Taylor & Francis
Taylor & Francis Group

Ergonomic evaluation of stereoscopic contents for a museum exhibition

N. Abe[a*], K. Ohta[b], T. Kawai[a], K. Ando[a], T. Kakinuma[b], K. Fujita[c] and N. Kudo[c]

[a]*Faculty of Science and Engineering, Waseda University, Tokyo, Japan;* [b]*Quality eXperience Design Co. Ltd., Tokyo, Japan;* [c]*Tokyo Broadcasting System, Inc., Tokyo, Japan*

This research entailed the production of stereoscopic 3D (S3D) contents using 2D-to-S3D conversion for exhibition at a museum and subjective evaluation. Hybrid production combining S3D images of existing live-action videos using the 2D-to-S3D conversion technology and computer graphic ones created via stereo rendering was conducted. Design and control of the chronological analysis of the parallactic angle was conducted on the produced contents, using binocular information as well as subjective evaluations, with the intent of conducting an investigation on the characteristics of such contents from the perspectives of the producers and viewers. An investigation was also conducted on the effects of the viewing position on the impressions of the S3D images.

Keywords: S3D contents; 2D-to-S3D conversion; parallactic angle; museum exhibition; subjective evaluation

1. Introduction

The number of cinemas installing stereoscopic 3D (S3D) systems is increasing rapidly, with a boost from the trend toward S3D movies in recent years. The development of S3D display devices for home use is thriving, and home electronics manufacturers are offering S3D-compatible television sets and S3D-compatible Blu-ray disc players. S3D broadcasts intended for general households have been launched in Japan, primarily by pay television channels and satellite broadcasting services. Compared with the past, the opportunities for people to view S3D contents have certainly increased. Researches on people's impressions of S3D contents have also been conducted, with some reporting positive evaluation ratings given by the viewers to image contents represented in S3D [1].

Issues remain, however, concerning some parties' reservations regarding safety in the production and application of S3D contents, which continue to increase. S3D images are considered to have a greater burden than 2D images due to the mismatch of the visual system with the contents. Accordingly, international standardization pertaining to the biomedical effects of S3D images is being carried forward [2]. '3DC Safety Guidelines' have also been established by a consortium of industry organizations in Japan [3]. The parallactic angle is often used as an index to provide a quantitative representation of the mismatch of the visual system with the contents. The parallactic angle is a representation of the gap between the convergence angles for viewing a screen and viewing an S3D image (Figure 1).

Inoue *et al*. took measurements of the eye movements when S3D images with varying parallactic angles were presented. They reported that the burden on the visual system remains low if the positive and negative parallactic angles are less than ± 1 degree [4]. The aforementioned '3DC Safety Guidelines' also recommend that the parallactic angles be kept within ± 1. Furthermore, Hanazato *et al*. indicated that the rapid changes in the parallactic angles between the frames contribute to the difficulty in viewing [5]. Producers of S3D contents are required to consider the problems peculiar to S3D images so they could produce contents that are safe and comfortable to view.

With a proper viewing environment for S3D contents almost in place, the stable supply of S3D contents that are safe and comfortable for viewers is required. There are issues, however, that still need to be resolved in the supply aspect of S3D contents. There are more cost and technical issues relating to the projection, reproduction, and editing of S3D images compared with 2D images. The supply of S3D contents is not sufficient even at present. 2D-to-S3D conversion, which converts 2D contents into S3D contents by adding binocular stereoscopic information, is drawing attention.

2. 2D-to-S3D conversion

2D-to-S3D conversion is a generic term referring to the technology used to convert 2D images into S3D images, through image processing conducted automatically or
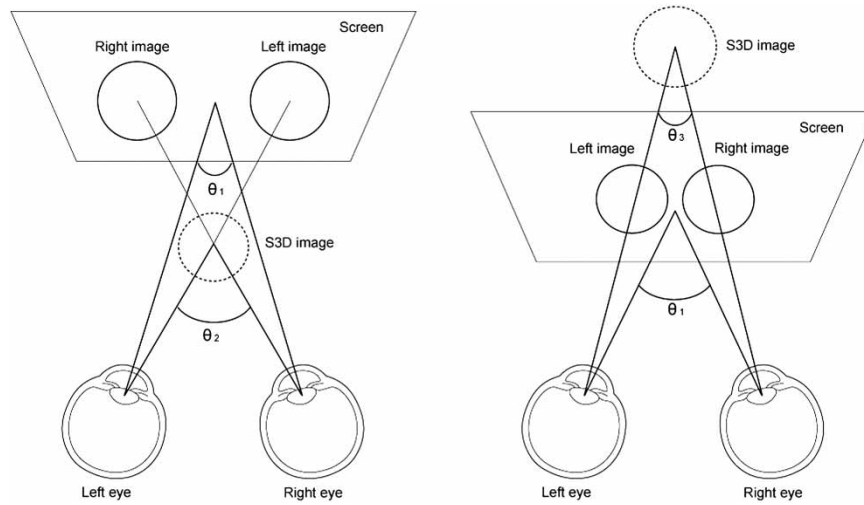
N. Abe et al.

Figure 1.   Parallactic angle: negative (θ^1–θ^2), positive (θ^1–θ^3).
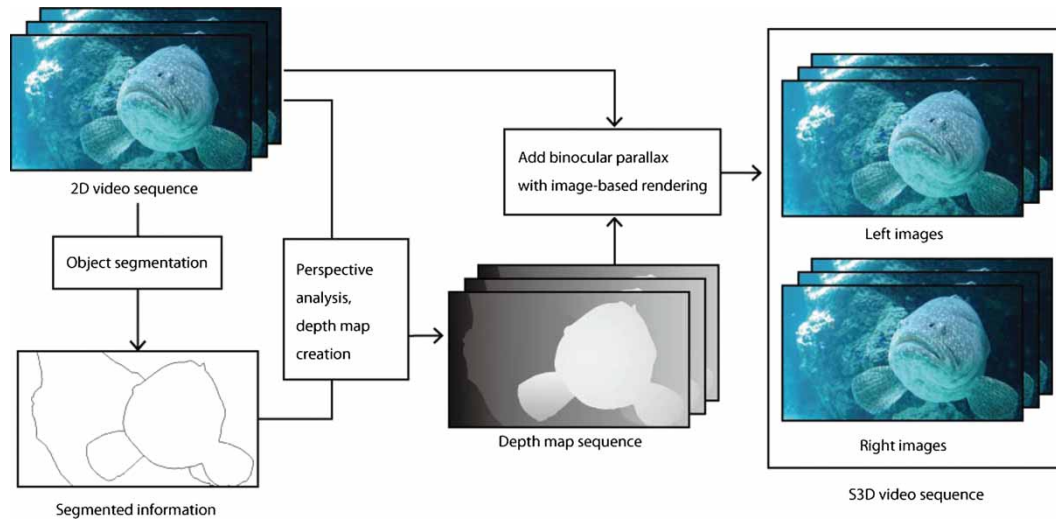

Figure 2.   Workflow of 2D-to-S3D conversion.

manually. This is a term that is widely used in recent years in relation to the production of S3D contents [6]. Spatial images have always been perceived by humans, using various types of information as cues. S3D information providing assistance for depth perception is categorized into two major groups: monocular information and binocular information. Monocular information is also recorded in 2D images, making it possible for the viewers to perceive depth within a projected image. S3D images captured with stereo cameras contain 'binocular parallax,' a type of binocular information, making it possible to offer richer spatial expressions. 2D-to-S3D conversion involves the addition of 'binocular parallax,' estimated from 2D images, to convert 2D images into S3D ones (Figure 2). As 2D-to-S3D conversion for creating S3D images is performed after capturing the images, it is considered possible to utilize the know-how accumulated in ordinary 2D content production, with lesser restrictions

for capturing images and producing visual expressions. The characteristics of the conversion technology also offer the advantage of being able to convert existing 2D contents into S3D ones. It is possible, therefore, to convert established 2D images into S3D ones by using the 2D-to-S3D technology, to supply such images as new S3D contents.

There is also a problem relating to a discrepancy that often occurs with monocular information, as binocular information is estimated from 2D images and then added on during 2D-to-S3D conversion. There are many issues, such as supplementation of the areas masked in the foreground, because information ordinarily obtained by capturing images using two cameras must be estimated from images that have been captured using only one camera. Few researches have been conducted to determine if any difference exists between contents that have been produced using 2D-to-S3D conversion and those that have been captured

and generated using cameras from the left and right perspectives. There has been no clear understanding, therefore, of such potential difference.

3. The 'Golden Capital of Sican' exhibition

The 'Golden Capital of Sican' exhibition was held at the National Museum of Nature and Science in Tokyo, Japan from July 14 to October 12, 2009. Over 200,000 people visited the exhibition [7]. The exhibition featured an S3D theater installed to provide viewing of S3D contents concerning the relics of Sican. The S3D contents consisted of a video recorded during the excavation survey conducted over several decades as well as descriptions of the relics that have been reproduced using computer graphics (CG). To provide the visitors with an abundantly realistic experience, S3D contents were decided to be used. This author and his associates were in charge of producing S3D contents for the Sican exhibition.

4. Research purpose

The research was aimed at producing contents that are safe and comfortable to view, by conducting a chronological analysis of S3D contents using a parallactic angle as an evaluation index. Furthermore, the effects of producing S3D contents by combining the 2D-to-S3D conversion technology, which is expected to be used at many content production sites in the future, with stereoscopic image capturing were investigated. Hybrid production, which combines S3D images of existing live-action videos using the 2D-to-S3D conversion technology, and CG ones created via stereo rendering, was conducted. The workflow of the content production is shown in Figure 3. Contents that are safe and comfortable to view were produced using quantitative standards such as the parallactic angle as indices, and a certain degree of reproducibility was maintained. Furthermore, design and control of the chronological analysis of the parallactic angle were conducted on the produced contents, using binocular 3D information as well as subjective evaluations, with the intent of conducting an investigation on the characteristics of such contents from the perspectives of the producers and viewers. Depending on the viewing position of the reproduced images, geometric distortions appear with S3D images, resulting in varied spatial effects [8]. An investigation was also conducted on the effect of the viewing position on the impressions of S3D images.
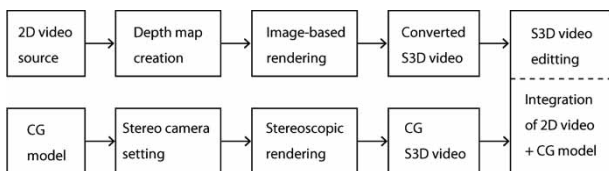

Figure 3. Workflow of the S3D content production for the Sican exhibition.

5. Production of S3D contents, and design and control of the parallactic angle

Two high-definition DLP projector (Panasonic PT-DZ6700, 1920 × 1080 pixels) units were used for image projection at an S3D theater, with the left and right images projected using the linear polarization method on a 300-inch screen. The synchronized left and right images were delivered by a cinema server (Digital Zone, HVP-4004). The viewers wore polarizer glasses to view the S3D images. The theater consisted of about 100 seats, with a design that placed the frontmost row at 5.5 meters from the screen and the backmost row at 8.5 meters. The viewers were also allowed to stand behind the last row of seats for viewing. The projection environment is shown in Figure 4.

A characteristic of the contents that were projected at the S3D theater was that they consisted of actual images captured at the excavation site of the Sican relics acquired in the past, and of CG materials that depicted the restored scene of the relics, which were re-rendered, respectively, into S3D videos. 2D live-action materials of images actually taken onsite were converted into an S3D video using 2D-to-S3D conversion, which was carried out by verifying the shapes of the relics with the director and experts, with the aim of creating S3D expressions that were natural and that had no discrepancies. The materials that were used for the CG video were produced by placing two virtual cameras into the model data and then rendering left and right videos. During the editing, the parallactic angle of the materials derived from live-action images and that had been processed via 2D-to-S3D conversion, and from the CG materials, was calculated to coordinate the balance between the two types of materials. The presentation time at the S3D theater, including the title and the end roll, was about 10 minutes.

The evaluation method for S3D contents proposed by Kishi *et al*. [9] was used for the analysis of the parallactic angle. The projection environment was set with a screen size of 300 inches and an observation distance of 5.5 meters to obtain the parallactic angle for each second. The farther the seat position was from the screen, the smaller the parallactic angle became. The parallactic angle
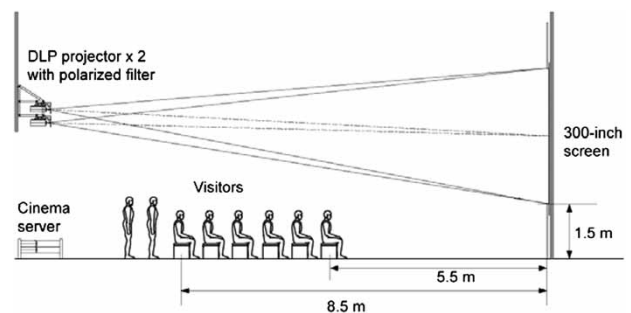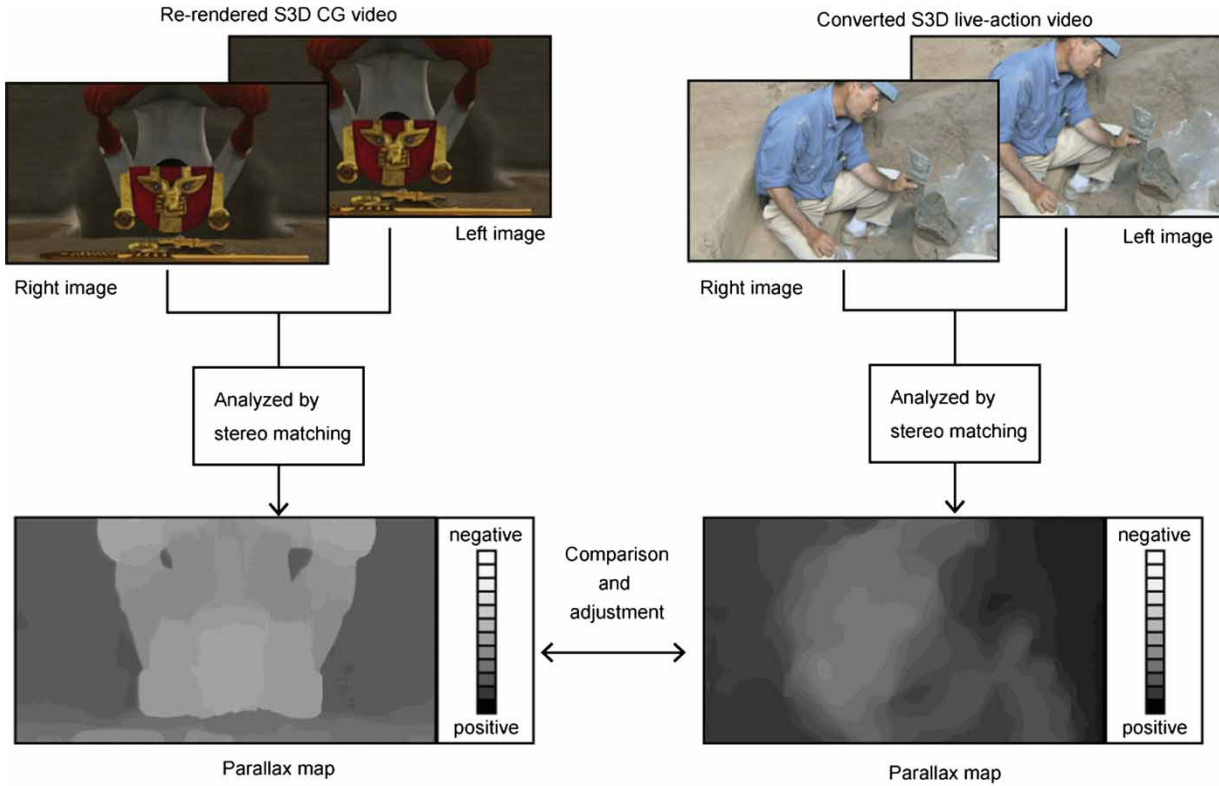

Figure 4. Projection environment.
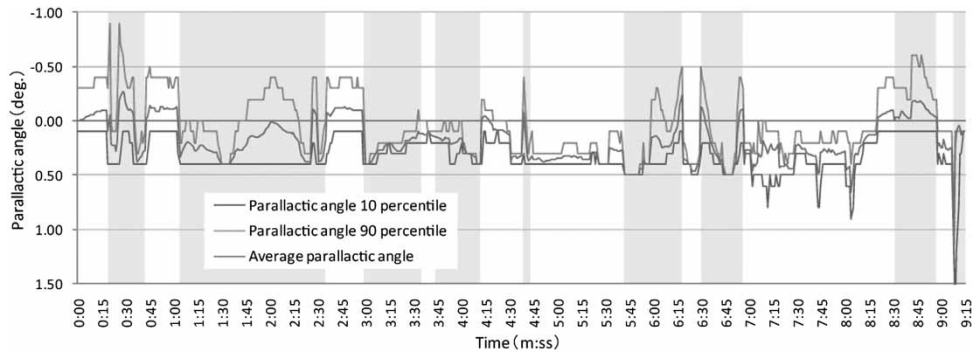
Figure 5. Analyzed parallax map.

Figure 6. Change in parallactic angle within the contents. The gray area is a converted S3D live-action video, and the white area is the re-rendered S3D CG video.

was derived by calculating the distance between the corresponding points of the left and right images (parallactic distance) for each pixel during the image processing by stereo matching. The parallactic distance was converted into the parallactic angle from the set projection environment. The analysis results of the S3D contents from the Sican exhibition are shown in Figure 5. The upper line represents the original image, and the lower line represents the analyzed distance image referred to as the 'depth map.' The distance distribution is represented, with the lighter segments inside the image representing the front portion and the darker segments representing the farther side. It is possible to calculate the parallactic angle from this depth map by entering the details of the projection environment. Design

and control of the stereoscopic effect were performed considering compatibility with the existing guidelines along with variation in the quantity of binocular information that corresponds to the progress of the story in the content, as well as the connection between the frames containing live-action materials and those containing CG materials, using the calculated parallactic angle. The chronological change in the parallactic angle of the contents is shown in Figure 6.

6. Subjective evaluation of the S3D contents

6.1. Method

Subjective evaluation was conducted by having the visitors at the S3D theater fill out a questionnaire after viewing the

contents. The survey was prepared with questions pertaining to the viewers' sex, age group, seat location, attributes concerning their S3D viewing experience, perceptions of the depth and naturalness of the contents, and experience of eye fatigue, using the rating scale method. Responses were requested using a five-level rating scale for the perceptions of the depth and naturalness of the S3D contents, and a four-level rating scale for eye fatigue. The seating locations for viewing were divided into four sections along the front and back directions of the front rows, middle rows, back rows, and standing, as well as three sections in the left and right directions of the right side, center, and left side, comprising a total of 12 blocks. The seating locations for viewing were then recorded (Figure 7).

6.2. Results

The number of survey respondents and the correlation with the seating locations are shown in Table 1. The number of valid responses was 447 (166 men and 281 women).

The results relating to the perceptions of the depth and naturalness of the S3D live-action video and the S3D CG video are shown in Figure 8. The average rating for the perception of the depth of the content was 3.5 for the S3D
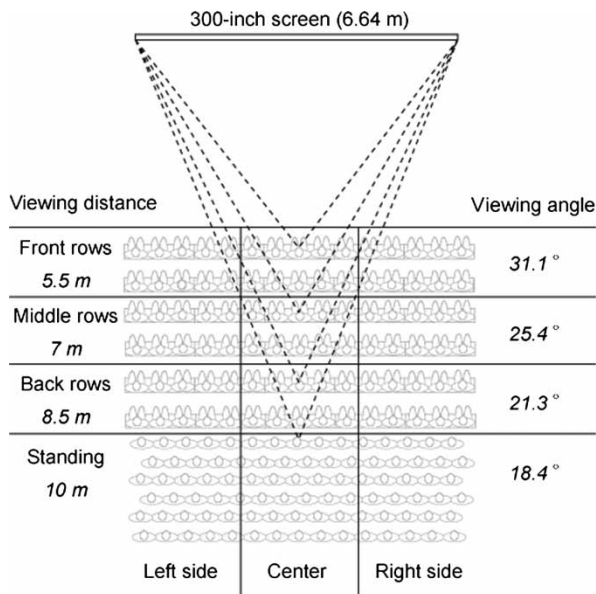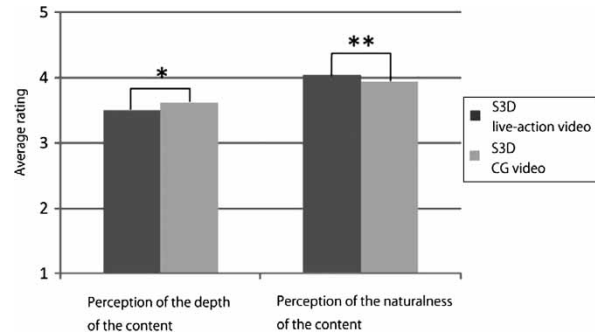
Figure 8. Evaluation results for the depth and naturalness of the S3D contents (**$p < 0.01$, *$p < 0.05$).

live-action video and 3.62 for the S3D CG video. The evaluation results of the S3D CG video were thus better. The average rating for the perception of naturalness of the content was 4.04 for the S3D live-action video and 3.95 for the S3D CG video. The evaluation results of the S3D live-action video were thus higher. In the t test on the respective items, significant differences were shown in both the perceptions of the depth of the content ($t(446) = 2.23, p < 0.05$) and of the naturalness of the content ($t(446) = -5.12, p < 0.01$). This suggests that the impressions differ according to the method used for S3D conversion.

The evaluation results for eye fatigue are shown in Figure 9. A total of 279 persons, representing more than 60% of the viewers, gave favorable ratings by indicating that their eyes did not get tired from viewing the contents, resulting in an average rating of 2.23.

To evaluate the impact of eye fatigue on the perceptions of the depth and naturalness of the contents, analysis of variance (ANOVA) was conducted on the extent of eye fatigue with respect to the perceptions of the depth and naturalness of the S3D live-action video and the S3D CG video. A significant difference ($F(3443) = 4.05, p < 0.01$) in the perception of the naturalness of the S3D live-action video was revealed as a result. Furthermore, a significant difference trend ($F(3443) = 2.20, p < 0.1$) was revealed for the perception of the naturalness of the S3D CG video. Upon obtaining this result, Scheffe's multiple comparison was conducted, and a trend was recognized, where those who
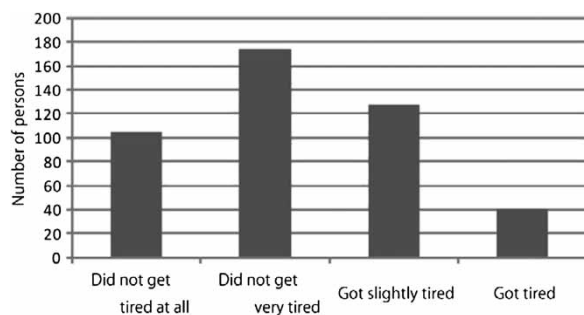
Figure 7. Seating locations for viewing.

Table 1. Seating locations of the survey respondents (unit: persons).

	Left side	Center	Right side	Total
Front rows	25	67	19	111
Middle rows	26	72	46	144
Back rows	35	96	40	171
Standing	4	14	3	21
Total	90	249	108	447

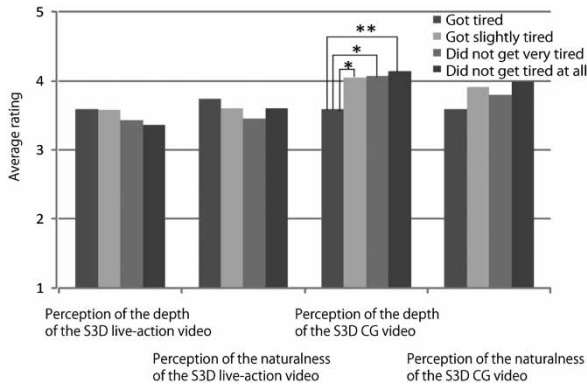Figure 9. Tabulation of the evaluation results for eye fatigue.

Figure 10.   Evaluation results for eye fatigue ($**p < 0.01$, $*p < 0.05$).

responded 'got slightly tired' ($p < 0.05$), 'did not get very tired' ($p < 0.05$), and 'did not get tired at all' ($p < 0.01$) evaluated the naturalness of the contents more favorably than those who responded 'got tired' (Figure 10). The fact that this trend was not significant with the perception of the naturalness of the S3D CG video implied that the perception of the naturalness of 2D-to-S3D conversion may potentially be connected to eye fatigue as this trend was not seen with the S3D CG video.

Furthermore, to evaluate the impact of the difference in seating location on the S3D content viewing experience, ANOVA was conducted on the seating location with the perceptions of the depth and naturalness of the S3D live-action video and the S3D CG video. The conditions for the standing viewers were excluded in this instance because the number of respondents was small. The results revealed no significant difference due to the impact of the seating location.

6.3.   Speculations

Subjective evaluation was conducted in this research by having the visitors fill out a questionnaire on their perceptions of the depth and naturalness of the S3D contents. The S3D CG video was rated higher for depth, and the S3D live-action video was rated higher for naturalness. This suggested that even when the S3D videos were produced using the parallactic angle of the same standard, a certain trend emerged with regard to the difference in impressions according to the method used for S3D conversion. Furthermore, no significant difference was found with regard to the impact of the seating location. For these reasons, it was suggested that the seating location has no impact on the stereoscopic effect or perception of naturalness for the parallactic-angle and image projection conditions of the S3D content, with which production and evaluation were conducted. It is speculated that the conditions that were used for this research were within the range in which humans can perceive the stereoscopic effect.

In terms of the impact of eye fatigue on the perceptions of the depth and naturalness of the S3D contents, naturalness

received lower evaluated ratings for the content processed with 2D-to-S3D conversion by those who responded 'got slightly tired,' 'did not get very tired,' and 'did not get tired at all.' As mentioned earlier, there is a discrepancy problem with the monocular information frequently occurring with images processed with 2D-to-S3D conversion. Even though there may be differences among individuals, it can be speculated that for 2D live-action videos converted into S3D videos via 2D-to-S3D conversion, the discrepancy with the monocular information can lead to a positive evaluation of naturalness, and this discrepancy among such S3D images may be connected with eye fatigue.

7.   Summary

This research entailed the production of S3D contents using 2D-to-S3D conversion for exhibition at a museum, and subjective evaluation.

- Proactive evaluation ratings were obtained for the depth and naturalness of the S3D contents, as a result of designing and controlling the binocular information using a parallactic angle for the 2D-to-S3D-converted videos and the CG ones created via stereo rendering.
- A difference in impression between 2D-to-S3D conversion used for the live-action video and stereo rendering used for the CG video was also recognized in terms of the perceptions of depth and naturalness.
- A certain trend was recognized with regard to the correlation between the perception of naturalness and eye fatigue for the images processed with 2D-to-S3D conversion.
- No impact of the seating location was revealed, however, by the evaluation results obtained in this research.

  Large theaters with seats arranged on inclines as well as contents intended for venues other than a museum will be considered for the future study on 2D-to-S3D conversion and viewer experience.

References

[1] M. Pölönen, M. Salmimaa, V. Aaltonen, J. Häkkinen and J. Takatalo, J. SID **17** (5), 459 (2009).

[2] International Standards for Business, Government and Society, *ISO/IWA 3: Image Safety − Reducing the Incidence of Undesirable Biomedical Effects Caused by Visual Image Sequences* (2005).

[3] Safety Guidelines Sub-Committee of 3D Consortium, *3D Consortium Safety Guidelines for Popularization of Human-Friendly 3D* (revised on April 20, 2010, compliant with ISO/IWA) Japan (2010).

[4] T. Inoue, K. Noro, T. Iwasaki and H. Ohzu, J. Inst. TV Eng. Jpn., **48** (10), 1301 (1994).

[5] A. Hanazato, Y. Yamanoue, Y. Nojiri and F. Okano, Inst. Image Inform. TV Eng. J. **27** (23), 37 (2003).

[6] T.K.H. Morikawa, K. Ohta and N. Abe, *Foundation of Three-Dimensional Image Expressions* (Ohmsha, Tokyo, Japan, 2010).

[7] *The 'Golden Capital of Sican' exhibition.* http://www.kahaku.go.jp/exhibitions/ueno/special/2009/sican/index.html.

[8] H. Yamanoue, M. Nagayama, M. Bitou, J. Tanada, T. Motoki, T. Mituhashi and M. Hatori, IEICE Transact. D-II, **J80-D-2** (9), 2522 (1997).

[9] S. Kishi, T. Yamazoe, T. Shibata, T. Kawai, T. Inoue, Y. Sakaguchi, K. Okabe and Y. Kuno, J. Inst. TV Eng. Jpn. **60** (6), 934 (2006).