gantry to move the foveal projection. A pioneering work by Yamaguchi et al. proposed using two projectors for peripheral and central vision, and applied an eye tracking system for following the central vision projection to the eye-gaze point on the projection screen [10]. DF+C displays with eye-tracking have been reviewed in [11]. Yang et al. proposed a reconfigurable multi-projector wall-display system by applying eight pan-

tilt projectors to demonstrate a DF+C display [12]. Rather than preparing multiple pan-tilt projectors, Chan et al. and Staadt et al. proposed installing a pan-tilt mechanism on one of the projectors, assuming a single user of the system [13], [14]. Not only pan-tilt projectors but also handheld projectors were used in DF+C displays. Cao et al. proposed a DF+C display as one of the potential applications of their handheld projection systems, in which a handheld projector displays a hi-res image, while the other handheld projector displays a low-res image [15]. On the other hand, Cotting and Gross proposed to implement a DF+C display by using a handheld projector for the foveal projection and a fixed projector for the peripheral projection [16].

Some researchers have evaluated the usability of DF+C displays through user studies. Hsiao et al. conducted an experiment to compare route tracing performance between different F+C conditions: one was a fixed F+C and the other was a DF+C [17]. In the experiment, participants traced with their fingers a wire on an IC board from a connection point to the other end, using either fixed F+C display or DF+C display. The results demonstrated that the completion time with the DF+C was significantly shorter than that with the fixed F+C. Chen et al. proposed using a DF+C display system in visual surveillance applications and evaluated its feasibility [18]. The user study results showed that for visual monitoring tasks, the DF+C approach was significantly faster than existing approaches and was preferred by users.

Although many different issues concerning the use of multiple projectors for implementing a DF+C display have been researched until now, no serious research study has taken into consideration ways to reduce the degradation of image quality caused by the motion blur of the dynamic foveal projection. Yamaguchi et al. found that a user cannot perceive the change of displayed image resolution within 300 ms after his/her eye movement [10]. Iwamoto pointed out this is caused by saccadic suppression [11]. This is the case for relatively fast eye movements, in particular, a single movement of 15 degrees. On the other hand, our method assumes much slower (sometimes even stopped) eye movements as well. Furthermore, as described in the previous section, we do not focus on the motion blur caused by the eye movement, which is sometimes not perceived because of saccadic suppression, but on that caused by the mechanical movement of the foveal projection. Because we assume that the movement of the foveal projection is independent from the eye movement, the saccadic suppression is not an issue in our method, which, consequently, considers only motion blur. In this paper, we try to reduce the motion blur artifacts and evaluate the resulting image quality both qualitatively and quantitatively through experiments.

III. Motion Blur Artifact Reduction Method

In this section, we propose to reduce motion blur artifacts of pan-tilt foveal projection that occur when the foveal projection image moves on a display surface. First, we explain the error function developed to assess the displayed image quality of an F+C display. Next, we introduce three techniques used to move the pan-tilt foveal projection. Then, we explain how an appropriate technique to move a dynamic foveal projection to reduce the motion blur artifacts can be selected. Figure 3 illustrates the overview of the proposed method. Without loss of generality, we describe in this paper the case where a single observer uses a DF+C system consisting of a single peripheral projector and a single foveal projector, but the approach is the same given any number of observers and peripheral and foveal projectors. Note that we use rectangular foveal projection, i.e., the whole projection image region, in order to minimize the difference between the displayed result and the original image.

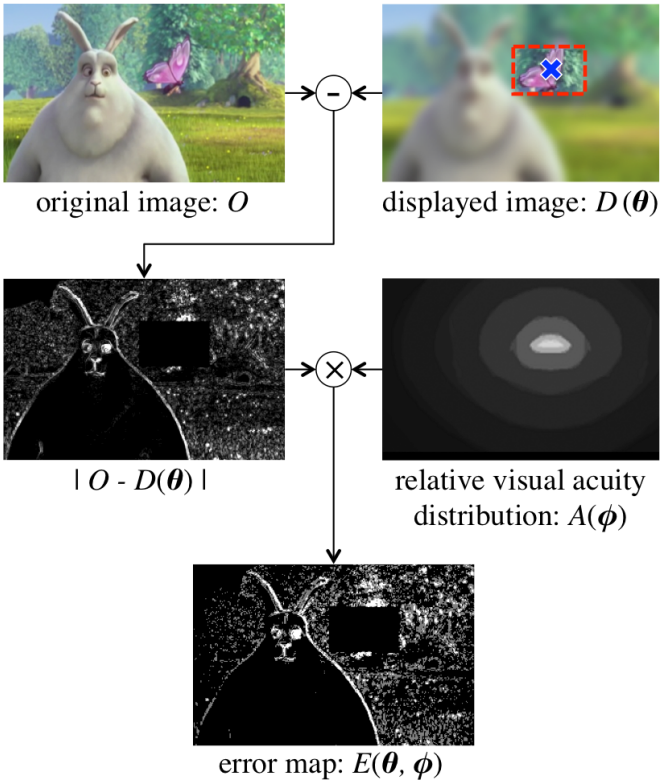

Fig. 4. Diagram of the error function computation. The blue cross and red dashed-line rectangle frame indicate the observer's gaze position and the foveal projection area, respectively.

A. Error Function

We define an error function to assess the displayed image quality of an F+C display as the difference between a displayed image and its hi-res original. Because humans can perceive hi-res images only at the fovea in the retina, it is reasonable to consider that the part of the displayed image perceived with the fovea region is more important than that perceived by other peripheral regions of the retina. Therefore, we consider the relative visual acuity as a weight in our error function, in which the acuity is multiplied by the above-mentioned difference between a displayed image and its hi-res original.

The error e is defined as the sum of weighted L1 norms over the entire display area. Suppose that O, D, A, and E represent the original image, the displayed image, the two-dimensional (2D) function of the relative visual acuity of human eyes, and

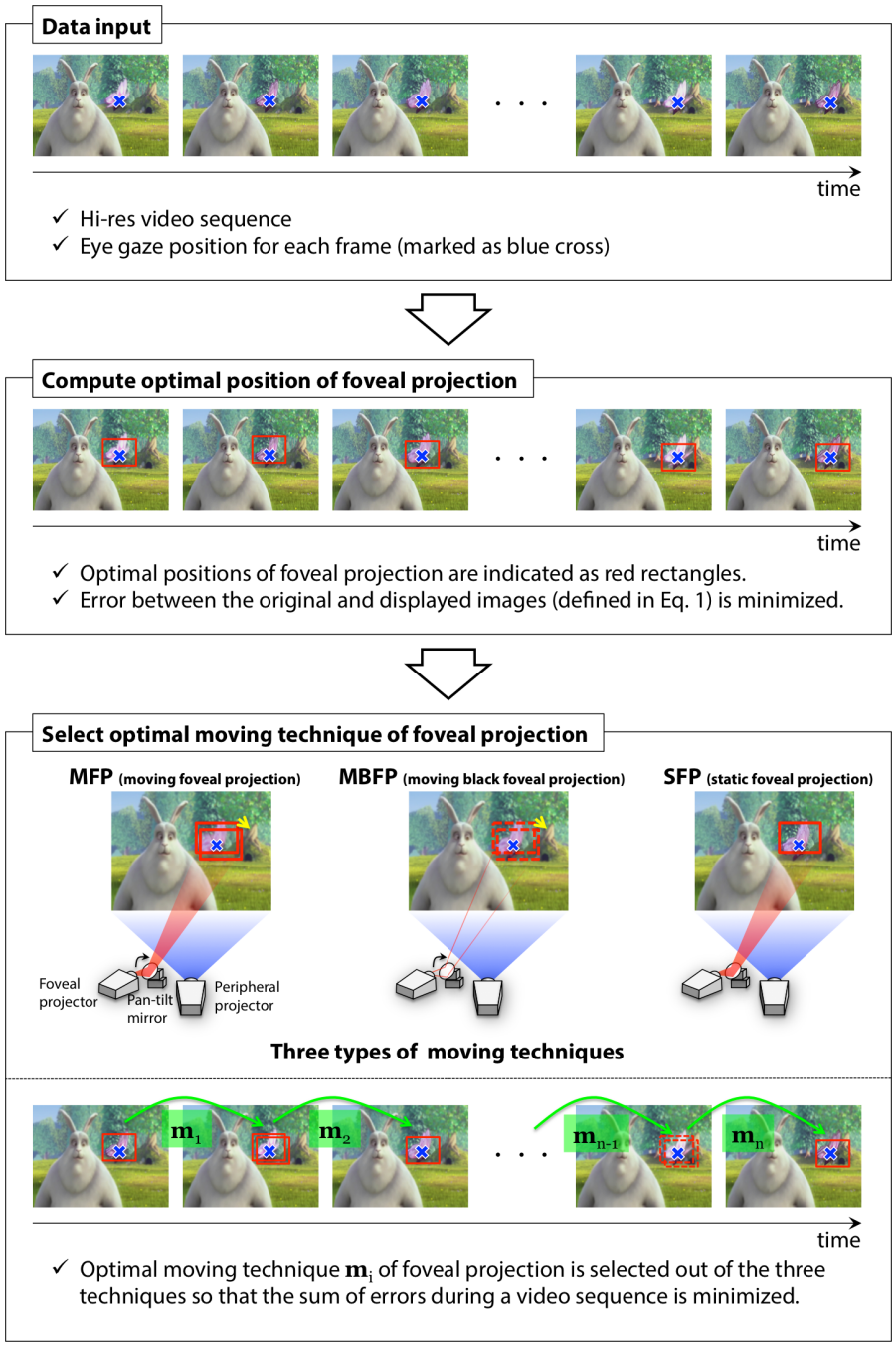

Fig. 3. Overview of the proposed technique.

the 2D weighted error map, respectively. Then

$$e(\boldsymbol{\theta}, \boldsymbol{\phi}) = \sum E(\boldsymbol{\theta}, \boldsymbol{\phi}), \tag{1}$$

where $\sum$ represents the summation operation of all the elements over E, and

$$E(\boldsymbol{\theta}, \boldsymbol{\phi}) = |O - D(\boldsymbol{\theta})|A(\boldsymbol{\phi}). \tag{2}$$

$\boldsymbol{\theta}$ and $\boldsymbol{\phi}$ represent the 2D position of the center of foveal projection in the display area and that of user's eye-gaze, respectively. Note that the process for creating the displayed image D will be explained in detail in III-B. Figure 4 illustrates the diagram of Eq. 2. The 2D weighted error map E in the figure shows error values; a higher (or brighter) intensity represents a larger error. We compute the error by calculating the difference between the displayed image D and the original image O, and then multiplying it with a weight defined by the eye position $\boldsymbol{\phi}$. Because the displayed image is not the same as the original image at the peripheral projection area, errors occur according to the positions of the foveal projection $\boldsymbol{\theta}$ and the user's eye position $\boldsymbol{\phi}$.

We apply the visual acuity model of Wertheim [19], [20] that assigns the highest value to the output at the central fovea and decreasing values to the peripheral areas. The central fovea in the eye corresponds to a gazing point on the display surface.

Therefore, the visual acuity function and, consequently, the error function are dependent on where the user's eye is gazing. In order to compute the error function correctly, we need to know where the user is gazing. There are several eye-gaze measurement methods, such as computer vision based techniques, which measure the directional orientation of eyeballs in images of faces captured either from a camera mounted on the user's head or one that is fixed on the environment.

Given a gaze point and a projection image, we search for the optimal position of the foveal projection $\boldsymbol{\theta}^{opt}$ that results in the minimum error value.

$$\boldsymbol{\theta}^{opt} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, e(\boldsymbol{\theta}, \boldsymbol{\phi}). \tag{3}$$

B. Techniques Used to Move Pan-Tilt Foveal Projections

In general, motion blur of a projected image inevitably occurs when the image is moved by a pan-tilt mechanism. Therefore, the displayed image quality is not always the best when hi-res foveal projection moves to the current optimal position $\boldsymbol{\theta}_i^{opt}$ from the previous one $\boldsymbol{\theta}_{i-1}^{opt}$ (where i represents the index of the current video frame). In some cases, a better image quality is achieved when the foveal projection stays at the same position without moving to the current optimal position. In this section, we present three different techniques used to move a foveal projection.

The first one is moving foveal projection (MFP), in which the foveal projection is moved to the current optimal point using a pan-tilt mechanism. The second one is moving black foveal projection (MBFP), in which the foveal projection displaying a uniform black image moves to the current optimal position, while the peripheral projection displays the whole image. The last one is static foveal projection (SFP), in which the foveal projection does not move.

The following part of this sub-section explains the image models of the moving techniques by which we can simulate displayed images when these techniques are applied. The displayed image $D(\boldsymbol{\theta}_i, m_i)$ at frame i consists of a peripheral projection image $P(\boldsymbol{\theta}_i, m_i)$ and a foveal projection image $F(\boldsymbol{\theta}_i, m_i)$:

$$D(\boldsymbol{\theta}_i, m_i) = P(\boldsymbol{\theta}_i, m_i) + F(\boldsymbol{\theta}_i, m_i), \tag{4}$$

where $\boldsymbol{\theta}_i$ represents the position of the foveal projection at frame i, and m_i is the index of the technique used to move the foveal projection as follows:

$$m_i = \begin{cases} 1, & \text{if moving technique is MFP} \\ 2, & \text{if moving technique is MBFP} \\ 3, & \text{if moving technique is SFP} \end{cases}.$$

The P_i and F_i of each moving technique are explained in the following paragraphs.

a) Moving foveal projection (MFP): Figure 5(a) shows the simulated displayed image $D(\boldsymbol{\theta}_i, 1)$ when MFP is applied. Figure 5(b) shows the diagram of our model to simulate the peripheral projection image $P(\boldsymbol{\theta}_i, 1)$ and the foveal projection image $F(\boldsymbol{\theta}_i, 1)$. L_i denotes the low resolution version of the displayed image, which has a spatial resolution identical to the peripheral projection. On the other hand, H_i denotes the high resolution version of that displayed image which has a spatial resolution identical to the foveal projection. We define a 2D binary mask M that is assigned the value 1 in the part of the projection image that is displayed by the peripheral projection, and the value 0 for the part that is displayed by the foveal projection. Because the foveal projection moves from the previous position $\boldsymbol{\theta}_{i-1}$ to the current position $\boldsymbol{\theta}_i$, the area for which $M = 0$ includes the areas corresponding to the previous and current positions of the foveal projection, as well as those along the path of the foveal projection. Note that the current position corresponds to the current optimal position $\boldsymbol{\theta}_i^{opt}$. Suppose that we denote this mask as $M(\boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_i)$; then, the peripheral and foveal projection images are modeled as follows:

$$\begin{aligned} \boldsymbol{\theta}_i &= \boldsymbol{\theta}_i^{opt}, \\ P(\boldsymbol{\theta}_i, 1) &= M(\boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_i) L, \\ F(\boldsymbol{\theta}_i, 1) &= (1 - M(\boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_i))(H * K(\boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_i)), \end{aligned} \tag{5}$$

where K represents the point spread function (PSF) kernel of motion blur corresponding to the path of the foveal projection. In this paper, K is a 2D matrix; some elements corresponding to the path take non-zero values and the other elements take zero (see Fig. 5(b)). The non-zero values are normalized so that their sum equals to 1.0. The motion blurred image can be computed as the convolution (denoted as "$*$" in Eq. 5) of the high resolution image and the PSF kernel.

b) Moving black foveal projection (MBFP): The MBFP technique moves the foveal projection without projecting

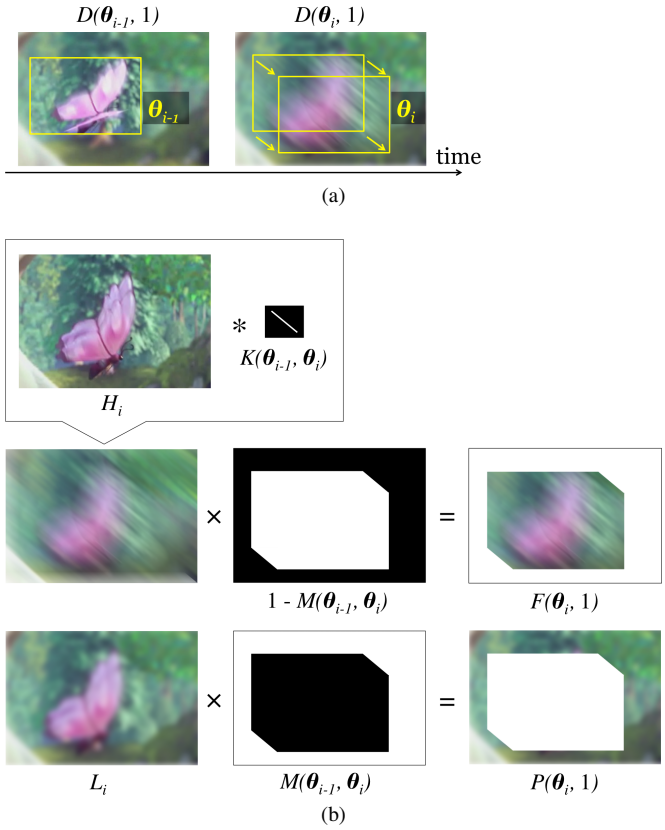

Fig. 5. MFP technique: (a) displayed pictures at frame $i-1$ and i, (b) peripheral and foveal projection images. The yellow rectangles indicate the positions of foveal projection.

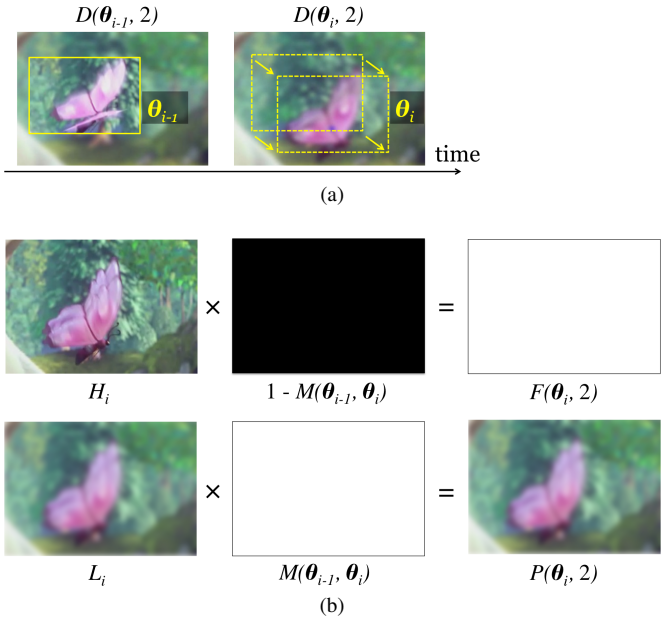

Fig. 6. MBFP technique: (a) displayed pictures at frame $i-1$ and i, (b) peripheral and foveal projection images. The yellow dashed-line frames indicate that the foveal projection displays a black image.

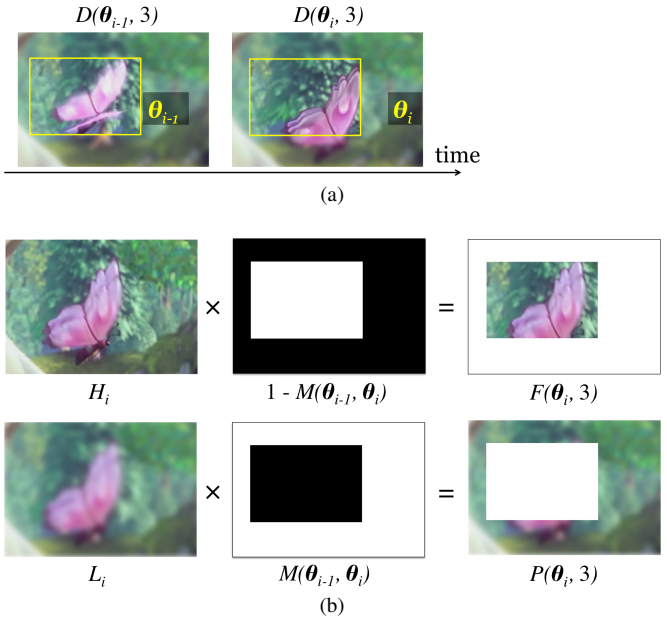

Fig. 7. SFP technique: (a) displayed pictures at frame $i-1$ and i, (b) peripheral and foveal projection images. The yellow rectangles indicate the positions of foveal projection.

anything (Fig. 6(a)). The current position corresponds to the current optimal position $\boldsymbol{\theta}_i^{opt}$ in this moving technique as well. Then, P_i and F_i are simply computed, as follows:

$$\begin{array}{lcl} \boldsymbol{\theta}_i & = & \boldsymbol{\theta}_i^{opt}, \\ P_i(\boldsymbol{\theta}_i, 2) & = & L_i, \\ F_i(\boldsymbol{\theta}_i, 2) & = & 0. \end{array} \tag{6}$$

In other words, the mask M is assigned the value 1 for the whole display surface in this moving technique (Fig. 6(b)).

c) Static foveal projection (SFP): The SFP technique does not change the position of the foveal projection (Fig. 7(a)). Therefore, the current position corresponds to the previous position in this technique. As shown in Fig. 7(b), this technique can be represented by the following equations:

$$\begin{array}{lcl} \boldsymbol{\theta}_i & = & \boldsymbol{\theta}_{i-1}, \\ P_i(\boldsymbol{\theta}_i, 3) & = & M(\boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_i) L_i, \\ F_i(\boldsymbol{\theta}_i, 3) & = & (1 - M(\boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_i)) H_i. \end{array} \tag{7}$$

C. Moving Technique Selection for Foveal Projection

We propose a technique to select an appropriate technique to move a dynamic foveal projection to reduce the motion blur artifacts during each transition from one frame to another. This technique requires the observer's gaze points for future frames. Leveraging the continuous nature of the eye-gaze movement, previous research proposed HMM (Hidden Markov Model)-based gaze prediction using Kalman filter [21]. It showed that the use's eye-gaze can be predicted 300 ms in the future with a prediction error in the viewing angle of less than four degrees. We assume that the viewing angle of foveal projection is sufficiently larger than four degrees in our DF+C system, where the above-mentioned prediction error does not cause a significant degradation of displayed image generated by the proposed technique. Therefore, this paper assumes a condition in which the gaze points of n future consecutive frames (less than a few hundred milliseconds) can be predicted by applying this technique.

Basically, our selection technique chooses the best set of moving techniques that will minimize the sum of errors (defined by Eq. 1) while moving through n frames. Suppose that $\mathbf{m}$ represents a set of moving techniques, each of which is applied to the foveal projection at each consecutive frame-to-frame transition for all the n frames $(m_i, i = 1, \ldots, n)$; then, we search for set $\mathbf{m}_{min}$, which minimizes the sum of the errors, as follows:

$$\mathbf{m}_{min} = \underset{\mathbf{m} \in m_1, \ldots, m_n}{\operatorname{argmin}} \sum_{i=1}^{n} e_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}_i m_i), \tag{8}$$

where

$$e_i(\boldsymbol{\theta}_i, \boldsymbol{\phi}_i, m_i) = \sum |O_i - D_i(\boldsymbol{\theta}_i, m_i)| A(\boldsymbol{\theta}_i, \boldsymbol{\phi}_i), \tag{9}$$

and $\boldsymbol{\phi}_i$ represents the eye-gaze position at frame i, according to Eq. 1 and 2.

There are three levels of speed for the foveal projection movement. The first level is the slowest movement, whose speed is below the threshold of seeing the movement, i.e., only imperceptible little blur is introduced. The second level movement is faster than the first and perceptible to the user, but, its motion blur is still smaller than the blur from the peripheral projection. The third level is the fastest movement, whose speed is larger than the blur from the peripheral projection. Basically, only when the third level movement occurs, the MBFP method should be used.

IV. Experiment

We evaluated the proposed method using qualitative and quantitative experiments, in which the subjects watched short video sequences projected with or without the proposed

Fig. 8. Video footages used in the simulation experiment.

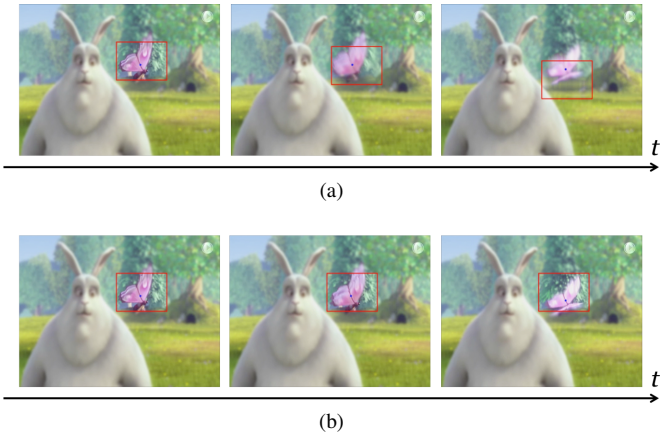

Fig. 9. Simulation results: (a) plain condition and (b) optimized condition. For clarity, red rectangles are superimposed to indicate the positions of foveal projections. They were not shown during the experiment.

TABLE I
NUMBER OF SUBJECTS WHO PREFERRED THE IMAGE QUALITY OF THE OPTIMIZED CONDITION AND THAT IN THE PLAIN CONDITION.

video	optimized condition	plain condition
butterfly	**10**	1
mouse	**10**	1
bunny	**6**	5
ski	**9**	2
football	**9**	2
artificial	**9**	2

method. We called the experiment that used the proposed method the **optimized condition**, and the one without the method the **plain condition**.

To make sure that all the subjects gazed at the same point while watching the video footage during both conditions, we superimposed a small blue circle on each frame of the video footage and asked the participants to gaze at it. We manually determined the movement of the blue circle for each video footage in advance so that it tracked a moving salient object. The display devices we used in the following experiments displayed image sequences at 60 Hz. Some of the displayed images already had software-based motion blurs.

The optimal set of techniques for moving a foveal projection to the optimized condition (i.e., $\mathbf{m}_{min}$) was computed by searching in the entire search space before the experiments. For the plain condition, only MFP was applied to the entire sequence, and thus, the indices of the moving techniques were fixed at 1, that is $m_i = 1,\ i = 1, \ldots, n$.

A. Qualitative Experiments 1: Simulation

We displayed simulated video footage on a normal LCD monitor and asked the participants to evaluate the image quality. The simulation was done based on the models described in Sec. III-B. We prepared six short video footages (which we called the "original footages"), with duration times ranging from 2.0 to 6.7 s with 30 fps. Three of the six videos were parts of a CG movie[1] (**butterfly**, **mouse**, and **bunny**); two were recorded home videos (**ski** and **football**); and one was an artificial movie (**artificial**), in which a sin wave pattern was moving horizontally over a static pink noise background (Fig. 8). The resolution of each movie was either 1280×720 pixels (butterfly, mouse, bunny, ski, and football) or 1024×768 pixels (artificial). The resolution of the sin wave was 100×100 pixels. Note that the pink noise is an image with a frequency spectrum in which the power spectral density is inversely proportional to the frequency of the signal (i.e., $1/f$). Initially, we manually determined the users' gaze points (i.e., the positions of the blue circles). Then, we computed $\mathbf{m}_{min}$ for each footage and simulated the displayed image sequence of each video footage for both optimized and plain conditions. We divided each original footage into multiple groups, each consisting of ten successive frames (i.e., $n = 10$) for the computation of $\mathbf{m}_{min}$, and then we combined them together again. The spatial resolution of the foveal projection was set to be four times higher than that of the peripheral projection. As a result, 12 (six original footages × two conditions) stimuli were simulated. Figure 9 shows an example of the computed image sequence.

The procedure of the experiment is described in the following. The subjects performed six trials using different original footage. During each trial, they watched two stimuli for each condition one by one on an LCD monitor and specified which image quality they preferred. The two stimuli were presented in a pseudo-random order. Six trials of different original video footages were also performed in a pseudo-random order.

Eleven subjects aged from 23 to 24 (9 males, 2 females) were recruited from a local university. The position and posture of each subject's head was fixed throughout the experiment. The spatial resolutions of the displayed image were 12 cycles/degree (c/d) to the subject's viewing angle for

[1]Big Buck Bunny: http://www.bigbuckbunny.org/

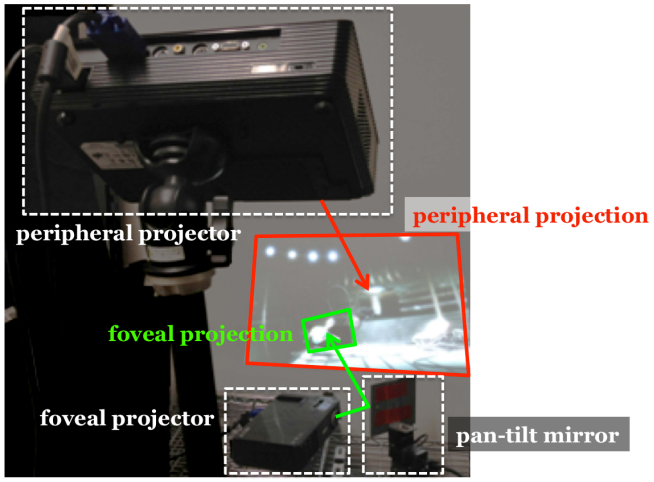

Fig. 10. Overview of the experimental system.

TABLE II
NUMBER OF SUBJECTS WHO PREFERRED THE IMAGE QUALITY OF THE OPTIMIZED CONDITION AND THAT IN THE PLAIN CONDITION.

video	optimized condition	plain condition
robot	**10**	1
butterfly	**9**	2
ski	**8**	3
artificial	**8**	3

the peripheral projection, and 48 c/d for the foveal projection. Table I shows the number of subjects who preferred the image quality in the optimized condition and those who preferred that in the plain condition. We confirmed that the subjects preferred the image qualities of the optimized condition to those of the plain condition for all the original video footages.

B. Qualitative Experiments 2: Real Projection

A real projection experiment was conducted on a system consisting of a peripheral projector and a dynamic foveal projector (Acer X1261, 1024×768 pixel) with a pan-tilt mirror (PTU-D47). Figure 10 shows the outlook of the system. Geometric and photometric corrections of projection images were required for seamless connection between foveal and peripheral projections. For the geometric correction, we used simple homography transformation for each frame. Because geometric correction is not the focus of our paper, we did not apply sophisticated and complicated techniques such as the ones presented in [22], [23]. For the photometric correction, we adjusted the white balance and maximum brightness of the foveal projector so that both the projectors displayed same color values. In addition, the edges of the foveal projection were blended for smooth connections [24].

The flow of the experiment is the same as that of the simulation experiment, except that this experiment used four original video footages. Three of the four videos were the same movies used in the simulation experiment (**butterfly**, **ski**, and **artificial**), and one was a part of a CG movie[2] (**robot**), as shown in Fig. 11. The resolution of the new video (robot) was 1280×720 pixels. The same subjects who participated

[2]Elephants Dream: http://www.elephantsdream.org/

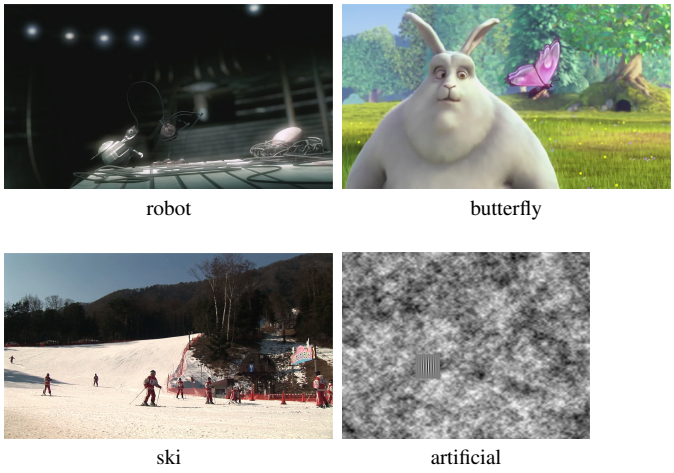

Fig. 11. Video footages used in the real projection experiment.

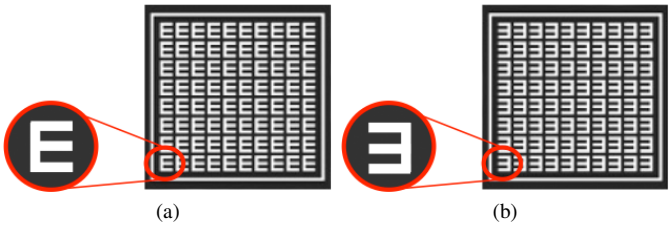

Fig. 12. Detailed views of (a) Image E and (b) Image 3.

in the simulation experiment participated in this experiment. The spatial resolutions of the displayed image were 10 c/d for the peripheral projection and 40 c/d for the foveal projection. Table II shows the number of subjects who preferred the image quality in the optimized condition and the number who preferred that in the plain condition. We confirmed that the subjects preferred the image qualities of the optimized condition to those of the plain condition for all the original video footages.

C. Quantitative Experiments 3: Real Projection

We quantitatively evaluated the image quality of the proposed method using a shape discrimination task. We horizontally moved an image consisting of multiple "E" characters (Fig. 12(a)) from the left edge to the right over the whole display area, where a pink noise image was displayed as a background (Fig. 13). While moving, the image E was randomly changed to an image consisting of multiple "3" characters instead of "E" for a short period (Fig. 12(b)). We set the center of the image E as the gaze point and computed the set of moving techniques of foveal projection $\mathbf{m}_{min}$ for the attaining optimized condition. The motion blurs of the images E and 3 caused by their horizontal movements make the discrimination of the images difficult, because the vertical lines of the images E and 3 become unclear because of the blurs, and consequently, they start to look similar. Images make up of multiple Es and 3s were also used in [17] to compare the identification performance between F+C and DF+C displays. Note that, through a preliminary experiment, we confirmed

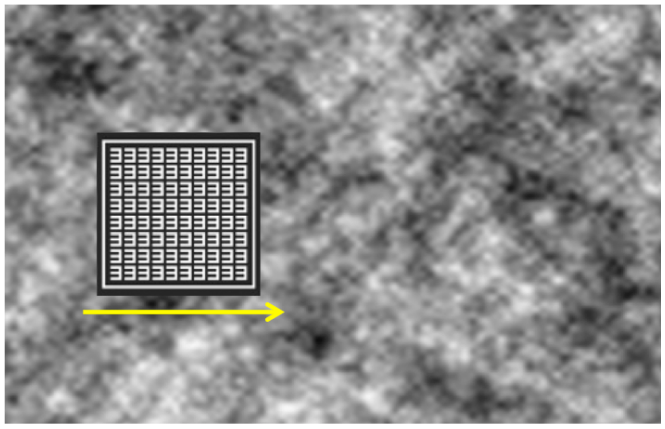

Fig. 13. Displayed image. The yellow arrow indicates the moving direction of image E and image 3, but it was not displayed during the experiment.

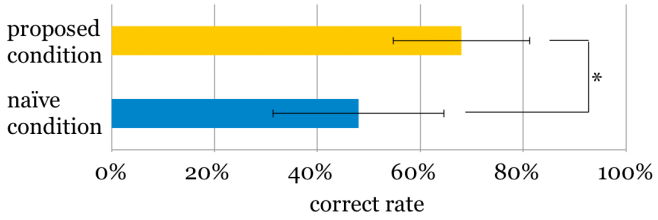

Fig. 14. Averages and standard deviations of the correct rates in the optimized and plain conditions (*: p <0.05).

that people correctly perceived the images E and 3 when they are not moving.

The procedure of the experiment is as follows. The subjects started each trial by pressing a keyboard button. Then, they gazed at image E moving from left to right in the display area and counted the number of times image E was changed to image 3. The trial was over when the image E reached the right edge of the display area. The subjects could not count correctly when significant motion blur occurred in image E. Therefore, we evaluated the image quality by comparing the correct rates between the optimized and plain conditions.

Ten subjects (males, aged from 23 to 24) were recruited from the local university. We used the same equipment used in the qualitative experiment, as described in Sec. IV-B. The spatial resolutions of the displayed image were 9 c/d for the peripheral projection and 36 c/d for the foveal projection. The speed of image E was set at 1.5 deg/s. We set the period during which image E was changed to image 3 as 120 ms. Each subject at first had three test trials, and then they performed 20 trials (10 trials for each of the two conditions). The order of trials was random among subjects.

We computed the correct rates for each subject and averaged them. Figure 14 shows the average and standard deviation of the quantitative experiment. The results show that subjects could correctly answer 48 % of the trials (at the chance level) during the plain condition, and 68 % of the trials during the optimized condition on average. In order to confirm if there is a statistically significant difference between the averages, we applied a paired t-test between them. The result of the test showed that the null hypothesis that the two average values are equal was rejected ($t_9 = 2.491,\ p < 0.05$); i.e., the difference of the correct rates between the optimized and plain conditions was significant. This confirms that the motion blur artifacts were significantly reduced in the optimized condition.

V. Conclusion

This paper presented a technique to reduce the motion blur artifacts of a foveal projection in a DF+C display. We defined an error function to assess the displayed image quality as the difference between an original hi-res image and the displayed image, taking the non-uniform spatial property of the human visual acuity into account. For a video sequence, we decided the optimal set of moving techniques of the foveal projection so that the sum of errors during the sequence was minimized. From the results of the qualitative and quantitative experiments, we confirmed that the proposed method provided better image quality when using a DF+C display and significantly reduced the motion blur artifacts compared to a conventional DF+C display.

One of our future tasks is to integrate into our system an eye-gaze tracking system that measures an observer's gaze point on a display surface and to apply the proposed method in a real user scenario. To implement this, we need to devise a better optimization method to compute $\mathbf{m}_{min}$ of Eq. 8 in real time.

References

[1] J. Kopf, M. Uyttendaele, O. Deussen, and M. F. Cohen, "Capturing and viewing gigapixel images," *ACM Transactions on Graphics*, vol. 26, no. 3, pp. 93:1–93:10, 2007.

[2] D. J. Brady, M. E. Gehm, R. A. Stack, D. L. Marks, D. S. Kittle, D. R. Golish, E. M. Vera, and S. D. Feller, "Multiscale gigapixel photography," *Nature*, vol. 486, no. 7403, pp. 386–389, 2012.

[3] A. Majumder and M. S. Brown, *Practical Multi-Projector Display Design*. A. K. Peters Ltd., 2007.

[4] P. Baudisch, N. Good, and P. Stewart, "Focus plus context screens: Combining display technology with visualization techniques," in *Proc. ACM UIST*, 2001, pp. 31–40.

[5] T.-T. Hu, Y.-W. Chia, L.-W. Chan, Y.-P. Hung, and J. Hsu, "i-m-top: An interactive multi-resolution tabletop system accommodating to multi-resolution human vision," in *Proc. IEEE TABLETOP*, 2008, pp. 177–180.

[6] S. Pirk, M. F. Cohen, O. Deussen, M. Uyttendaele, and J. Kopf, "Video enhanced gigapixel panoramas," in *Proc. ACM SIGGRAPH Asia Technical Briefs*, 2012, pp. 7:1–7:4.

[7] Y.-P. Tsai, Y.-N. Wu, and Y.-P. Hung, "Generating a multiresolution display by integrating multiple projectors," in *Proc. IEEE ProCams*, 2003.

[8] G. Godin, P. Massicotte, and L. Borgeat, "Foveated stereoscopic display for the visualization of detailed virtual environments," in *Proc. EGVE*, 2004, pp. 7–16.

[9] M. Ashdown and P. Robinson, "Escritoire: A personal projected display," *IEEE MultiMedia*, vol. 12, no. 1, pp. 34–42, 2005.

[10] H. Yamaguchi, A. Tomono, and F. Kishino, "Large visual field high resolution display employing eye tracking," *Transaction of the Institute of Electronics, Information and Communication Engineers*, vol. J73-C-II, no. 11, 1990, (in Japanese).

[11] K. Iwamoto, "Eye movement tracking type wide-angle high-resolution image display system," Ph.D. dissertation, The university of Tokyo, 2004, (in Japanese).

[12] R. Yang, D. Gotz, J. Hensley, H. Towles, and M. S. Brown, "Pixelflex: a reconfigurable multi-projector display system," in *Proc. IEEE Visualization*, 2001, pp. 167–174.

[13] L.-W. Chan, W.-S. Ye, S.-C. Liao, Y.-P. Tsai, J. Hsu, and Y.-P. Hung, "A flexible display by integrating a wall-size display and steerable projectors," *Springer LNCS*, vol. 4159, pp. 21–31, 2006.

[14] O. G. Staadt, B. A. Ahlborn, O. Kreylos, and B. Hamann, "A foveal inset for large display environments," in *Proc. ACM VRCIA*, 2006, pp. 281–288.
[15] X. Cao, C. Forlines, and R. Balakrishnan, "Multi-user interaction using handheld projectors," in *Proc. ACM UIST*, 2007, pp. 43–52.
[16] D. Cotting and M. Gross, "Interactive visual workspaces with dynamic foveal areas and adaptive composite interfaces," *Computer Graphics Forum*, vol. 26, no. 3, pp. 685–694, 2007.
[17] C.-H. Hsiao, L.-W. Chan, T.-T. Hu, M.-C. Chen, J. Hsu, and Y.-P. Hung, "To move or not to move: a comparison between steerable versus fixed focus region paradigms in multi-resolution tabletop display systems," in *Proc ACM CHI*, 2009, pp. 153–162.
[18] K.-W. Chen, C.-W. Lin, T.-H. Chiu, M. Y.-Y. Chen, and Y.-P. Hung, "Multi-resolution design for large-scale and high-resolution monitoring," *IEEE Transactions on Multimedia*, vol. 13, no. 6, pp. 1256–1268, 2011.
[19] T. Wertheim, "Über die indirekte sehschärfe," *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, no. 7, pp. 172–187, 1894.
[20] H. Strasburger, I. Rentschler, and M. Jüttner, "Peripheral vision and pattern recognition: A review," *Journal of Vision*, vol. 11, no. 5, pp. 1–82, 2011.
[21] Y. Feng, G. Cheung, W.-T. Tan, P. Le Callet, and Y. Ji, "Low-cost eye gaze prediction system for interactive networked video streaming," *IEEE Transactions on Multimedia*, vol. 15, no. 8, pp. 1865–1879, Dec 2013.
[22] M. Ashdown and Y. Sato, "Steerable projector calibration," in *Proc. of IEEE International Workshop on Projector-Camera Systems*, 2005.
[23] I. Mitsugami, N. Ukita, and M. Kidode, "Fixed-center pan-tilt projector and its calibration methods," in *Proc. of IAPR Conference on Machine Vision Applications*, 2005.
[24] R. Raskar, G. Welch, K.-L. Low, and D. Bandyopadhyay, "Shader lamps: Animating real objects with image-based illumination," in *Proc. Eurographics Workshop on Rendering*, 2001, pp. 89–102.

Daisuke Iwai received his B.S., M.S., and Ph.D. degrees from Osaka University, Japan, in 2003, 2005, and 2007, respectively. He was a visiting scientist at Bauhaus-University Weimar, Germany, from 2007 to 2008, and a visiting Associate Professor at ETH Zurich, Switzerland, in 2011. He is currently an Associate Professor at the Graduate School of Engineering Science, Osaka University. His research interests include human-computer interaction and projection-based mixed reality.

Kei Kodama received his B.S. and M.S. degrees from Osaka University, Japan, in 2010 and 2012, respectively. He is currently working for Asahi Kasei Corporation. His research interests include projection-based mixed reality.

Kosuke Sato received his B.S., M.S., and Ph.D. degrees from Osaka University, Japan, in 1983, 1985, and 1988, respectively. He was a visiting scientist at the Robotics Institute, Carnegie Mellon University, from 1988 to 1990. He is currently a Vice Dean of the Graduate School of Engineering Science, Osaka University. His research interests include image sensing, 3D image processing, digital archiving, and virtual reality. He is a member of ACM and IEEE.