The Twenty-Eighth International Symposium on Artificial Life and Robotics 2023 (AROB 28th 2023),
The Eighth International Symposium on BioComplexity 2023 (ISBC 8th 2023),
The Sixth International Symposium on Swarm Behavior and Bio-Inspired Robotics 2023 (SWARM 6th 2023)
B-Con Plaza, Beppu, Japan and ONLINE, January 25-27, 2023

Comparison of Head-Mounted Display versus Four-sided Screen Displays on Passive Viewing Experience for Panoramic Video

Hirotoshi Asano[1] , Junko Ichino[2] , Yusuke Tokunaga[3] and Masayuki Wada[3]

[1]Department of Computer Science, Kogakuin University, Tokyo, Japan
(Tel : +81-3-3340-2704; E-mail: hirotoshi@cc.kogakuin.ac.jp)
[2]Department of Information Systems, Tokyo City University, Kanagawa, Japan
(Tel : +81-45-910-2570; E-mail: ichino@tcu.ac.jp)
[3]Department of Electronics and Information Engineering, Kagawa University, Kagawa, Japan
(Tel : +81-3-3340-2704; E-mail: hirotoshi@cc.kogakuin.ac.jp)

Abstract: This study investigates how user experience is affected during the viewing of panoramic videos in both virtual reality (VR) environments and non-VR environments. In particular, we investigate the differences between the effects of two viewing environments on the physiological and psychological, memory aspects of the users. We conducted experiments and compared three display-type conditions: a head-mounted display (HMD), four-sided screen display (SCREEN), and SCREEN with a narrow field of view when using goggles (SCREEN-GOGGLE). We found that the SCREEN users remembered the content of the videos more accurately than the HMD users. These results are consistent with the nasal skin temperature results, suggesting that, compared to SCREEN users, HMD users had a lower mental load.

Keywords: Head-mounted display, Four-sided Screen displays, Passive panoramic viewing, Physiological and psychological measurements.

1. INTRODUCTION

In recent years, head-mounted displays (HMDs) have become a popular way to watch 360° videos, including VR and panoramic videos [2]. An HMD, which is a typical display device for virtual reality, is worn on the user's body. HMDs have been receiving attention in the fields of entertainment (e.g., games), medical care, and ICT- based education. HMDs are also popular for watching movies through passive cinematic VR (CVR). In addition, a cave automatic virtual environment (CAVE) [1], which is a technology for projecting images on a polyhedral screen, is also widely used in industry, education, psychological research, and the arts [5]. CAVE uses the wall and the surrounding space as a display device with multiple projectors. Therefore, it does not completely block the view of the real world, such as with an HMD. In other words, users can judge or act on an event or object using their own body as a scale of reference while maintaining their own bodies within their field of vision. In particular, projection mapping technology, which uses a projector to project images onto a space or object and provides various visual effects to layered images, has been widely used in exhibitions. As HMDs become more popular with the general public, more people will become interested in enjoying 360° videos using CAVE, utilizing projectors and walls. With this diversification of displays, users need to choose the display types according to their purposes. Content developers also need to understand the characteristics of each display for application to educational, medical, entertainment, and other fields. It is therefore important to evaluate the differences between HMD and CAVE displays to guide users and developers toward the appropriate technology. Many studies have been conducted on understanding the impact of different display types on users [6], [7], [23]. The main problem is to accurately assess the differences between these display types. We carefully reviewed the previous studies on display types and found that in the simple viewing of panoramic videos, the difference between the influence of HMD setups and CAVE setups on users is still not completely clear. J.Ichino [4] used five indicators to evaluate cognitive activities from both performance and process perspectives. On the other hand, this manuscript defines four new research questions for presence, memory, simulation sickness, and point of view and evaluate them using nine indicators. In addition, two types of 360°videos, documentary and narrative, were used to evaluate the effect of different videos on users. The next section provides an overview of relevant studies comparing HMDs and CAVEs using an evaluation index.

2. RELATED WORK

2.1 Comparison of HMDs and CAVEs based on evaluation index

2.1.1 Presence

Using a task of finding two words cards, Kim et al. [8] compared three setups, that is, an HMD, a 6-screen screen similar to CAVE, and a desktop display. It was found that the sense of presence of the six-screen CAVE-like setup significantly increased. Tcha-Tokey et al. [9] compared an HMD setup with a four-screen CAVE-like setup. The user conducted a task while experiencing Egyptian history. The results showed that the sense of presence in the CAVE setup was significantly higher than that in the HMD setup. Other studies have also supported these results. In Juan et al.'s study [10], a task was conducted, during which the floor upon which the user was standing suddenly dropped. It was reported that the users felt more anxiety and presence in the CAVE setup than in the HMD setup. In a study by Krijin et al. [11], patients with a fear of heights carried out tasks such as climbing stairs at a high altitude and releasing their hands from a handrail. It was reported that the presence in the CAVE setup was higher than that in the HMD setup. However, the study by He et al. [13] reported different results. In the study by He et al., users watched panoramic videos and compared the HMD and screen setups. They reported that the sense of presence in the HMD setup was higher than that in the screen setup. As a common feature of these studies, the feeling of presence in the HMD setup is higher than that in the CAVE setup. There are two reasons why their conclusions differ. One is based on the type of task. The tasks assigned in Kim et al.'s and Tcha-Tokey et al.'s studies [9] are interactive, in which the users move within the VR space. By contrast, in He et al.'s study the users were assigned the passive task of watching a video. This suggests that the type of task influences the evaluation of the display type. The other reason is the difference in the vision applied. Juan et al. pointed out that the difference in visibility may have affected the results of the display-type evaluation. Users in a CAVE environment can interact with their own body as a reference of scale, while keeping their own body in perspective. By contrast, a user wearing an HMD will have a restricted field of view. This suggests that differences in viewing angles can affect the evaluation of the display type. Therefore, we need to investigate the impact of the user's restricted field of view on the evaluation of the sense of presence when simply viewing a passive panoramic video.

2.1.2 Memory

Numerous previous studies, including those by Luong et al. [17] and Chowdhury et al. [18]. have evaluated user memory. MacQuarries et al. [5] compared four different setups, that is, an HMD, SurroundVideo + (SV), a technology for creating panoramic 360°3D images, and television (16:9). They re- ported that the HMD setup resulted in fewer memories than the CAVE setup. Polcar et al. [14] also evaluated the effects on user memory. The HMD setup demonstrated less knowledge than a screen or desktop setup. In addition, Philpot et al. [16] compared an HMD setup with a CAVE-type display setup. There were no differences between the setups in their results or in the memories of the users. The results of Rizzo et al.'s work [15], i.e., also could not be confirmed same result. MacQuarries et al. have an interesting take on these differences in results. They report that narrative content is a good choice, whereas images of horror are not. Fonseca et al. [12] also argued that watching narrative videos can enhance the emotions of the users. Therefore, documentary and narrative content needs to be selected instead of horror to correctly evaluate the display type during the viewing of passive panoramic video.

2.1.3 Simulation sickness

Several previous studies have examined simulation sickness. Weidner et al. [6] evaluated the relationship between non-VR (2D and 3D) and VR (HMD) setups and the physiological response of the user, the occurrence of simulation sickness, and motion performance in lane-change tasks. They reported that HMD setups significantly increase the amount of simulation sickness compared to non-VR 3D setups. Sharples et al. [19] also reported that an HMD setup causes simulation sickness. Eudave et al. [20] and Ahmaniemi et al. [21] reported that an HMD setup affects the physiological parameters such as the heart rate, electrodermal activity, and galvanic skin response. Chowdhury et al. [18] evaluated the sense of presence and reported that an HMD setup improves the information recall (MSQ score) but increases the SSQ. Kim et al. also reported that an HMD setup causes simulation sickness, whereas a CAVE setup does not. By contrast, Polcar et al. [14] reported that simulation sickness occurs in both CAVE and HMD setups. A common result of these studies is that an HMD setup is prone to simulation sickness when a user is assigned the task of moving within a space. There have been few reports on whether differences occur between an HMD setup and a CAVE setup in the case of panoramic videos, which are less prone to simulation sickness.

2.1.4 Viewpoint

Petkova et al. [7] and Slater et al. [22] argued that in general VR research, "whether you are looking from a first-person perspective or a third-person perspective" is important. Considering the function of the visual cortex, it is possible that the difference in viewpoints influenced Juan's experimental results. However, few studies have evaluated the differences in viewpoints depending on the display type in the task of viewing panoramic videos. Evaluating the differences in perspectives by display type may provide a more detailed understanding of the impact on presence.

2.2 Research questions for passive panoramic video viewing experience

This paper investigates in detail how user experience is affected during the passive viewing of panoramic videos in two different types of environments, that is, VR and non-VR environments, based on the problem-related task. We set up four research questions (RQs) for viewing panoramic videos based on related studies.

· Do field-of-view limitations affect the evaluation of display types in terms of presence? (RQ1)
· Does content affect the display type ratings in terms of user memory? (RQ2)
· Does a passive panoramic viewing experience affect the display type in terms of user simulation sickness? (RQ3)
· Does a passive panoramic viewing experience affect the display type in terms of the points of view of the users? (RQ4)

3. MATERIALS AND METHODS

3.1 Experiment system

We focused on three types of viewing environments, i.e., an experimental system related to the viewing of videos using an HMD (hereafter, HMD), a four-sided screen display where the user wears goggles such that his or her field of vision is identical to that of an HMD (hereafter, SCREEN-GOGGLE), and a four-screen display with no restrictions on the user's field of view (hereafter, SCREEN). Fig. 1 shows the experiment system used for the HMD and SCREEN GOGGLE environments. In the case of the SCREEN-

Fig.1 (Upper) HMD and (lower) SCREEN-GOGGLE, which restricts the vision of the user by applying goggles.

GOGGLE system, the user does not wear goggles. This was the only difference between the two systems. First, we describe the HMD system. The system is composed of an HMD (HTC Vive with a resolution of 1080 × 1200, monocular only), a skin temperature sensor (Gram LT-USB1), a blood pressure monitor, and an electroencephalograph (g.tec medical engineering GmbH g.USBamp). The HMD allows the user to view the video across a 360° horizontal field of view through monoscopic vision. The user sits in a legless rotating chair with their eyes rested and open. The user then wears the HMD on his or her head, and infrared sensors are positioned in two spots toward the head of the HMD wearer so that they are recognized by the PC. The user is also fitted with electrodes for an electroencephalograph (EEG), electrodes to monitor the skin temperature on the nose and forehead, and electrodes on the right palm and index finger to monitor the pulse rate. The EEG electrodes were affixed in accord with the international 10-20 system at three points (sites Fz, Cz, and Pz), with a GND electrode (at Oz) and a reference electrode (on the left earlobe). The sampling frequency for the skin temperature monitor was 0.5 Hz; for the blood pressure monitor, 1 kHz; and for the EEG, 100 Hz. Next, we describe the SCREEN-GOGGLE and SCREEN systems. The measurement equipment installed, sampling frequencies, and electrode fitting methods were identical to those used in the experimental system for HMD viewing. However, to gauge the impact of the user's field of vision during HMD viewing, a set of goggles were prepared to restrict the user's field of field, similar to when wearing an HMD. Four screens were used, covering all surfaces, except the ceiling and floor. The four screens are 2.3 m high × 2.3 m wide.

3.2 Procedure

The user first watched a practice video during the "practice" session, and then responded to a Self-Assessment Manikin (SAM) [24] and a simulator sickness questionnaire (SSQ)[25] during the first "questionnaire" session. As the genre of video used during the practice session, the users watched a music video. The viewing time was approximately 90 s. The users were given 240 s to respond to the questionnaire. Next, during the "session viewing" section, the video was watched. The viewing media consisted of an HMD and four screens. However, in order to examine the viewing angle of the HMD and the four screens, three varieties of viewing setup were used: the HMD, a four-screen setup where users were equipped with goggles to restrict their field of vision, and a four-screen setup where no goggles were used. In addition, all videos, including the practice video, were viewable monoscopically in a horizontal 360° field of view, with a uniform vertical field of view of 90°. To examine the impact of the video genre, two types of videos were viewed, i.e., documentary and narrative. However, to avoid having each user see the same movie on different media, three types of documentary and narrative videos were prepared. The documentaries consisted of videos entitled "Syria," "Congo," and "Chen-nai." The narrative videos consisted of the titles "Rain Shine," "Pearl," and "Delivery". After viewing the video, a second "questionnaire" session was conducted in which the users responded to an SSQ, an SAM, an Igroup Presence Questionnaire (IPQ) [26][23], a first/third-person perspective questionnaire, and a comprehension test conducted through a computer monitor. The questionnaire response time was 240 s. One trial required a minimum of 1590 s and a maximum of 1650 s. Furthermore, we also changed the order in which the video content was viewed for each subject to account for counterbalancing. Each user was given six trials (three types of viewing media × two types of videos).

3.3 Subject requirements

The users consisted of 15 healthy men and women aged 21 to 25 years (average age of 23.1 years, with a standard deviation of 1.5 years, and vision with the naked eye or with a visual correction of 1.0). The experimental period was 6 h or more after subjects awoke for the day, with subjects receiving 8 h or more of sleep the day before. The experiment was conducted 2 h before mealtime, and the ingestion of medicine, smoking, alcohol intake, and excessive exercise were prohibited on the day of the activity. Furthermore, in accord with a review by Kagawa University's Research Ethics Committee, the step- by-step procedure followed by the experiment was described in detail to the users, and the experiment was conducted only after user agreement to cooperate was obtained. In addition, owing to the impact of the circadian rhythms, the experiment was conducted during the daytime (from 9 am to 1 pm). Taking into consideration the potential that the subjects might become acclimated to the experiment, a period of at least 3 days was allowed to elapse between experiment sessions.

4. EVALUATION MEASUREMENTS

4.1 Memory measurements

A comprehension test was used to evaluate the introspection of the user, wherein the user was given a questionnaire containing nine questions per movie regarding the content of the panoramic video. The correct answer rate, confidence factor, and answer time were used as evaluation indicators of memory. Preliminary experiments were ducted to reduce the variation in the difficulty level, and the difficulty level of the problem was adjusted.

4.2 Physiological measurements

4.2.1 Nasal skin temperature

As the skin temperature changes according to changes in blood flow, physiological changes are strikingly apparent in the nasal skin temperature (NST) [27]. When the sympathetic nerves are stimulated, the nasal skin temperature decreases, depending on the decrease in blood flow through the AVA. When sympathetic nerves are inhibited, NST increases owing to the increased blood flow through the AVA. Therefore, by using NST as an index, it is possible to evaluate the activity in the sympathetic nervous system. In general, NST decreases when people have a high mental load [28]. In Insko et al.'s study [29], the authors used the measurement of skin temperature to evaluate the sense of presence when using an HMD setup. However, they did not compare HMDs with CAVE. In addition, the user was moving in the VR environment. Thus, their experimental conditions differed from ours.

4.2.2 Plethysmogram

In the study by Meehan et al. [30], [31], the heart rate was measured to evaluate the HMD setup. A plethysmogram (PTG) includes heart rate information, and thus it is possible to calculate the heart rate variability index from the RR interval and the A-A interval, which is the equivalent interval between initial positive waves. In particular, after the baseline fluctuation component was removed from the pulse wave and the pulse wave component was extracted, the A-A intervals were interpolated. In addition, the low-frequency (LF) component power spectrum is extracted using an FFT analysis and used as an index of heart rate variability [28].

4.2.3 Electroencephalograms

In Linjia et al.'s study [32], the authors selected the measurement of electroencephalograms (EEGs) to evaluate the HMD and screen setup. Baka et al. [33] also used EEGs to compare the presence of virtual and physical environments. With EEGs, as the subject's depth of various activities, e.g., coma, sleep, rest, agitation, and awaking, increases, the EEG waves transition from high amplitudes and low speeds to low amplitudes and high speeds. Through a frequency analysis, EEGs are classified as being in the theta, alpha, beta, or gamma bands and are used as an index of one's psychological state, such as one's level of stimulation or emotion [34]. After a band-pass filter from 0.5 to 30 Hz is applied through the EEG, the power spectral density is classified into three bands: theta, alpha, and beta bands. Next, the baseline fluctuation component is removed from the PSD for each frequency component, and after the PSDs are interpolated, the contents of the theta, alpha, and beta bands are calculated. In addition, the content based on the percentage of the content in each band to the content in all bands from 0.5 Hz to 30 Hz is calculated and used as an index of activity of the central nervous system [35]–[38].

4.3 Psychological measurements

4.3.1 First/Third-person perspective questionnaire

The first-person/third-person perspective questionnaire evaluates the impact that the viewed media has on the user perspectives when the videos are viewed. The user was asked two questions in total, one question regarding a first-person perspective and one regarding a third-person perspective, to which the user responded on a 7-point scale.

4.3.2 SAM questionnaire

The SAM [24] is a questionnaire evaluating the impact of the viewed media and video on the user's pleasure, arousal, and dominance (whether or not the user has control). A questionnaire using images representing the three types of conditions was displayed on a computer monitor, and the users were asked to choose a score on a 9-point scale. The incremental difference in the questionnaire responses before and after the experiment was used as the evaluation index.

4.3.3 IPQ questionnaire

The IPQ [26] was administered after viewing the 360° video. The user responded to 14 questions regarding the media's sense of realism on a 7-point scale ranging from 0 to 6. The questions covered four areas: general presence (GP), spatial presence (SP), involvement (INV), and experienced realism (REAL). In general, GP is an index expressing how real the virtual experience feels, whereas SP is an index expressing the sense that one's body is present in the virtual space. INV is an index of the degree to which a user is charmed and fascinated while in the virtual space. REAL is an index that expresses how close the experience of being in the virtual space is to reality. An average score for each question assigned to a given index was determined, and the scores were used as evaluation indices.

4.3.4 SSQ questionnaire

The SSQ [25] is a questionnaire that evaluates the impact that the media viewed and its video has on video motion sickness of the user. The maximum scores for each evaluation index were as follows: total score, 239.36; nausea, 267.12; oculomotor, 212.24; and disorientation, 389.76. The minimum score for each evaluation index was zero.

5. RESULTS

5.1 Comprehension test

The results of the comprehension test based on the percentage of correct answers are shown in Fig. 2-A. A two-factor ANOVA revealed the main effect of the display type ($F(2, 28) = 4.927$, $p = 0.015$). There was a significant difference between HMD and SCREEN-GOGGLE ($p = 0.040$). The higher the percentage of correct answers, the more correctly the user understood the video. The results of the two-factor ANOVA showed that the main effect was recognized. As a result of multiple comparisons using Bonferroni's method, HMD was understood to have a

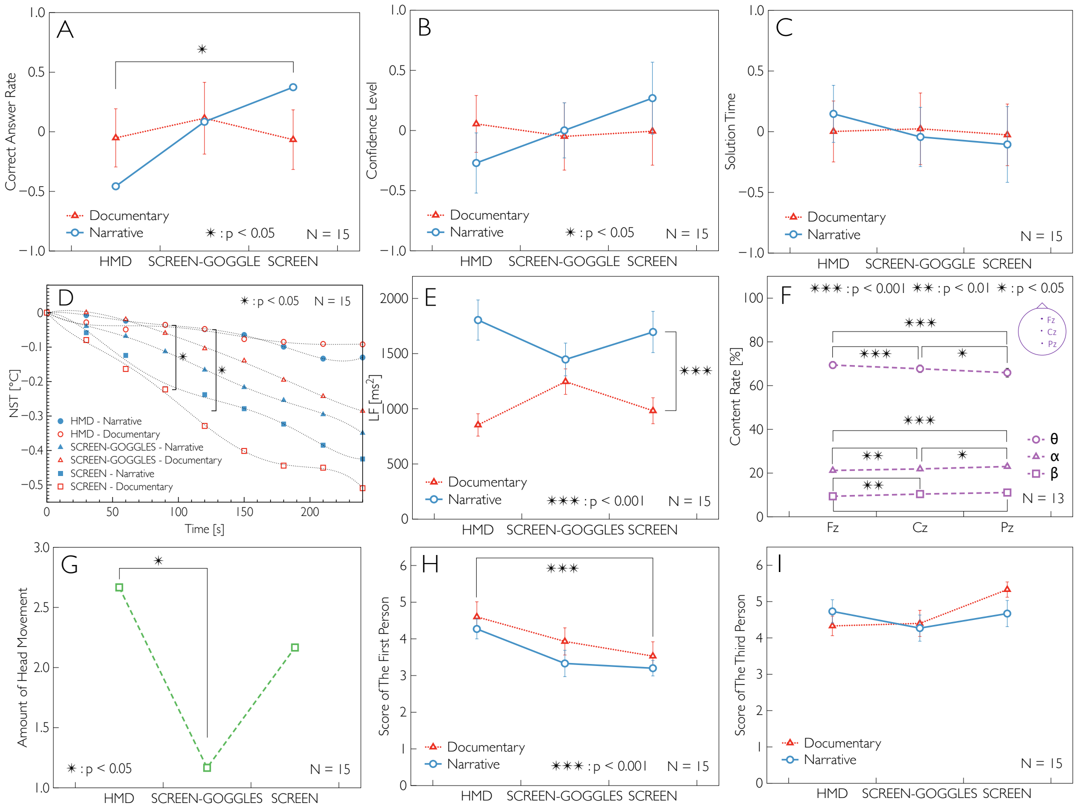

Fig.2 A, B, and C show the correct answer rate, recall rate, and response time on average. D shows the time-series change in the NST. E shows the LF on average. F shows the contents of theta, alpha, beta on average for the parts Fz, Cz, and Pz. The image on the upper right shows the positions of the electrodes. Fz is located in the frontal region, Cz is located in the center region, and Pz is located in top of the head region. A frontal lobe related to thought and cognition exists in the frontal region. We could not measure the data of two of the subject's owing to a malfunction of the measuring equipment. G and H are average scores of the first person/third person perspective questionnaire. All error bars show standard error of the mean.

relatively low significance compared to SCREEN-GOGGLE.

5.1 Comprehension test

Fig. 2-B shows the confidence level (below the "recall" ratio) for the comprehension test. The higher the confidence level, the stronger the participant's confidence in his/her own responses. The results of a two-factor ANOVA showed no significant main effects. Fig. 2-C shows the response time for the comprehension test. The longer the response time, the stronger the possibility that the user's grasp of the video is dubious. As with the confidence level, a two-factor ANOVA did not show any significant main effect.

5.2 Nasal skin temperature

The chronological changes in the NST are shown in Fig. 2-D. A two-factor ANOVA revealed the main effect of the display type at 90 s ($F(2, 28) = 5.664$, $p = 0.009$) and 120 s ($F(2, 28) = 5.701$, $p = 0.008$). There was a significant difference between HMD and SCREEN at 90 s ($p = 0.043$) and 120 s ($p = 0.033$). For the temperature changes, the start of the experiment was used as a baseline. A low value indicates a strong trend toward stimulation of the sympathetic nervous activity when viewing the videos based on constricted blood vessels. It is clear that, under all scenarios, the NST decreased from the start of the experiment, indicating that the sympathetic nervous activity had been stimulated. A two- factor ANOVA was performed for every 30-s interval, and the results indicate that a main effect was present in the display type. Multiple comparisons were conducted using Bonferroni's method, and the results indicate that the users viewing a video with an HMD experienced significantly less change in NST than users viewing a video on the screens. Therefore, it was clarified that the stimulation of the user's sympathetic nervous activity was significantly less prominent with the use of HMDs.

5.3 Plethysmogram

The LFs calculated using the PTGs are shown in Fig. 2-E. The two-factor ANOVA revealed the main effect of the video genre ($F(1, 12) = 28.565$, $p = 0.000175$) and an interaction effect ($F(2,24) = 7.144$, $p = 0.004$). In addition, there was a significant difference in the video genre ($p = 0.000175$). A high value indicates activity in the autonomic nervous system connected to the cardiovascular system. A two-factor ANOVA was conducted, and the results indicate that a main effect was present in the video genre. It was clear that the activity was significantly lower when the users were viewing documentaries than when viewing narratives.

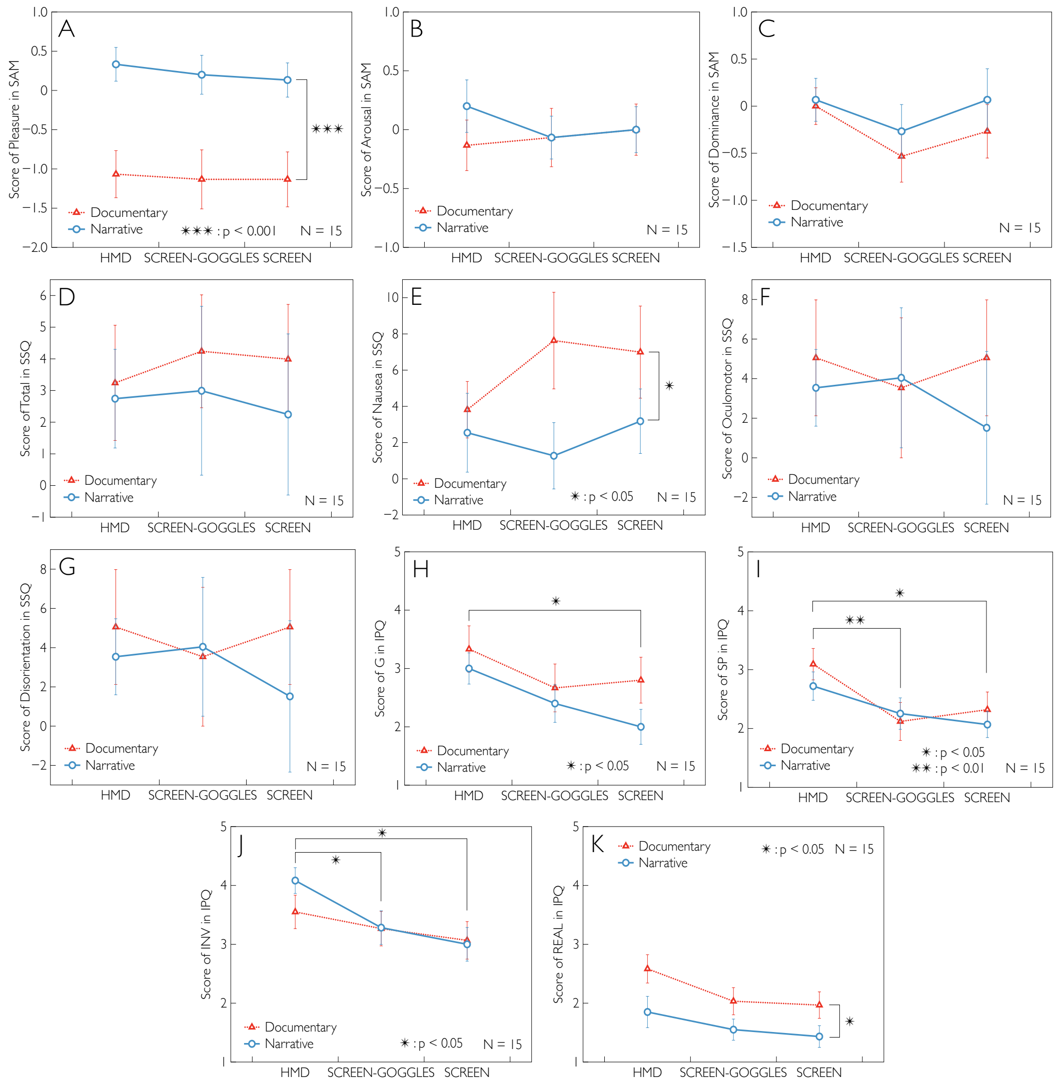

Fig.3 A, B, and C show average scores of pleasure, arousal, and dominance in SAM. D, E, F, and G show the average total, nausea, oculomotor, and disorientation scores of SSQ. H, I, J, and K show the average scores of general presence, spatial presence, involvement, and experienced realism for IPQ. All error bars show standard error of the mean.

Stimuli that greatly stimulated the sympathetic nerves inhibited the LFs; therefore, it is highly probable that viewing documentaries significantly stimulated the sympathetic nervous activity in relation to the cardiovascular system. In addition, an inter- action effect was observed, and multiple comparisons were conducted using Bonferroni's method. As a result, it was clear that differences in LFs were produced through the combination of HMDs and documentaries and of SAs and documentaries.

5.4 Electroencephalograms

The theta, alpha, and beta contents calculated from the EEGs for each part are shown in Fig. 2-F. The two-factor ANOVA revealed the main effects of the brain type within the brain region between Fz and Cz ($p = 0.000227$), Fz and Pz ($p = 0.000424$), and Cz and Pz ($p = 0.029587$) in the theta re- gion regarding the theta content ($F(1.310, 32.739) = 15.889$, $p = 0.000118$), alpha content ($F(1.333, 33.313) = 14.450$, $p = 0.000195$), and beta content ($F(1.230, 30.758)=5.945$, $p=0.016$). In addition, there was a significant difference in content. There was a significant difference in the type of brain region between Fz and Cz ($p = 0.000227$), Fz and Pz ($p = 0.000759$), and Cz and Pz ($p = 0.01627$) in terms of the alpha content. There was a significant difference in the type of brain region between Fz and Cz ($p = 0.001299$) and Fz and Pz ($p = 0.014901$) in the beta content. The theta content reflects cognitive activity and sleep. In particular, as the frontal lobe is the part of the brain connected to cognitive function, the closer this activity is to the frontal lobe, the more likely that cognitive activity takes place during video viewing. In addition, the alpha content is said to reflect the interruption of cognitive processing and a standby state in the brain. The closer this activity is to the frontal lobe, the more strongly it is suggested that functions inhibiting cognition and physical activity will be active. In particular, the alpha content tends to increase as awareness declines. The beta content is said to reflect the active cognitive processing in a resting state, i.e., the closer this activity is to the frontal lobe, the greater the possibility that the brain is in an idling state. When the parts were compared to each other, the experiment results showed that for content close to the frontal lobe, the percentage of theta content was significantly higher, and the percentages of alpha and beta were significantly lower. These trends were maintained across display types.

5.5 First/Third-person perspective questionnaire

The results of the first-person perspective questionnaire from the first-person/third-person perspective questionnaire are shown in Fig. 2-H. The two-factor ANOVA revealed the main effect of the display type in the first-person perspective questionnaire ($F(1.438, 20) = 7.358$, $p = 0.008$). There was a significant difference between the HMD and SCREEN groups ($p = 0.0009$). A high score here means that the user felt as if he or she was viewing the video from a first-person perspective. The results of the two-factor ANOVA indicated that a main effect was present in the display type. The results of using the Bonferroni method to conduct multiple comparisons indicated that the score for the HMD was significantly higher than that for SCREEN. However, no significant difference was observed between HMD and SCREEN-GOGGLE, and there was a difference in first-person perspective results depending on whether or not the user had a limited field of vision. The results of the third-person perspective questionnaire from the first-person/third-person perspective questionnaire are shown in Fig. 2-I. A high score here indicates that the user felt as if he or she was viewing the video from a third- person perspective. The results of the two-factor ANOVA did not indicate the presence of a main effect or interaction effects, and no significant differences were observed between the display or video types.

5.6 SAM questionnaire

The results of the SAM for pleasure are shown in Fig. 3-A. A two-factor ANOVA revealed the main effect of video genre ($F(1, 14) = 19.091$, $p = 0.0001$). There was a significant difference between documentary and narrative ($p = 0.001$). A high pleasure rating indicates that the user has pleasant feelings toward the video. No differences were observed between display types. By contrast, video-wise, narrative video received a significantly higher pleasure score than documentary video. The SAM results for arousal are shown in Fig. 3-B. A high arousal rating indicates that the user found viewing the video stimulating. No significant differences were observed in this value between the display or video types. The SAM results for dominance are shown in Fig. 3- C. A high dominance rating means that the user felt a sense of dominance over the space while viewing the video. No significant differences were observed in this value between the display or video types.

5.7 IPQ questionnaire

The IPQ is a questionnaire that evaluates the immersion, sense of presence, and realism of the media viewed and its video. The questionnaire was administered after viewing the 360° video. The users responded to 14 questions regarding the media's sense of realism on a 7-point scale ranging from 0 to 6. The questions covered four areas: general presence (GP), spatial presence (SP), involvement (INV), and experienced realism (REAL). In general, GP is an index expressing how real the virtual experience feels, whereas SP is an index expressing the sense that one's body is present in the virtual space. In addition, INV is an index of the degree to which a user is charmed and fascinated by being in the virtual space, and REAL is an index that expresses how close the experience of being in the virtual space is to reality. An average score for each question assigned to a given index was determined, and the scores were used as evaluation indices [26].

5.8 SSQ questionnaire

The total scores for the SSQ are shown in Fig. 3-D. A high total score means that the user felt a sense of motion sickness while viewing the video. No significant differences were found between display or video types. The SSQ results for nausea are shown in Fig. 3-E. The two-factor ANOVA revealed the main effect of the video genre ($F(1, 14) = 5.362$, $p = 0.036$). There was a significant difference between documentary and narrative ($p = 0.036$). A high nausea score indicates that the user felt significant stomach discomfort while viewing the video. No significant differences were found between display types; however, for the video types, nausea scores were significantly higher for documentary videos than for narrative videos. The SSQ results for the oculomotor are shown in Fig. 3-F. A high oculomotor score indicates that the user's eyes grew tired while watching the video. No significant differences were found between the display or video types. The SSQ results for disorientation are shown in Fig. 3-G. A high disorientation score means that the user felt a sense of dizziness or vertigo while viewing the video. No significant differences were found between display or video types.

6. DISCUSSION

6.1 Do field-of-view limitations affect the evaluation of display types in terms of presence?

For RQ1, we shall refer to the results of the IPQ (Figs. 3-H, 3-I, 3-J, and 3-K). The experimental results assessing GP, SP, and INV indicate that there is a significant difference between the HMD and SCREEN. For both comparisons, the HMD produced a higher sense of presence and immersion, suggesting that viewing a video with the HMD tends to instill a stronger sense of presence in the user than viewing it with a screen setup. This finding could be related to the idea that a high sense of presence and immersion is characteristic of media created from transmitted light and heightens and introspection [5]. Furthermore, there is a significant difference between HMD and SCREEN-GOGGLE in GP. This suggests that the limitation of the visual field may affect the presence. By contrast, the results related to the REAL index of the IPQ show the presence of significant differences in video genre, with documentaries demonstrating a significantly higher sense of realism than narrative videos. The video compared in this experiment consists of rather unrealistic narratives and documentaries with serious content, and thus a difference can be considered to be present here.

6.2 Does content affect the display type ratings in terms of user memory? (RQ2)

For RQ2, we referred to the experiment results of the memory tests (Figs. 2-A, 2-B, and 2-C), NST (Fig. 2-D). The memory test results show that users of SCREEN provided significantly higher correct answers than users of the HMD. In addition, the content type did not affect the percentage of correct answers. These results indicate that content type may not affect users' memory during a passive panoramic video viewing. Surprisingly, a considerable difference was observed in the rate of correct answers provided during the memory test; however, the resulting conviction of the degree of user understanding during the memory test did not differ depending on the display-type conditions. This suggests that, although the understanding level of the users change depending on the display type, the user's may not realize it (i.e., the HMD users do not realize that their understanding is low). In addition, the results were partially consistent with the results of the NST. The NST decreases when people experience high mental loads [3]. The NST of users with the HMD did not decrease compared to users of SCREEN and SCREEN. In other words, users wearing the HMD may have had a lower mental load than users wearing the SCREEN, and they may have viewed the images while in a daze. Therefore, the user's memory during a passive panoramic video viewing may not be affected by a content type. It also indicates that only the display type may affect the user's memory.

6.3 Does a passive panoramic viewing experience affect the display type in terms of user simulation sickness? (RQ3)

For RQ3, we shall refer to the experiment results related to LF (Fig. 2-E) of the SAM (Figs. 3-A, 3-B, and 3-C), SSQ (Figs. 3-D, 3-E, 3-F and 3-G), and the REAL index of IPQ (Fig. 3-K). A significant difference between the genres of the video content was found in the nausea score of the SSQ. Similarly, the difference in the effect by video genre is also shown in the REAL scores of LF and IPQ. However, there was no significant difference between video content genres in indices other than nausea for SSQ. In other words, symptoms caused by VIMS, such as eye fatigue, dizziness, and lightheadedness, did not occur. We believe that there are two potential factors that influence nausea in the SSQ. One is the narrative. MacQuarries et al. [5] reported that narrative content is a good choice. The difference between documentary and narrative video content is the presence or absence of a narrative. The second is the video content itself. MacQuarries et al. also reported that horror images do not fare well. Documentaries are realistic because they include current events. Narratives, by contrast, are unrealistic because they are animated. In fact, the REAL scores on the IPQ were significantly different based on the video content. Therefore, we speculate that the potential factors that influence the nausea score on the SSQ are narrative and realism. In addition, the reason for the high nausea score when watching the documentary video can be explained by the LF component. In addition, del Paso et al. [39] reported that the physiological and psychological manipulations that greatly increase the sympathetic activity conversely decrease the LF power. This effect on the cognitive process may have led to user discomfort.

6.4 Does a passive panoramic viewing experience affect the display type in terms of the points of view of the users? (RQ4)

For RQ4, we shall refer to the experiment results of the first-person/third-person perspective test (Figs. 2-G and 2-H). Both sets of results indicate that viewing video using the HMD produces a greater sense of first-person perspective compared to viewing video using the SCREEN, and makes the user feel as if he or she is really in the world of the video. He et al. [13] experiment results found a significant difference of a sense of presence between the HMD and one-sided SCREEN. Therefore, the sense that the user is actually in the virtual space might also be connected to a sense of presence between the HMD and four-sided SCREEN. There were also no significant differences found between using the HMD and using the SCREEN-GOGGLE, which suggests that the user's field of view during the viewing experience can be considered a primary factor impacting the user's sense of perspective. Based on the above results, the display type can be considered to affect the user's perspective.

7. CONCLUSION

This study examined the impact of two different environments, a VR environment and a non-VR environment, on users when viewing a panoramic video. The study examined this issue comprehensively from perspectives: physiological and psychological, and memory aspect. Three different viewing environments were compared: HMD, SCREEN-GOGGLE, and SCREEN. In this study, we discovered that differences in the viewing environment impacted the users' sense of perspective, memory, sense of presence, and simulation sickness.

REFERENCES

[1] C. Cruz-Neira, D. J. Sandin, and T. A. De. Fanti, "Surrounded-screen projection-based virtual reality: the design and implementation of the CAVE" *In Proceedings of the 20st Annual Conference on Computer Graphics and Interactive Techniques*, pp. 135-142, 1993.
[2] I. E. Sutherland, "A head-mounted three-dimensional display" *In Proceedings of The American Federation of Information Processing Societies*, pp. 757-764, 1968.
[3] H. Genno, K. Ishikawa, O. Kanbara, M. Kikumoto, Y. Fujiwara, R. Suzuki, and M. Osumi, "International Journal of Industrial Ergonomics" *In Proceedings of the 20st Annual Conference on Computer Graphics and Interactive Techniques*, Vol. 19, No. 2, pp. 161-171, 1997.
[4] J. Ichino, H. Asano, Y. Tokunaga, and M. Wada, "Effects of passive panoramic video viewing using immersive displays on user's cognitive activity", *Transactions of Information Processing Society of Japan*, Vol. 61, No. 11, pp. 1757-1770, 2020.
[5] A. MacQuarries, and A. Steed, "Cinematic virtual reality.: Evaluating the effect of display type on the viewing experience for panoramic video", *In Proceedings of 2017 IEEE Virtual Reality*, pp. 45-54, 2017.

[6] F. Weidner, A. Hoeschy, S. Poeschlz, and W. Brollx, "Comparing VR and non-VR driving simulations: An experimental user study", *In Proceedings of 2017 IEEE Virtual Reality*, pp. 281-282, 2017.
[7] V. I. Petkova, M. Khoshnevis, and H. H. Ehrsson, "The perspective mat- ters! Multisensory integration in egocentric reference frames determines full-body ownership", *Frontiers in Psychology*, pp. 2-35, 2011.
[8] K. Kim, M. Z. Rosenthal, D. Zielinski, and R. Brady, "Comparison of desktop, head mounted display, and six wall fully immersive systems using a stressful task" *In Proceedings of 2012 IEEE Virtual Reality Workshops (VRW)*, pp. 143-144, 2012.
[9] K. Tcha-Tokey, E. Loup-Escande, O. Christmann, and S. Richir, "Effects on user experience in an edutainment virtual environment: comparison between CAVE and HMD", pp. 1-8, 2017.
[10] M. C. Juan, and D. Pérez, "Comparison of the levels of presence and anxiety in an acrophobic environment viewed via HMD or CAVE", *Presence: Teleoperator Virtual Environments*, Vol. 18, No. 3, pp. 232-248, 2009.
[11] M. Krijin, P. M. G. Emmelkamp, R. Biemond, C. W. Ligny, and M. J. Schuemie, "CC.A.P.G. van der Mast. Treatment of acrophobia in virtual reality: The role of immersion and presence", *Behaviour Research and Therapy*, Vol. 43, pp. 229-239, 2004.
[12] D. Fonseca, and M. Kraus, "A Comparison of head-mounted and hand- held displays for 360° videos with focus on attitude and behavior change", *In proceedings of Academic Mindtrek '16*, pp. 287-296, 2016.
[13] L. He, H. Li, T. Xue, D. Sun, S. Zhu, and G. Ding, "Am I in the theater? Usability study of live performance based virtual reality", *In Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pp. 1-11, 2018.
[14] J. Polcar, and P. Horejsi, "Knowledge acquisition and cyber sickness: A comparison of VR devices in virtual tours", *Science Journal*, pp. 613-616, 2013.
[15] A. A. Rizzo, L. Pryor, R. Matheis, M. Schultheis, K. Ghahremani, and A. Sey, "Memory assessment using graphics-based and panoramic video virtual environments" In Proceedings of 5th International Conference of Disability, *Virtual Reality and Association*, pp. 331-338, 2004.
[16] A. Philpot, M. Glancy, P. J. Passmore, A. Wood, and B. Fields, "User experience of panoramic video in CAVE-like and head mounted display viewing conditions" *In Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*, pp. 65- 75, 2017.
[17] T. Luong, N. Martin, F. Argelaguet, and A. Lécuyer, "Studying the mental effort in virtual versus real environments" *In Proceedings of IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 809-816, 2019.
[18] T. I. Chowdhury, S. M. S. Ferdous, and J. Quarles, "Information recall in a virtual reality disability simulation", *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, Vol. 37, pp. 1-10, 2017.
[19] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson, "Virtual reality induced symptoms and effects (VRISE). Comparison of head mounted display (HMD) desktop and projection display systems", *Displays*, Vol. 29, No. 2, pp. 58-69, 2007.
[20] L. Eudave, and M. Valencia, "Physiological response while driving in an immersive virtual environment" *In Proceedings of 2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pp. 145-148, 2017.
[21] T. Ahmaniemi, H. Lindholm, K. Muller, and T. Taipalus, "Virtual reality experience as a stress recovery solution in workplace", *In Proceedings on 2017 IEEE Life Sciences Conference*, pp. 206-209, 2017.
[22] M. Slater, B. Spanlang, M. V. Sanchez-Vives, and O. Blanke, "First-person experience of body transfer in VR", *PLoS ONE*, Vol. 5, No. 5, pp. 1-27, 2010.
[23] M. Cordeil, T. Dwyer, K. Klein, B. Laha, K. Marriott, and B. Thomas, "Immersive Collaborative Analysis of Network Connectivity: CAVE-style or Head-Mounted Display?", *IEEE Trans. Vis. Comput. Grap*, Vol. 23, No. 1, pp. 441-450, 2017.
[24] M. M. Bradley, and P. J. Lang, "Measuring emotion: The Self-assessment manikin and the semantic differential", *Journal of Behavior Therapy and Experimental Psychiatry*, Vol. 25, No. 1, pp. 49-59, 1994.
[25] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simula- tor sickness questionnaire: An enhanced method for quantifying simulator sickness", *The International Journal of Aviation Psychology*, Vol. 3, No. 3, pp. 203-220, 1993.
[26] H. Regenbrecht, and T. Schubert, "Real and illusory interactions enhance presence in virtual environments", *Presence Teleoperators and Virtual Environments*, Vol. 11, No. 4, pp. 425-434, 2002.
[27] A. Nozawa, M. Uchida, and H. Ide, "Estimation Method of Information Understanding in Communication by Nasal Skin Thermogram", *IEEJ Transactions on Fundamentals and Materials*, Vol. 216, No. 9, pp. 909- 914, 2006.
[28] T. Mizuno, T. Sakai, S. Kawazura, H. Asano, K. Akehi, S. Matsuno, K. Mito, Y. Kume, and N. Itakura, "Facial skin temperature fluctuation by mental work-load with thermography", *In Proceedings of the International Conference on Electronics and Software Science*, Vol. 136, No. 11, pp. 1581-1585, 2016.
[29] B. Insko, "Measuring presence: Subjective, behavioral and physiological methods", *Ios Press*, pp. 109-119, 2003.
[30] M. Michael, B. Insko, M. Whitton, and F. P. Brooks Jr, "Physiological measures of presence in stressful virtual environments", *ACM Transactions on Graphics*, Vol. 21, No. 3, pp. 645-652, 2002.
[31] S. L. Calvert, and S. L. Tan, "Impact of virtual reality on young adults' physiological arousal and aggressive thoughts: Interaction versus observation", *Journal of Applied Developmental Psychology*, Vol. 15, No. 1, pp. 125-139, 1994.
[32] L. He, H. Li, T. Xue, D. Sun, S. Zhu, and G. Ding, "Am in the Theater? Usability Study of Live Performance Based Virtual Reality", *In Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, Vol. 28, pp. 1-11, 2018.

[33] E. Baka, K. E. Stavroulia, N. M. Thalmann, and A. Lanitis, “An EEG- based evaluation for comparing the sense of presence between virtual and physical environments”, *In Proceedings on Computer Graphics International*, pp. 107-116, 2018.

[34] A. Gevins, M. E. Smith, H. Leong, L. McEvoy, S. Whitfield, R. Du, and G. Rush, “Monitoring working memory load during computer-based asks with EEG pattern recognition methods. Human factors”, Vol. 40, No. 1, pp. 79-91, 1998.

[35] H. Asada, Y. Fukuda, S. Tsunoda, M. Yamaguchi, and M. Tonoike, “Frontal midline theta rhythms reflect alternative activation of prefrontal cortex and anterior cingulate cortex in humans”, *Neuroscience Letters*, Vol. 274, No. 1, pp. 29-32, 1999.

[36] T. Inouye, K. Shinosaki, A. Iyama, Y. Matsumoto, S. Toi, and T. Ishihara, “Potential flow of frontal midline theta activity during a mental task in the human electroencephalogram”, *Neuroscience Letters*, Vol. 169, No. 1-2, pp. 145-148, 1994.

[37] W. Klimesch, P. Sauseng, and S. Hanslmayr, “EEG alpha oscillations: The inhibition-timing hypothesis”, *Brain Research Reviews*, Vol. 53, No. 1, pp. 63-88, 2007.

[38] H. Laufs, K. Krakow, P. Sterzer, E. Eger, A. Beyerle, A. Salek-Haddadi, and A. Kleinschmidt, “Electroencephalographic signatures of attentional and cognitive default modes in spontaneous brain activity fluctuations at rest”, *In Proceedings of the National Academy of Sciences*, Vol. 100, No. 19, pp. 11053-11058, 2003.

[39] G. A. R. d. Paso, W. Langewitz, L. J. M. Mulder, A. v. Roon, and S. Duschek, “Psychophysiology”, Vol. 50, No. 5, pp. 477-487, 2013.

