ITE Trans. on MTA Vol. 3, No. 2, pp. 149-154 (2015)

Paper

Eye-Position Distribution Depending on Head Orientation when Observing Movies on Ultrahigh-Definition Television

Yu Fang †, Masaki Emoto (member)††, Ryoichi Nakashima †,
Kazumichi Matsumiya †, Ichiro Kuriki (member)†, Satoshi Shioiri (member)†

Abstract We investigated the relationship between eye-position distribution and head-orientation while viewing dynamic images in an 85 *inch* 8K ultrahigh-definition television (UHDTV). Participants watched video clips of several types of scenes made for UHDTV without any restriction of head movements. The result showed a correlation between eye and head orientation, which is similar to the coordinative movements of the eye and head for static-image viewing in previous studies. The distribution of eye positions was biased in the direction of head orientation both in horizontal and vertical directions. When the head orientated to the left/right (or up/down), the eyes also tended to direct in the left/right (or up/down) relative to the head. These findings suggest that eye-head coordination is similar between static and dynamic images and that head orientation could be used to evaluate the viewing condition and image contents of wide field of view display such as that of an UHDTV.

Key words: eye-head coordination, ultrahigh-definition television, fixation, eye position, head orientation.

1. Introduction

Visual attention selects potentially important information from a vast amount of the visual input we obtain in everyday life. The visual system analyzes objects attended. If the location of attention focus could be predicted from scene information, it is useful in many fields, such as evaluation of TV programs, advertisements, web design, and traffic environments and so on. Indeed, several visual attention models have been proposed to predict attention focus (or gaze location), calculating visual salience from stimulus images, since a pioneer work by Itti et al [1, 2]. However, the model predictions are successful only in limited cases, and researchers have found difficulties to build a general-purpose model.

Some of authors of the present study have proposed a technique to improve the prediction of attention location, using the information of head orientation [3]. This proposal is based on the fact that head orientation biases eye position distribution and information of head orientation narrows the area of attention focus predicted by an attention model. Although this model requires head orientation does not as same as the other models, it can improve the accuracy of prediction in general situations. The model can be used to estimate, for example, customers attention in shops, where the data of head orientation can be obtained from monitoring cameras. Systematic relationship between the eyes and head during gaze movements has been reported in several studies [4-7]. When we shift our gaze to far peripherally (>30°), an eye movement is frequently accompanied by a head movement, whereas head is stationary for small-gaze shifts [8, 9]. The head-movement amplitude is proportional to the gaze-shift amplitude for the large gaze shifts. However, gaze shifts sequentially during everyday life tasks such as reading or driving [5, 10] and the eye-head coordination is not predicted from that of single gaze shifts. To overcome this problem and to propose a method of gaze prediction with head orientation, a previous study investigated the eye-head coordination while viewing natural but stationary scenes [3].

To generalize the method, it is necessary to investigate the eye-head coordination with dynamic images. Dynamic image is different from static image in at least two aspects. First, moving objects or transient events attract attention, which is classified as exogenous attention. The eyes tend to move before head to shift

Received November 5, 2014; Revised December 17, 2014; Accepted Janualy 3, 2015
† Tohoku University
(Sendai, Japan)
†† NHK Science and Technology Research Laboratories
(Tokyo, Japan)

gaze by bottom-up attention [4]. This contrasts to the eye movements after head movements to shift gaze by top-down attention or during a task-related gaze shift as viewing static images[11]. Second, the eyes track the object smoothly (pursuits eye movements) when viewers look at a moving object [12, 13]. And the eyes can ignore the motion caused by self-motion [14]. Both expected and unexpected object scan be calculated in high degree of saliency using the method as the static image studies. No studies have investigated the eye-head coordination in the natural condition, where there are pursuit and saccadic eye movements. These aspects indicate that eye-head coordination can be very different between watching the dynamic images and viewing the static images.

To measure the eye and head movements in the condition as naturalistic as possible, the present study investigated the eye-head coordination under viewing the dynamic natural scenes, using ultrahigh-definition video system in a wide field of view display. The ultrahigh-definition video system has been named Super Hi-Vision [15, 16] in Japan. The Super Hi-Vision is a video format of wide field of view display, which is $7,680$ *pixels* wide by $4,320$ *pixels* tall (8K) and is known to provide realistic dynamic images. This is suitable for our purpose to investigate the eye and head movements in everyday life.

2. Materials and Methods

2.1 Subjects

Experiments were performed on twenty adult human subjects (5 males, 15 females; overall age ranging from 22 to 50 years; mean age: 37.6 years; SD: 8.74 years) with no known eye or head movement disorders and with normal or corrected to normal vision (only contact lenses were allowed in the experiment to avoid interference with the eye tracker). All subjects were naïve and paid for the experiment.

2.2 Experimental Setup

Experiments were performed in a darkroom of the NHK Science and Technology Research Laboratories using an 85 inch display ($189cm \times 110cm$, 16:9), which has a resolution of $7,680 \times 4,320$ *pixel* (33.2 *megapixels*), whose refresh rate is 60Hz (Fig.1). An electromagnetic motion tracking system (FASTRAK; Polhemus, USA) was used to measure the orientation (Azimuth and elevation, static accuracy orientation: 0.15°) of two small, light-weighted sensors. One of the sensors was put on the subjects head, and the other one was fixed on the back of chair on which the subject was sitting. These two sensors were used to record the orientation of the head related to body at a 60 Hz sampling frequency. Eye movements were recorded at 60 Hz by a wearable glasses-typed eye tracker (EMR-9; NAC, Japan), which contains three cameras; two are for recording the position of the eyes, while the scene camera in the middle has a field-of-view of 92° visual angle. All the eye, head, and chair orientation signals were synchronously recorded by a computer. The stimulus movie presented on UHDTV display was controlled by another computer. Recording and start playing of movie were synchronized by hand, pressing buttons simultaneously to initiate the processes of the two PCs. The time was confirmed by the video recorded from the EMR-9.

Fig. 1 Experimental setup.

2.3 Stimuli and Procedure

A fifteen minutes soundless ultrahigh-resolution movie of real-world scenes was presented on the display. The movie included the people such as walking in streets, surfing, paragliding, and playing football, and nature scenes such as sun rising, sea of clouds, flowers, animals and so on.

The subject was asked to perform two simple tasks, while watching the movie with free eye and head movements. One was to press a key when they noticed the change of scene (usually between clips), and the other one was to estimate the degree of blur by pushing the joystick forward when they noticed blurring of the image (sometimes fast motion created image blur). Each subject watched the same movie twice with two different viewing distances: 55 and 110 centimeters, in which the visual angles of the display were $120° \times 90°$ and $80° \times 53°$ respectively. Half of the subjects watched with $55cm$ viewing distance first and then $110cm$ distance and the other half did it in opposite order. The subjects took a rest of about 5 minutes after the first viewing distance of movie watching.

The eye tracker was calibrated before watching the movie. The eye and head movement data were recorded during the movie was being displayed and analyzed later offline.

2.4 Data Analysis

The eye positions were recorded as the angle of eyes relative to the head. The head orientation relative to the body was obtained by the difference between the two sensors, one of which was put on the subjects head, and the other one was fixed on the back of chair. Gaze position was derived by the sum of the eye positions and the data of head orientation relative to the body.

To determine the saccade and fixation of gaze movements, the velocity and acceleration of gaze movements were calculated. The onset of each saccade was defined as the first frame at which both gaze velocity and gaze acceleration exceeded thresholds: $30°/s$ for velocity and $50°/s^2$ for acceleration. In a similar fashion, the end of saccade was the first frame at which both velocity and acceleration fell below the thresholds. The period of fixation was determined as the time between the end of the previous saccade and the onset of the next saccade. The eye position and head orientation were calculated by averaging the data of eye and head orientation in each fixation period. Using the pairs of head orientation and eye position data, eye-position distribution was obtained after classifying the head orientation for several bins with interval of 3° for head movements within $\pm 7.5°$ (the data beyond $\pm 7.5°$ were pooled into one bin for each direction).

Note that the fixation period here implies the period between two saccades, and we did not distinguish fixation and pursuit periods in the present study.

3. Results

Fig. 2 illustrates the distribution of head orientation. The distribution varies with viewing distances. The ordinate shows the proportion of the number of fixations with each head orientation bin with the interval of 3°, and the abscissa shows the head orientation relative to the body. Fig. 2a and 2b show the distribution of head orientation in horizontal direction, where positive numbers indicate the right. Fig. 2c and 2d show the vertical orientation, where positive numbers indicate the up. The numbers above the histogram indicate the viewing distance, $55cm$ or $110cm$.

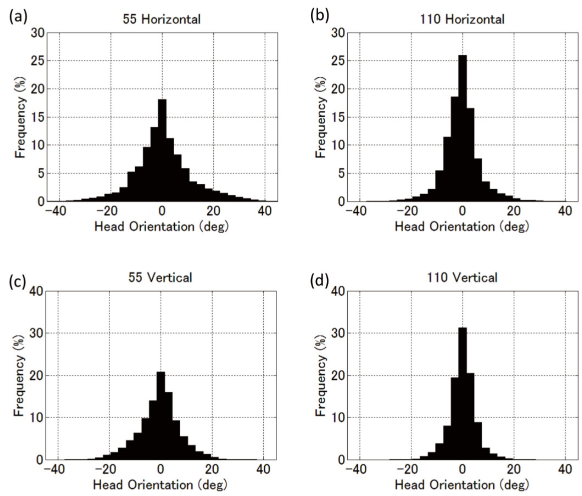

Fig. 2 Head-orientation distribution

Head orients more in the center than in the lateral region in all four combinations of viewing distance and movement directions. This shows clear tendency that the head keep the same orientation with the body while watching the dynamics images, which is similar to the previous study with the static images [17].

The range of head movement, that is how largely head-orientation distributes from body direction, can be estimated by the standard deviation of the distribution of head orientation. The standard deviation analysis demonstrated that head moved more widely with a $55cm$ viewing distance than with a $110cm$ viewing distance both in horizontal [4.87 vs. 6.76, paired t-test, $t(19) = -5.76$, $p < 0.001$] and in vertical [5.75 vs. 7.73, paired t-test, $t(19) = -4.35, p < 0.001$] directions. These results indicate that the head moves more with closer viewing distance, that is, with larger size of display in visual angle.

The effect of display size on head movements found in this experiment is consistent with that for using simple LED stimuli and tasks. Experiments with simple stimuli and tasks reported that head moves more with larger gaze shifts [8, 9]. In the present experiment, the display size is larger in visual angle with closer viewing distances, and thus larger gaze shifts are expected to occur to look at the same locations in the images in two viewing distances. Therefore, larger head movements are expected with closer viewing distances because of the larger range of gaze shifts.

Fig. 3 shows the distribution of eye positions for different ranges of head orientation separately in different panels; the data labeled 0° has eye positions with head orientation between ±1.5°. Similarly, ±3° indicates between ±1.5° and ±4.5°, ±6° indicates between ±4.5° and ±7.5°, and ±9° ±21° indicates between ±7.5° and ±22.5°. Positive in horizontal direction indicates right relative to the body, and positive in vertical indicates up relative to the chest direction defined as the facing direction of the chair. The ordinate shows the percentage of the number of fixations and the abscissa shows the eye direction relative to the head. We fitted a Gaussian function to each set of data using a least squares method to approximate the distribution of eye positions. The red line in each histogram shows the function fitted to the distribution of eye positions.

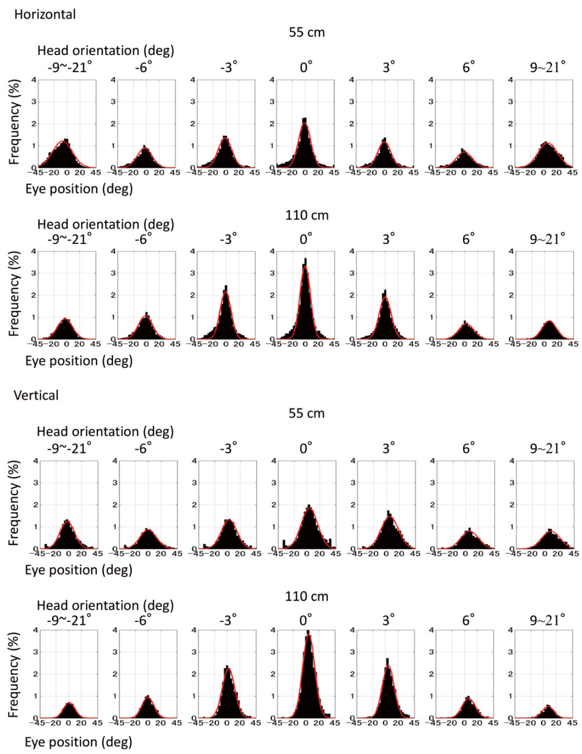

Fig. 3 Head-orientation effect on eye-position distribution.

The distribution of eye positions was biased according to head orientation for all four combinations of two directions (horizontal/vertical) and the two viewing distances (55/110cm). Specifically, the eyes tended to direct in the left/right (or up/down) relative to the head when the head orientated to the left/right (or up/down). To illustrate this effect clearly, we plotted the peak position of the eye-position distribution as a function of head orientation in Fig. 4. They were obtained from the fitted Gaussian function for each subject and averaged over 20 subjects. Different colors of symbol represent the different viewing distances: blue represents 55cm and red represents 110cm viewing distance. The line shows a linear function fitted to the data points. Error bars represent the standard error of mean.

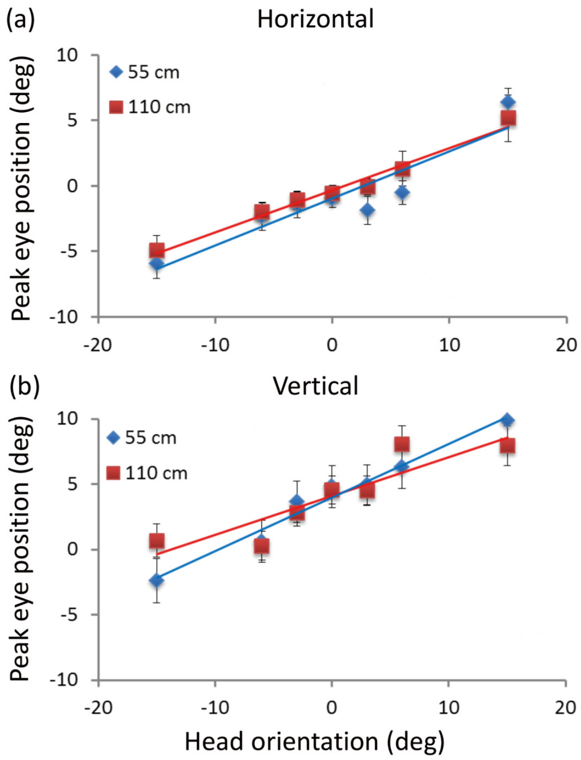

Fig. 4 Relationship between the peak eye position and head orientation.

A relatively simple relationship is seen between the peak eye position and head orientation in Fig. 4. Peak position monotonically shifts with head orientations, and the effect is similar on the horizontal and vertical sides. The slope of the fitted lines shows magnitude of head orientation effect on eye positions: steeper indicates larger contribution. Although the slope averaged over subjects is larger for 55cm viewing distance than for 110cm viewing distance (0.56 and 0.57 for 55cm; 0.39 and 0.37 for 110cm), a paired t-test shows no significant difference either in horizontal [paired t-test, $t(19) = 1.18, p = 0.25$] or in vertical [paired t-test, $t(19) = 1.01, p = 0.33$] direction.

To avoid the influence of head movements, we analyzed fixations without any head movement. The slope averaged over subjects seems larger for 55 cm viewing distance than for 110 cm viewing distance (horizontal and vertical slopes are 0.36 and 0.45 for 55cm and 0.29 and 0.30 for 110cm), a paired t-test shows there is significant difference both in horizontal [paired t-test, $t(19) = 2.87, p < 0.01$] and in vertical [paired t-test, $t(19) = 5.88, p < 0.001$] direction. The effect of head orientation on the eye position distribution was larger in 55cm than that in 110cm of viewing distance.

4. Discussion

The experiment showed that the relationship between eye and head orientations is helpful to estimate gaze locations in dynamic images. The distribution results indicate that the head orientation provides information about eye positions. When the head was orientated to the left (or right), the eyes also tended to direct to the left (or right) relative to the head (Fig. 3). This relationship is consistent with that from the experiment using visual search and static scene viewing tasks [17, 18]. This information can be used to improve prediction of the gaze location with saliency map models as has been indicated by our previous report [3]. In the report, we have proposed a method to use head orientation to narrow the possible gaze locations from salient locations in images calculated by a saliency model. Since eye position distributed mostly within a range ±20° (see Fig. 3) around the peak, multiplying the distribution function to the salience map as weights provide better gaze estimations. Two dimensional eye positions distribution was modeled based on horizontal and vertical components and applied to the saliency map calculated by the attention model proposed by Itti et al [1, 2]. The evaluation result revealed that the accuracy of the gaze prediction model is improved when the saliency map and the head orientation information are combined during static natural scene viewing. The present results indicate that head orientation can also be used for gaze prediction on natural movies in wide field of view displays.

Eye-head coordination in vertical direction was similar to that in horizontal direction in the present experiment. When the head was orientated to the upward (or downward), the eyes also tended to direct to the upward (or downward) relative to the head (Fig. 3). The slope in vertical direction was similar to that in horizontal direction, 0.35 and 0.57 for 110cm (53°) and 55cm (90°). This indicates that vertical head orientation can also be used for gaze prediction. We will be working on prediction of attention focus by the model using the effect of head orientation on the eye movement from the present experiment.

The magnitude of head orientation effect can be evaluated by the slope of the function of eye distribution peak against head orientation. The slope in the present experiment was 0.39 for 110 cm (80°in visual angle) and 0.56 for 55cm (120°) for the head orientation range of ±22.5° for horizontal direction, and 0.35 and 0.57 for 110cm (53°) and 55cm (90°) for vertical direction. Although there was no significant difference shown while head movements were included between the two viewing distances for both directions, the head orientation effect were larger in 55cm than that in 110 cm of viewing distance while stationary head. A previous study investigated eye-head coordination during visual search and estimated the slope of the eye-position peak vs head orientation function. The slope in the visual search experiment was around 0.96 in the region of about ±20°[18]. The visual stimuli were presented in a 360° visual angle display system in the visual search experiment.

There may be a general trend that the head orientation effect, that is, the slope of head orientation against peak position of eye position distribution, increases with the size of stimulus in visual angle. To examine whether the head orientation effect varies with the stimulus size, we divided the slope of head orientation effect by each stimulus size, 80° (110cm viewing distance), 120° (55cm viewing distance) and 180° (visual field size, instead of 360°, as the field size of simultaneous visible field in a 360° display). The number is about constant (0.0049, 0.004, and 0.0053 for 110cm viewing distance, 55cm viewing distance and 360° display). The head orientation effect to the eye position distribution may be determined relative to the size of stimulus field. In the other words, the head orientation effect may be in a high correlation with the size of stimulus field. Although this needs further investigation, it is likely that head movements are more important for wider field of view and recoding head movement as well as eye movements helps evaluation of display and contents for wide field of view displays.

5. Conclusion

We investigated eye-head coordination while watching a natural movie on ultrahigh-definition television. The present study reveals that head orientation influences on the distribution of eye positions for dynamic-images observation both in horizontal and vertical directions. This information can be used to improve prediction of the attention focus by saliency map models.

6. Acknowledgments

This study was supported by the Core Research for Evolutional Science and Technology (CREST) Program of the Japan Science and Technology Agency (JST) to SS.

References

1) L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," Ieee Transactions on Pattern Analysis and Machine Intelligence, vol. 20, pp. 1254-1259 (Nov 1998)
2) L. Itti, "Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes," Visual Cognition, vol. 12, pp. 1093-1123 (Aug 2005)
3) R. Nakashima, Y. Fang, A. Hiratani, K. Matsumiya, I. Kuriki, and S. Shioiri, "Gaze estimation with a saliency map and head movements.," presented at the International Joint Workshop on Advanced Sensing / Visual Attention and Interaction (ASVAI2013), Okinawa, Japan (November, 2013)
4) W. H. Zangemeister and L. Stark, "Types of gaze movement: variable interactions of eye and head movements," Exp Neurol, vol. 77, pp. 563-77 (Sep 1982)
5) M. Land, N. Mennie, and J. Rusted, "The roles of vision and eye movements in the control of activities of daily living," Perception, vol. 28, pp. 1311-1328 (1999)
6) D. Guitton, "Control of Eye Head Coordination during Orienting Gaze Shifts," Trends in Neurosciences, vol. 15, pp. 174-179 (May 1992)
7) E. G. Freedman and D. L. Sparks, "Eye-head coordination during head-unrestrained gaze shifts in rhesus monkeys," J Neurophysiol, vol. 77, pp. 2328-48 (May 1997)
8) D. Tweed, B. Glenn, and T. Vilis, "Eye-head coordination during large gaze shifts," J Neurophysiol, vol. 73, pp. 766-79 (Feb 1995)
9) J. S. Stahl, "Amplitude of human head movements associated with horizontal saccades," Exp Brain Res, vol. 126, pp. 41-54 (May 1999)
10) E. Kowler, Z. Pizlo, G. L. Zhu, C. Erlekens, R. Steinman, and H. Collewijn, "Coordination of head and eyes during the performance of natural (and unnatural) visual tasks," in The Head-Neck Sensory Motor System, ed Oxford: Oxford University Press, pp. 419-426 (1992)
11) A. Doshi and M. M. Trivedi, "Head and gaze dynamics in visual attention and context learning," in Computer Vision and Pattern Recognition Workshops, 2009. CVPR Workshops 2009. IEEE Computer Society Conference on, pp. 77-84 (2009)
12) A. Z. Khan, P. Lefevre, S. J. Heinen, and G. Blohm, "The default allocation of attention is broadly ahead of smooth pursuit," J Vis, vol. 10, p. 7 (2010)
13) G. R. Barnes, "Cognitive processes involved in smooth pursuit eye movements," Brain Cogn, vol. 68, pp. 309-26 (Dec 2008)
14) A. Hiratani, R. Nakashima, K. Matsumiya, I. Kuriki, and S. Shioiri, "Considerations of Self-Motion in Motion Saliency," in The 2nd IAPR Asian Conference on Pattern Recognition (ACPR2013), Okinawa, Japan, pp. 783-787 (November 5-8, 2013)
15) M. Sugawara, M. Kanazawa, K. Mitani, H. Shimamoto, T. Yamashita, and F. Okano, "Ultrahigh-definition video system with 4000 scanning lines," Smpte Motion Imaging Journal, vol. 112, pp. 339-346 (Oct-Nov 2003)
16) M. Emoto, K. Masaoka, M. Sugawara, and Y. Nojiri, "The viewing angle dependency in the presence of wide field image viewing and its relationship to the evaluation indices," Displays, vol. 27, pp. 80-89 (Mar 2006)
17) R. Nakashima, Y. Fang, K. Matsumiya, R. Tokunaga, I. Kuriki, and S. Shioiri, "Eye position distribution depends on head orientation in natural scene viewing.," presented at the APCV, Incheon, Korea (2012)
18) Y. Fang, R. Nakashima, K. Matsumiya, R. Tokunaga, I. Kuriki, and S. Shioiri, "Eye position distribution depends on head orientation," presented at the VSS, Naples, USA (2012)

Yu Fang obtained his bachelor degree in engineering at Guangdong Ocean University of China in 2006. He then went to Japan and received his Master degree of Information Sciences at Tohoku University in 2012. He is currently in the third year of his Ph.D. course at the laboratory of visual cognition and systems in the Research Institute of Electrical Communication at Tohoku University. His study interest is the mechanisms of human eye-head coordination.

Masaki Emoto received his B.S. and M.S. degrees in electronic engineering and his Ph.D. degree in human and environmental studies from Kyoto University in 1986, 1988, and 2003, respectively. He has been working on future television systems and human science, especially human visual system, at NHK Science and Technical Research Laboratories. His research interests include future television systems incorporating desirable effects such as a heightened sensation of presence, and eliminating undesirable side effects to the viewer such as photosensitive response, visually induced motion sickness, and visual fatigue.

Ryoichi Nakashima received his B.E. and M.E. degrees in urban engineering from The University of Tokyo, Japan in 2004, 2006, respectively. He received his Ph.D. degree in psychology from The University of Tokyo, Japan in 2011. From 2011 to 2014, he was a postdoctoral fellow in Research Institute of Electrical Communication, Tohoku University. He is currently a postdoctoral fellow in RIKEN. His research interests include human visual attention, visual perception and visual cognition.

Kazumichi Matsumiya received Ph.D. degree from Tokyo Institute of Technology in 2000. Then, he was a postdoctoral researcher at York University in Canada until December, 2001. From January, 2002 to December, 2003, he was a postdoctoral researcher at Imaging Science and Engineering Laboratory, Tokyo Institute of Technology. From January, 2004 to March, 2005, he was a research fellow at Human Information Sciences Laboratory, ATR. He moved to Research Institute of Electrical Communication, Tohoku University as a research associate in April, 2005. He is currently working there as an associate professor from 2014. His research interests are in visual psychophysics, multimodal integration, motion perception, eye movements.

Ichiro Kuriki received the B.E. degree from University of Tokyo in 1991, and received M.E. and Ph.D. degrees from Tokyo Institute of Technology in 1993 and 1996, respectively. During 1996-1998 he was with Tokyo Institute of Technology, 1998-2000 with the University of Tokyo, and 2000-2005 with NTT Communication Science Laboratories. He is now with Research Institute of Electrical Communication, Tohoku University. His main interest in research is the mechanisms of human color perception, especially in the visual cortex.

Satoshi Shioiri received Dr.Eng. in 1986 from Tokyo Institute of Technology. Then he was a postdoctoral researcher at University of Montreal and at ATR before moved to Chiba University in 2005 as an assistant professor, an associate professor, and a professor. He moved to Tohoku University in 2005 as a professor of Research Institute of Electrical Communication of Tohoku University. His research interests include motion perception, depth perception, color vision, mechanisms of visual attention and eye movements, and modeling of visual functions.