Evaluation of Discomfort Degree Estimation System with Pupil Variation in Partial 3D Images

Shoya Murakami$^{(\boxtimes)}$, Kentaro Kotani, Satoshi Suzuki, and Takafumi Asao

Department of Mechanical Engineering, Kansai University, Osaka, Japan
{k202465,kotani,ssuzuki,asao}@kansai-u.ac.jp

Abstract. The purpose of this paper was to examine whether the changes in pupil diameter can refrect on the degree of discomfort by various levels of partial 3D images as well as other validated characteristics. Moreover, we discuss the effectiveness of the systems while guiding visual attention by partial 3D images. Images chosen from IAPS (International Affective Picture System) were used to make 3D images. Power spectrum ratio of the pupil variation, called S/C value, generated by stimuli images to those by control images was calculated. The relationship between VAS scores for the impression regarding projected images and the S/C values was set to the major concern for this study. As a result, the average S/C values in 2D neutral images ranged 0.634 to 1.318, whereas the average S/C values in partial 3D neutral images ranged 0.412 to 1.552. VAS scores in 2D neutral images ranged 3.6 to 8.5 and that in partial 3D neutral images ranged 1.2 to 7.4. Moreover, correlation coefficients between VAS scores and S/C values in 2D neutral images was 0.116 and those in partial 3D neutral images was −0.114. In partial 3D images, this negative correlation coefficient was found in consistent with the previous study, whereas the correlation coefficients for both images were relatively low. It was suggested that S/C value was likely to use as a candidate for discomfort measure with a modification of collection technique for VAS scores during the experiment.

Keywords: Partial 3D · Pupil variation · Visual attention guidance

1 Introduction

In recent years, numerous visual information is sent to us through the medium of outdoor advertising and web pages, however, we cannot process all the information due to limited capacity for visual and memory functions, resulting in ineffective collection of useful information [1]. Providers for such advertising information were eager to optimize the technology to send such information effectively and thus, visual attention guidance has been receiving attention among providers as well as researchers.

The technology of visual attention guidance can make users focusing on an intended part of visual images for a longer period of time than usual and leave a significant impression on users [2]. Therefore, both of users and publishers can get the benefits by using the visual attention guidance [1].

However, it was reported that people felt the guidance to be factitive when they found themselves being forced to guide their visual attentions, and that spoiled the impression of the image of information turning into negative impressions [3]. This led to the study about visual attention guidance without users' awareness. Hagiwara et al. [4] reported a study on the visual attention guidance by applying change of colors. However, the disadvantages of this method were that the quality of information at a guidance destination were changed and users may be confused by finding the color of the part changing if they know its original color [1]. Hata et al. [1] also reported a study on visual attention guidance using image resolution control. The disadvantages of this method were that the information was forced to be changed by lowering resolutions except the area of guidance destination. Both of these methods apparently change the quality of information.

In our study, use of partial 3D images was used for a possible solution to maintain the quality of information in applying visual attention guidance. Partial 3D image refers to an image converted to 3D by augmenting cross parallax to a specific area of the 2D image [5]. To eliminate any undesired parallax at the other areas, a fixed level of non-cross parallax is augmented [5]. It was reported that the area augmented parallax was being focused on [6] and such cases were happened regardless of whether the viewers were aware of 3D at the area [7]. Therefore, use of partial 3D images was a potential alternative to make users guide their visual attentions without giving a feeling like "forcing to be guided". By using this method, the modification of information at a guidance-destination was minimum and the information shown other than the area of a guidance destination stayed unchanged.

However, how much negative feeling can be surpressed by using the partial 3D images should be evaluated. This study focused on the effect of use of partial 3D images for avoiding unpleasant feelings about visual attention guidance.

The final purpose of our study is to investigate usefulness of the unaware visual attention guidance by partial 3D images by comparing the degree of discomfort between the aware and unaware visual attention guidances. However, the subjective evaluation may suffer from lower reliability of unpleasant feelings because people have different standards for positive or negative feelings between aware and unaware visual attention guidance. Thus, it is necessary to construct the objective evaluation to estimate the degree of discomfort.

In this study, we used the pupil diameter variation to estimate the emotion. The pupil is controlled by the automatic nervous systems and reflects emotional variation [8]. Kawai et al. [9] reported that the pupil variation was changed by whether positive or negative feelings induced by images visually given to the participants.

To estimate the degree of discomfort, based on the previous study [9], we built the systems to estimate the degree of discomfort with the power spectrum of the pupil variation.

The purpose of this paper is to examine whether or not the evaluation systems for the degree of discomfort is valid at partial 3D images. We further discuss the effectiveness of the systems while guiding visual attention by using partial 3D images.

2   Pupillary Response

When the light with constant brightness is given to the pupil, a pupil contraction is generated [8]. If the light level becomes low, the pupil size is back to the normal size [8]. The higher the light level is, the larger the amount of the pupil contraction is and the longer the pupil stays in its contraction [8]. The pupil size gets smaller when the light level gets high [8]. In the pupillary response to light, in case the light level is gradually increasing, the pupil size does not vary [8]. However, in case of the low light level, the pupil size varies when the light level suddenly increases [8]. In case the light stimulus keeps being given to the pupil with a duration of one second, the pupil generally starts contracting after the latency for 0.2–0.3 s [10]. Moreover, the pupillary contraction reaches to the maximum a second after the onset of the light capture, then the pupil dilates and it is back to normal [10].

Kawai et al. [9] calculated the power spectrum ratio of the pupil diameter variation while presenting control images (non-stimuli) and stimuli (positive, neutral, negative) images. As a result, they found that the power spectrum ratio in each stimulus was different below the frequency components of 0.4 Hz [9]. Moreover, Kawai et al. [11] reported that the pupil diameter was contracted after presenting positive images and dilated after presenting negative images. However, Hess [12] reported that the pupil was dilated not only when we were being given the workload but also the arousal was increased, and the pupil was contracted when the arousal was decreased. Murakawa et al. [13] reported that the pupil diameter was dilated when the visual stimuli were of particular interests. Thus, it is important to give participants the workload and the stimuli, which gives minimum interest to focus on positive/negative impressions to participants.

## 3   Methods

3.1   Participants

Five undergraduate students (M_{age} = 21.8 years; 5 males) participated in this experiment. All participants had corrected-to-normal vision.

3.2   Experimental Setup

Figure 1 shows the schematic view of experimental device. Pupillary responses, specifically pupil area, were recorded by using an eye-tracker (EyeLink II, SR-Research) at 250 Hz of sampling rate. EyeLink II was a head-mounted device, as shown in Fig. 2 and detected the pupil by infrared ray radiation. When infrared rays reached the iris, the part except for the pupil reflected the rays [14]. The eye tracker detected the non-reflection part as the pupil and generated output the area as pixel data. Camera units provided an accurate measure of pupil size across variations in eye shapes and camera angles.
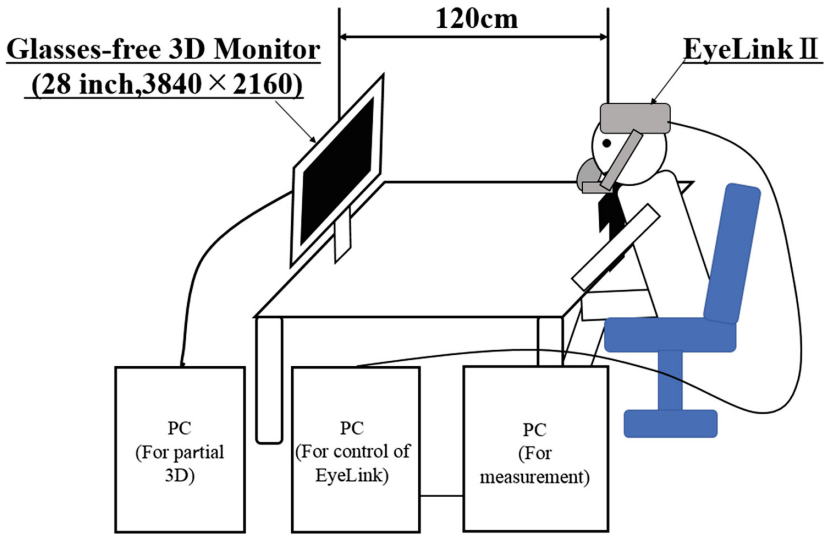
**Fig. 1.** Schematic view of experimental device
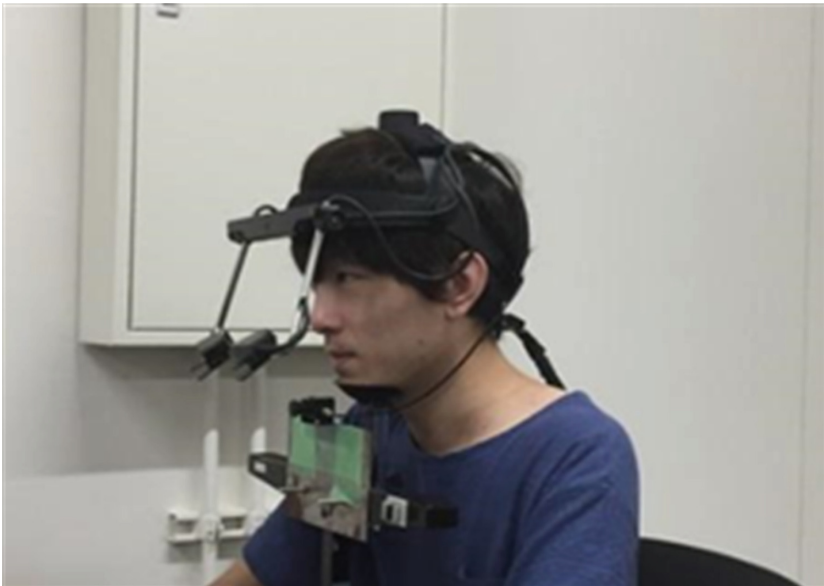
Fig. 2. Posture during experiment using eye tracker

A total of four neutral images were used in this experiment. The images were chosen from the IAPS (International Affective Picture System) [15]. They were expressed by three criteria (valence, arousal, and dominance), each of which were evaluated with nine grades. Based on the previous study [16], using these three criteria,

the neutral images were chosen as valence of 5.0, arousal of 3.0 and dominance of 6.0 in this study. The size of visual stimulus was set to 1024×768 pixels. The distance from participants to the display was 120 cm. 3D images were displayed with a 4 K glasses-free 3D monitor (Let's Corporation).

Image processing

Control and partial 3D images: The pupil diameter varied with the luminance change of the presented image. Therefore, the control image of each stimulus image was created such that there were brightness differences between the control and stimulus images. Figure 3 shows the example of control images. Control images were generated by using the procedure by kawai et al. [9].

(a) Stimulus image selected from the IAPS   (b) Control image
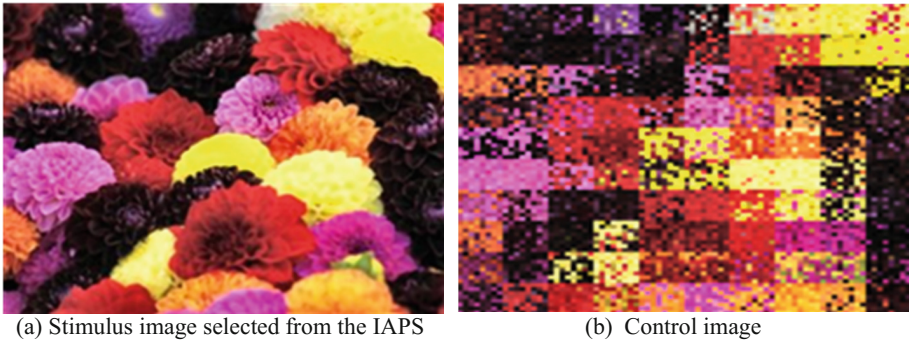
Fig. 3. Conversion to control image

Figure 4 shows the example of partial 3D images. In this study, the partial 3D images were created by Tridef 3D Photo Transformer.

(a) Original image selected from the IAPS   (b) Depth map
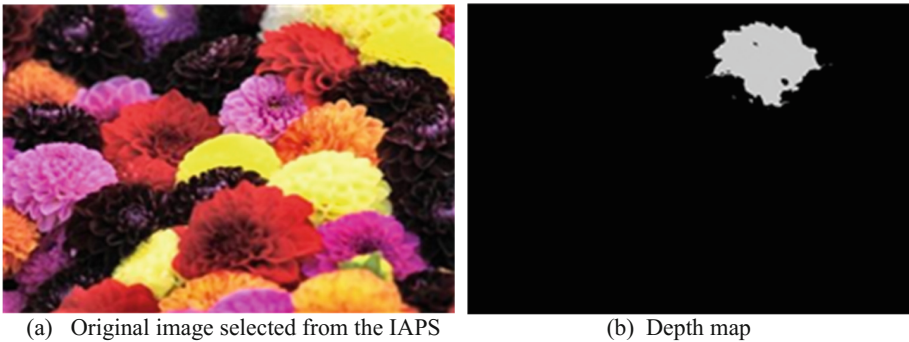
Fig. 4. Sample image of partial 3D conversion, the area for 3D conversion is shown in grey on depth map (right)

3.3 Experimental Procedure

Before the experiment, we showed participants several partial 3D images to minimise their emotional changes caused by the initial interest in partial 3D. Figure 5 shows the time chart of the experimental session including the time sequence for presenting images. In the experiment, this session was performed for each image. A control image, two stimuli images (2D, partial 3D) and three masking images were presented in each session. Participants were allowed to blink while the monitor was displaying the masking images, but we instructed them not to blink while the monitor was displaying the control images and stimuli images in order to obtain secure data for pupil area. Participants were asked to watch the images while this session. After finishing all the sessions, Visual Analogue Scale (VAS) was given to participants for collecting the degree of positive/negative feelings for the images.
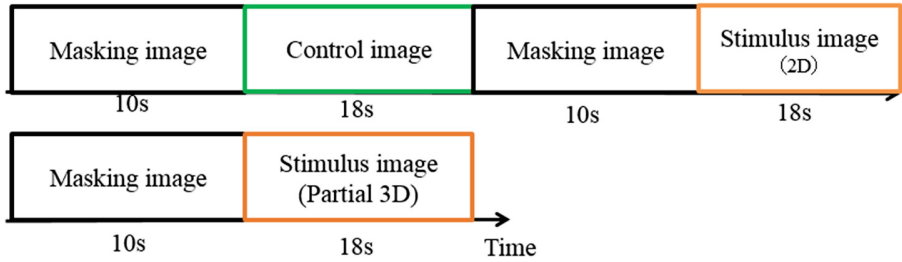

Fig. 5. Time chart of the experimental session

3.4 Data Analysis

The pupil area partially lost by blink were linearly interpolated between 100 ms around the time of blink based on the previous study [17] and a 5-points moving average was applied to the time course of the pupil areas for smoothing [9]. Fast Fourier Transformation for 2048 units in time-series data of the pupil area (8.192 s) was proceeded. Latency for the changes in pupil diameter was reported as 0.2–0.3 s by Kondo et al. [18]. We, however, set the latency for two seconds after presenting the images since the pupil variation was affected by the emotional changes generated by bright changes when images were changed. Based on the previous study [9], the power spectrum ratio of the pupil variation in stimuli images to that in control image was calculated from Eq. (1).

$$S/C = \frac{\sum_{f=0.3\text{Hz}}^{1.6\text{Hz}} P(t)}{\sum_{f=0.3\text{Hz}}^{1.6\text{Hz}} P(c)} \tag{1}$$

P(t):   Power spectrum of the pupil variation caused by presenting stimulus image
P(c):   Power spectrum of the pupil variation caused by presenting control image

Generally, it was suggested that the peak frequency range for most pupil responses were lower than 1.6 Hz [14]. Takahashi and colleague also reported that pupil variation contained 0.05–0.3 Hz of frequency components was considered as pupillary noise [14]. Thus, the frequency range of the pupil variation was defined between 0.3 and 1.6 Hz in this study. In the previous study [9], when positive or neutral stimuli were presented, S/C value showed below 1.0 and when negative stimulus was presented, it was reported that S/C value was above 1.0. In this study, the relationship between VAS score and S/C value was quantitatively evaluated.

4 Results

Figure 6 shows temporal changes in Participant A's pupil area. Figure 7 shows the power spectrum of Participant A's pupil area variation. Figure 8 shows the relationship between VAS score and S/C value in 2D image. Figure 9 shows the relationship between VAS score and S/C value in partial 3D image. The colleration coefficients between VAS score and S/C value in Figs. 8 and 9 were 0.116, −0.114, respectively, which were unexpectedly low. Participant E's pupil area was not able to be recorded while Image *d* was presented, due to excessive noises.
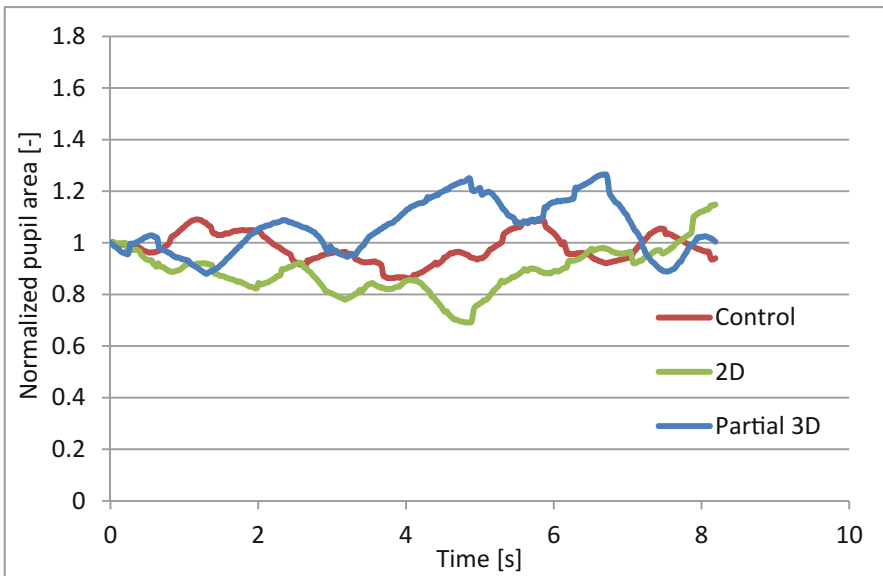
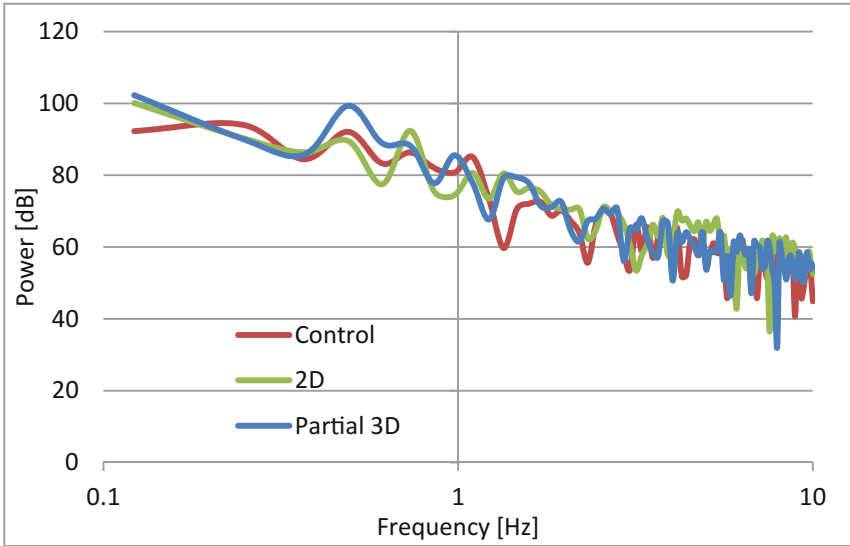**Fig. 6.** An example of temporal change in Participant A's pupil area (Color figure online)

Fig. 7. Power spectrum of (Participant A) pupil area variation (Color figure online)
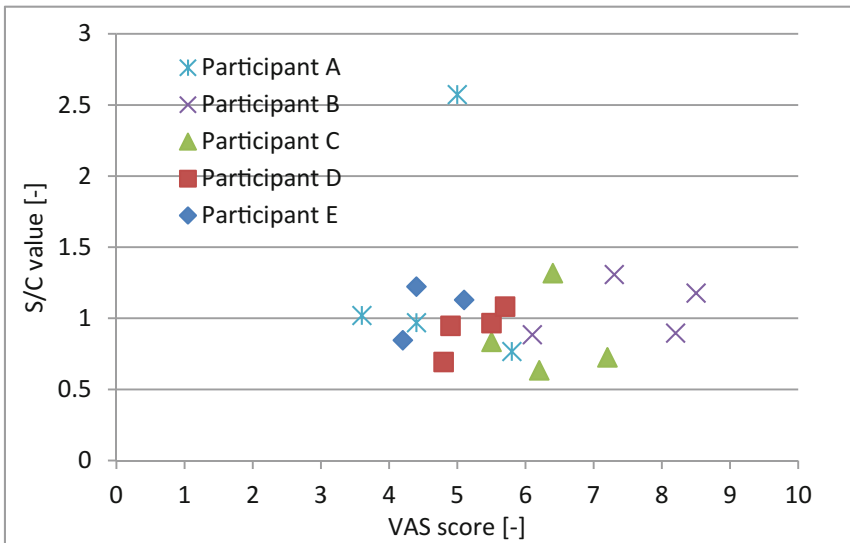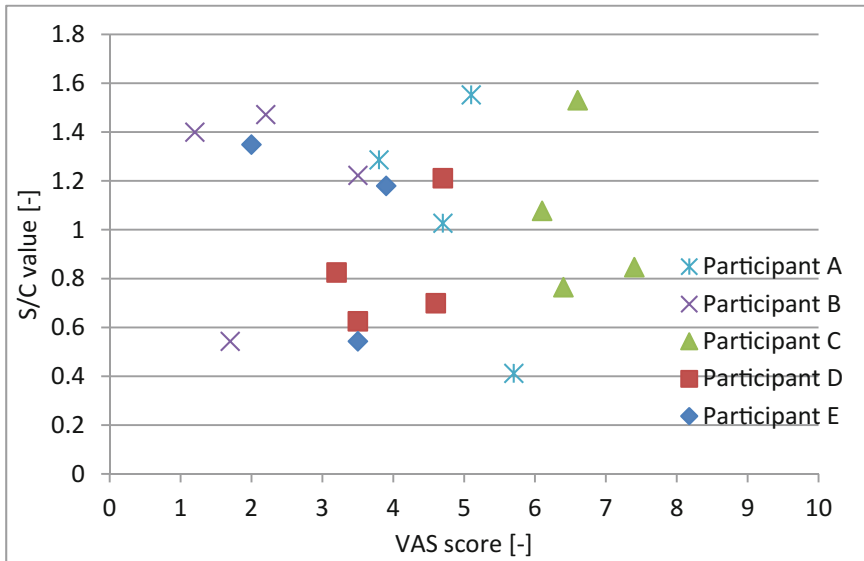

Fig. 8. Relationship between VAS scores and S/C values for each participant when participants monitored in 2D images
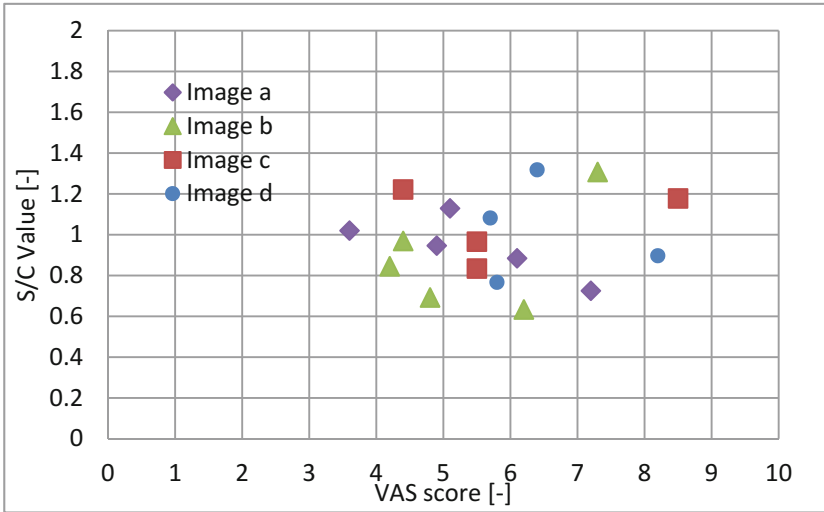
Fig. 9. Relationship between VAS scores and S/C values for each participant in partial 3D images

5   Discussion

Figure 8 revealed that most S/C values were distributed near 1.0 except for the participant A, where a high S/C value (= 2.573) was observed. The result can be interpreted that the emotion was hardly varied by showing neutral images. It was also observed that S/C values were not correlated with VAS scores when 2D neutral images were shown. These unexpected results were possively due to the potential polarization caused by glasses-free 3D monitor. Figures 8 and 9 indicated that some VAS scores in partial 3D images were lower than those in 2D images and S/C values did not seem to relate to VAS scores. Thus, we concluded that impression induced by partial 3D images may affect both VAS scores and S/C values. When partial 3D images were presented, relatively high S/C values were obtained at low VAS score conditions and it was suggested that S/C value was likely to use as a candidate for discomfort measure. Reason for independency between S/C values and VAS scores when partial 3D neutral images were presented was that VAS scores did not correspond to the intended responses because VAS scores were obtained after finishing all the sessions, resulting in unfocused responses.

Figure 10 shows the relationship between VAS scores and S/C values (except for a high value outliers) of each image when 2D images were given. Figure 11 shows the relationship between VAS scores and S/C values of each image when partial 3D images were given. Comparing Figs. 10 and 11, S/C values in Image *a* and Image *c* tended to be above 1.0 and the values in Image *a* and Image *c* tended to be below 1.0 in Fig. 11. Thus, it was concluded that the compartments given by partial 3D conversion may affect their feelings.

Fig. 10. Relationship between VAS scores and S/C values for each 2D images
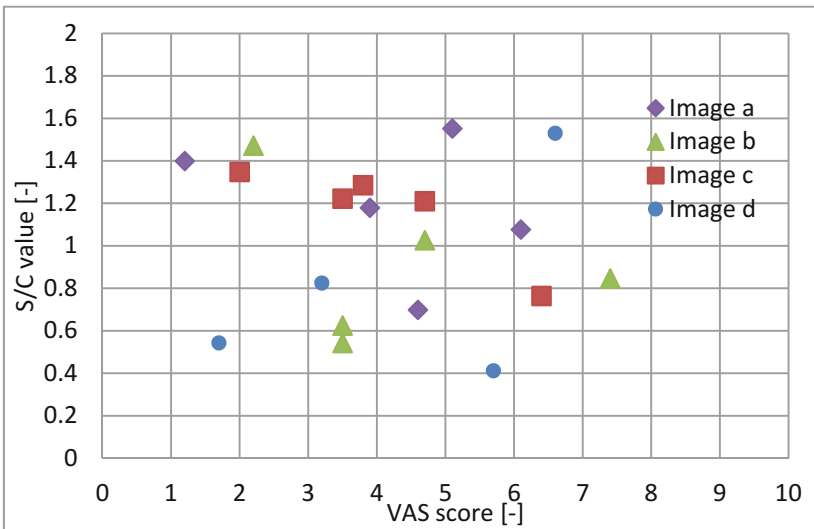
Fig. 11. Relationship between VAS scores and S/C values for each partial 3D images

In this study, whether or not participants made a specific impression, irrespective of comfort or discomfort, in neutral images was not clarified. It was possible that participants happened to gaze at the partial 3D image, which made such specific impression although selected images were designed to be free from impression of

comfort or discomfort for all participants. Thus, it is possible that the responses obtained by VAS was not matched with the intended responses associated with changes in pupil sizes.

6   Conclusion

In this study, we examined whether or not the systems to estimate the degree of dis comfort based on the previous study [9] was valid at partial 3D image presentations. Moreover, we discussed the effectiveness of the systems while guiding visual attention by partial 3D images. As a result, VAS scores did not have correlation with S/C values. Therefore, it was not clear whether or not the systems were valid at partial 3D images at this point.

The goal of our study was to investigate usefulness of the visual attention guidance by partial 3D image. The results cannot judge whether the negative emotion was due to the visual attention guidance. Thus, we need to evaluate the effectiveness of the systems to the partial 3D with modification of strategy to obtain VAS scores more clearly to show their impressions for each image.

References

1. Hata, H., Koike, H., Sato, Y.: Visual attention guidance using image resolution control. 18th Inf. Process. Soc. Japan Interact. 57–64 (2014)
2. Veas, E.E., Mendez, E., Feiner, S.K., Schmalstieg, D.: Directing attention and influencing memory with visual saliency modulation. In: CHI 2011, pp. 1471–1480 (2011)
3. McCay-Peet, L., Lalmas, M., Navalpakkam, V.: On saliency, affect and focused attention. In: CHI 2012, pp. 541–550 (2012)
4. Hagiwara, A., Sugimoto, A., Kawamoto, K.: Saliency-Based Image Processing for Guiding Visual Attention. Information Processing Society of Japan SIG Technical Report, vol. 2011-CVIM-177(5), pp. 1–8 (2011)
5. Kim, S., Takahashi, M., Watanabe, K., Kawai, T.: The effects of functional binocular disparity on route memory in stereoscopic images. In: IS&T International Symposium on Electronic Imaging, vol. 2016, no. 5, pp. 1–6. Society for Imaging Science and Technology (2016)
6. Koido, Y., Kawai, T.: Cognitive characteristics of partial stereoscopic image conversion. Papers of Japan Ergonomics Society Proceeding In: The 52th Conference of Japan Ergonomics Society, pp. 252–253. Japan Ergonomic Society (2011)
7. Kandachi, H., Koido, Y., Kim, S., Morikawa, H., Mitsuya, R., Kawai, T., Watanabe, K.: Partially converted stereoscopic images and the effects on attention and memory. Ergon. Jpn. Ergon. Soc. **49**(Supplement), 34–35 (2013)
8. Matsunaga, K.: Psychology of the Pupil Movement. pp. 7–30. Nakanishiya publishing (1990)

9. Kawai, S., Takano, H., Nakamura, K.: Development of Emotional Estimation System by Pupil Diameter Variation. The Institute of Electronics, Information and Communication Engineers Technical Report MBE, ME and Bio Cybernetics, 112(3), pp. 7–11 (2012)
10. Lowenstein, O., Loewenfeld, I.E.: The pupil. In: Davson, H. (ed.) The Eye, vol. 3, pp. 255–337. Academic Press, New York (1969)
11. Kawai, S., Takano, H., Nakamura, K.: Emotional Estimation by Pupil Diameter Variation. 11th Forum Inf. Technol. **11**(3), 421–422 (2012)
12. Hess, E.H.: Attitude and Pupil Size. Sci. Amerian **212**(4), 46–54 (1965)
13. Murakawa, S., Nishina, D., Ueki, M., Yokota, M.: Relationships between the image features on a river landscape. Trans. AIJ J. Archit. Plan. Environ. Eng. **64**(524), 53–60 (1999)
14. Takahashi, K., Tsukahara, M., Toyama, K., Hisada, M., Tamura, H.: Neural Networks and Biological Control, pp. 224–233. Asakura shoten, Tokyo (1976)
15. Lang, P.J., Bradley, M.M., Cuthbert, B.N.: International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-8, University of Florida (2008)
16. Yamakawa, Y., Takano, H., Nakamura, K.: Estimation of trait anxiety using pupil diameter variation during visual stimuli. Proc. 13th Forum Inf. Technol. **13**(2), 373–374 (2014)
17. Nakamori, S., Mizutani, N., Yamanaka, T.: Changes in pupil size to visually liking for pictures of faces. Int. J. Affect. Eng. **10**(3), 321–326 (2011)
18. Kondo, Y., Nishimura, Y., Ishii, H., Shimoda, H., Yoshikawa, H.: A study on an objective examination method of eye strain by using eye-sensing display. Hum. Interface Symp. **2**, 643–648 (2006)