

Immersion Analysis Through Eye-Tracking and Audio in Virtual Reality

Jihoon Lee and Nammee Moon*

Department of Computer Science, Hoseo University, Asan-si, 31499, Korea
*Corresponding Author: Nammee Moon. Email: nammee.moon@gmail.com
Received: 09 January 2021; Accepted: 12 March 2021

Abstract: In this study, using Head Mounted Display (HMD), which is one of the biggest advantage of Virtual Reality (VR) environment, tracks the user's gaze in 360° video content, and examines how the gaze pattern is distributed according to the user's immersion. As a result of analyzing the gaze pattern distribution of contents with high user immersion and contents with low user immersion through a questionnaire, it was confirmed that the higher the immersion, the more the gaze distribution tends to be concentrated in the center of the screen. Through this experiment, we were able to understand the factors that make users immerse themselves in the VR environment, and among them, the importance of the audio of the content was shown. Furthermore, it was found that the shape of the gaze distribution for grasping the degree of immersion by the subject of the content was different. While reviewing the experimental results, we also confirmed the necessity of research to recognize specific objects in a VR environment.

Keywords: Virtual reality; eye-tracking; immersion analysis; 360° video

1 Introduction

As technology has progressed, smart cities and the Internet of Things (IoT) have been actively studied [1–3]. Several services such as remote healthcare services and IoT remote-home control services based on 5G mobile networks have been developed. These services have certain specific characteristics: they need higher network speeds and well-organized networks while improving mobility and providing flexibility and freedom of location. In particular, online services such as non-face-to-face lectures are a major breakthrough accomplished using smart city technology. First, they give people the means to use their time efficiently and freely. Second, these services enable people to maintain their regular routines even in the event of unforeseen emergencies such as epidemics or natural disasters. Given these advantages, several efforts have focused on improving the quality of online services, one of which is to use Virtual Reality (VR) technology.

VR using a Head-mounted Display (HMD) provides more immersive content than existing 2D content. An HMD can provide user immersion in virtual environments. Research on HMDs is underway to increase user immersion by various means. Through a study that deduces facial expressions based on pupil movements and the shape of eyes from cameras for eye-tracking, we can see that the quality of interactions between users and VR content has improved [4–6].



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Several technological innovations have been applied in response to COVID-19. To prevent the spread of the virus, telecommuting, non-face-to-face lectures, and online meetings are being actively employed [7,8]. These measures have been positively evaluated regarding work efficiency. However, these measures have received more criticism in the educational domain. A typical reason is that it is difficult for the teacher to measure the students' attention unless the teacher and the student are in the same physical space. In this same context, students may have difficulty in establishing immediate communication with the teacher.

In this study, the degree of immersion through the user's gaze tracking coordinate data in a virtual environment through 360-degree video is investigated. Based on researches through eye-tracking data and immersion in existing 2D contents, we check whether they are applied equally to VR contents. In addition, as well as the effect of sound for 2D contents' immersion improvement, we checked how the sound of VR contents shows the difference in immersion improvement compared to 2D contents.

2 Related Works

To improve user immersion in 2D content, research was conducted on optimal localization with sound effects, gaze tracking, and environmental changes. Among them, descriptions of the content were added through sound and it was confirmed that there is a clear difference in the degree of immersion associated with the user's environment. Reference [9] proposed user-modeling according to the propensity by combining the user's location-positioning data and gaze-tracking data using Wi-Fi Received Signal Strength (RSS) technology [9]. This was done through experiments on smartphones, tablet PCs, and laptops to determine the difference in the disposition and screen size according to the distance between the user and the media. In addition, by classifying spaces such as public places or study rooms, the correlation between immersion score and concentration was derived based on the number of people sharing a space.

The COVID-19 epidemic is driving non-face-to-face activities such as social distancing and remote classes. As a result, research to improve content immersion in VR environment has become more active. VR content can provide content to users with better immersion with HMD device. Therefore, combining VR with various types of content such as games, education, and advertisements has been studied [10–14]. Specifically, from an educational perspective, VR can improve the quality of online lectures. Reference [15] researched VR for education focused on the use of immersive VR [15]. In addition, research to experience 360° content in VR is also underway. To apply the recent trend, studies for faster and simpler real-time streaming and algorithms for video-frame stabilization and focus assistance for stabilization of 360° content are also being conducted [16–18].

Eye-tracking technology tracks user gaze and measures the eyeball position and movement. It is used in research fields such as psychology and cognitive linguistics [19]. Based on these studies, eye-tracking was examined in Human–Computer Interaction (HCI) for applications such as marketing. It was also used to determine what types of content users found more interesting in a desktop environment rather than a VR environment and what elements were eye-catching on a webpage [20]. There has also been a considerable amount of research into applying these eye-tracking techniques to education and medical fields to remove the limitations of locations [21–23].

Reference [24] was cited for the questionnaire in the experiment and several modifications were made since the content cannot be manipulated [24]. Reference [24] conducted a total of three experiments on VR game content. Through their experiment, they confirmed that both objective

measurement through eye-movement and subjective measurement through a questionnaire can be effective. In addition, both positive and negative emotions were strong influences when participants engaged in immersion. The experiment in this paper visually confirms participants' eye-movement directly and shows the difference in the distribution of gazes in the HMD according to the degree of immersion through the questionnaire items that have proven effectiveness. Fig. 1 below is the heatmap on a 2D webpage and 3D 360° video.

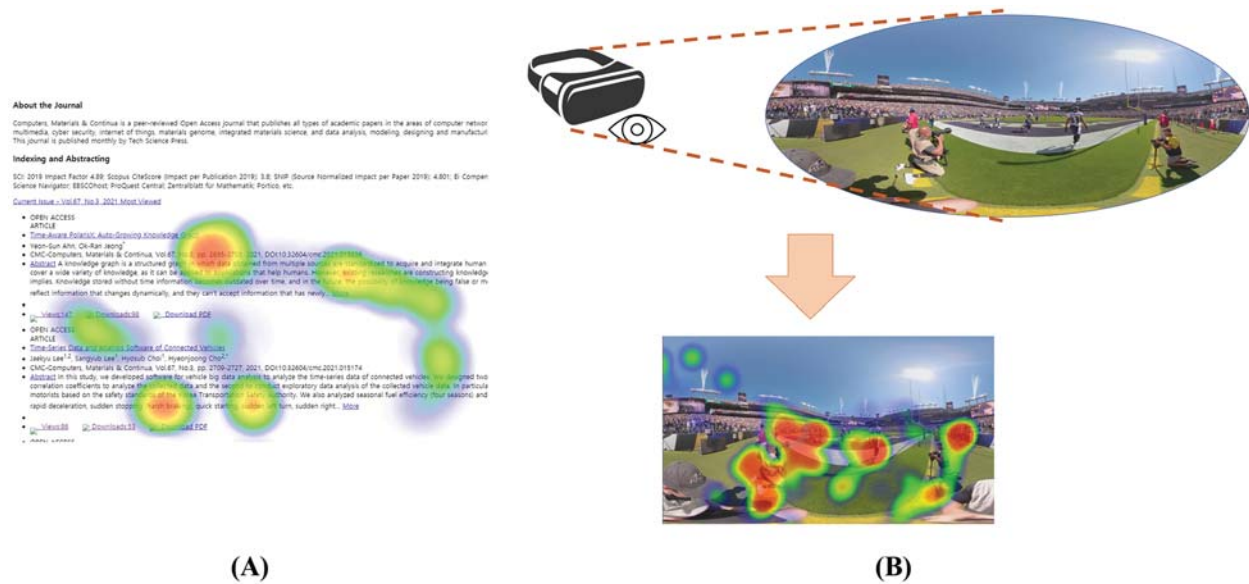


Figure 1: Eye-tracking on a 2D webpage (A) and a 3D 360° video (B)

3 Experiment

There were 20 experiment participants in total, all of whom were in their 20s. There were 13 males and seven females. Due to COVID-19 restrictions, the experiment was conducted in a separate room. Participants always wore masks for the experiment and their masks did not interfere with the HMD device or earphones. The experiment was conducted following social distancing protocols; during the experiment, no one including the author was allowed to enter the separate room except where unavoidable. Participants were able to immediately stop the experiment if they felt abnormal in any way or were severely unable to concentrate.

3.1 360° Video Contents and Questionnaire

The subjects of the 360° video contents for the experiment were sports and travel. Each participant experienced the two pieces of content depending on whether it included sound and four contents were experienced in total. Tab. 1 above shows example contents. And Fig. 2 below is the images of participant's eyes, HMD's location and gazed coordination.

Table 1: Contents information

Type	Content 1	Content 2	Content 3	Content 4
Played with audio	O	X	O	X
Category	Sports	Sports	Travel	Travel

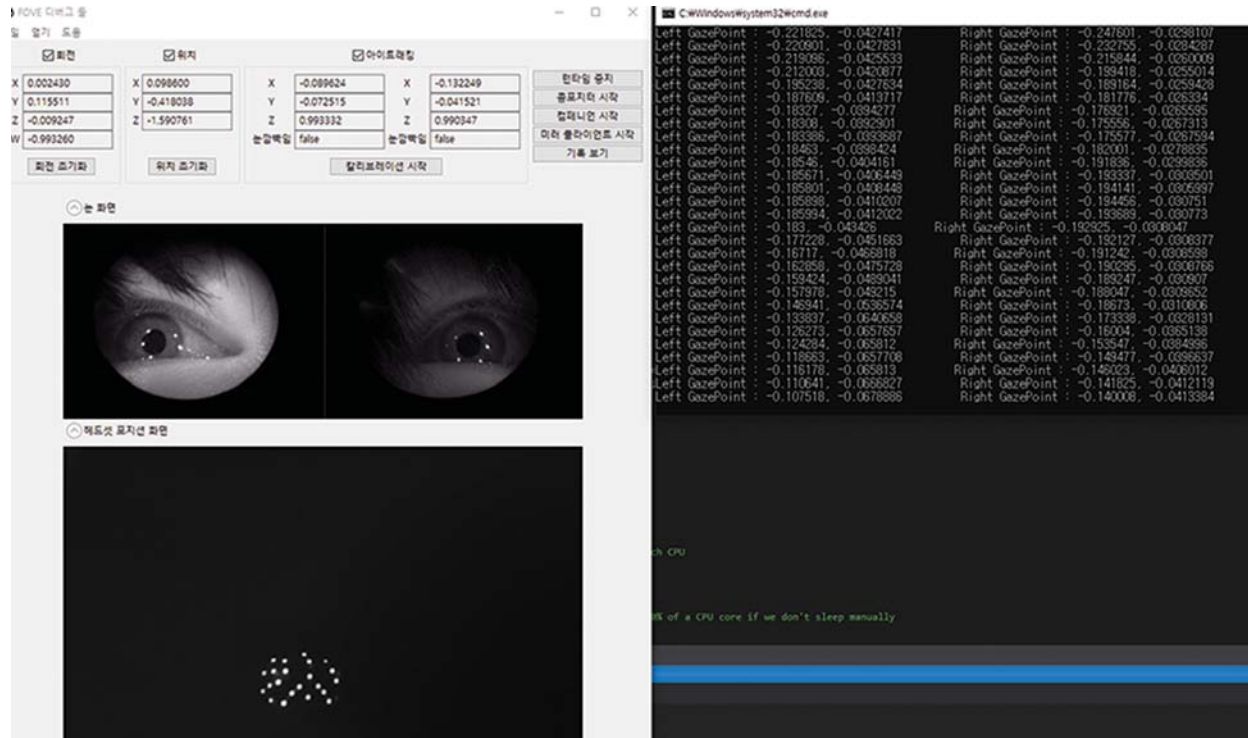


Figure 2: Participation in the experiment

While subjects were experiencing the content, they could quit out and end the experience at any time they wanted. In addition, the subjects were allowed sufficient rest time until the next content was executed. Whenever each content piece was finished, a survey was conducted, the content of which was a modified version of the survey contents in Reference [21]. Immediately after the end of the content, the subject was asked to respond to the questionnaire, the content of which can be found in the Appendix.

3.2 Experiment Environment

In this study, we experimented with two subjects of 360° content in a VR environment. Tab. 2 shows the environment of our experiment. We used FOVE VR Eye-tracking HMD as the main display equipment and Apple’s Bluetooth earphone, AirPods Pro, as the main audio equipment.

Table 2: Experiment environment

Type	Content
OS	Windows 10
CPU	Intel i9-9900K
GPU	TITAN RTX
Programming language	C++ and Python 3.7
VR development tool	Unity 4.4.3f
VR HMD device	FOVE VR
Audio device	Apple airpods pro
Number of experiment volunteers	20

3.3 Data Collection

FOVE’s SDK was used to gather eye-tracking data used in the experiment. Fig. 3 below is the screen that can be seen when executing FOVE’s debug tool.

Tab. 3 describes the values of X, Y, Z, and W that can be seen in the (1) box of Fig. 3. These values are for the location of the HMD and indicate the horizontal, vertical, front and rear, and height positions, respectively.

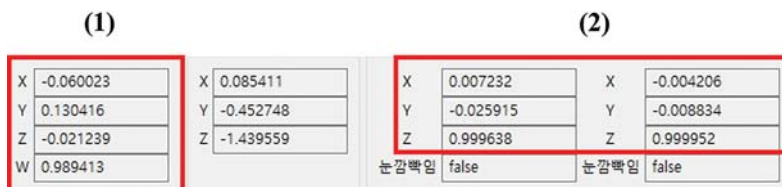


Figure 3: FOVE Debug tool screen

Table 3: Meaning of each value of HMD

Value	Meaning
X	HMD’s horizontal position
Y	HMD’s vertical position
Z	HMD’s front and rear position
W	HMD’s height location

Tab. 4 is a description of the values in box (2) in Fig. 3. Both the left and right sides have values of X, Y, and Z, which are values for the horizontal position of the eyeball, the vertical position of the eyeball, and the pupil dilation and contraction. The left and right are divided into left and right eyes.

Table 4: Meaning of each value of the tracked-gaze value

Value	Meaning
X	Horizontal position of an eyeball
Y	Vertical position of an eyeball
Z	Pupil dilation and contraction

We used the FOVE SDK to get the gaze coordinates of the FOVE debug tool. The gaze coordinate data was received directly from the HMD via C++. Figs. 4 and 5 below show the output of the C++ program and the part of csv files that shows the gaze points of each frame. The Gaze Point indicates the “Screen Gaze” of the FOVE SDK. In Screen Gaze, the center of the screen where the eye is looking is (0, 0), and the maximum coordinate value for each is 1. Fig. 6 below presents an example of Screen Gaze.

The columns in Tab. 5 correspond to the x and y coordinates for the left and right eyes on the HMD screen coordinates, respectively.

Fig. 6 below shows the mapping process for the extracted gaze data and the specific frame of the content.

When mapping the coordinates and video, there are a few things worth keeping in mind. First, it is necessary to prepare for a case where the HMD cannot recognize the eyes, such as when blinking. In this paper, for instances of None Data, which is the input for the eye blinking, the frame at the time is removed and mapped by five frames from the input point. Meanwhile, when None Data is checked as output, the HMD screen also progresses to a blank screen. In the 360° image, the part that the user can actually see from the HMD is not the entire content but only a 180° section that is the user's viewpoint. Therefore, the image to be used for mapping only proceeded to the screen viewed as the HMD's mirror client. Fig. 7 below shows the user's perspective when using the HMD.

```

C:\Windows\system32\cmd.exe
Left GazePoint : -0.221825, -0.0427417   Right GazePoint : -0.247601, -0.0298107
Left GazePoint : -0.220901, -0.0427831   Right GazePoint : -0.232755, -0.0284287
Left GazePoint : -0.219096, -0.0425533   Right GazePoint : -0.215844, -0.0260009
Left GazePoint : -0.212003, -0.0420877   Right GazePoint : -0.199418, -0.0255014
Left GazePoint : -0.195236, -0.0427634   Right GazePoint : -0.189164, -0.0259428
Left GazePoint : -0.187609, -0.0413717   Right GazePoint : -0.181776, -0.026334
Left GazePoint : -0.18327, -0.0394277   Right GazePoint : -0.176921, -0.0265595
Left GazePoint : -0.18308, -0.0392901   Right GazePoint : -0.175556, -0.0267313
Left GazePoint : -0.183386, -0.0393687   Right GazePoint : -0.175577, -0.0267594
Left GazePoint : -0.18463, -0.0398424   Right GazePoint : -0.182001, -0.0278835
Left GazePoint : -0.18546, -0.0404161   Right GazePoint : -0.191836, -0.0299836
Left GazePoint : -0.185671, -0.0406449   Right GazePoint : -0.193337, -0.0303501
Left GazePoint : -0.185801, -0.0408448   Right GazePoint : -0.194141, -0.0305997
Left GazePoint : -0.185898, -0.0410207   Right GazePoint : -0.194456, -0.030751
Left GazePoint : -0.185994, -0.0412022   Right GazePoint : -0.193689, -0.030773
Left GazePoint : -0.183, -0.043426   Right GazePoint : -0.192925, -0.0308047
Left GazePoint : -0.177228, -0.0451663   Right GazePoint : -0.192127, -0.0308377
Left GazePoint : -0.16717, -0.0466818   Right GazePoint : -0.191242, -0.0308598
Left GazePoint : -0.162858, -0.0475728   Right GazePoint : -0.190295, -0.0308766
Left GazePoint : -0.159424, -0.0489041   Right GazePoint : -0.189247, -0.030907
Left GazePoint : -0.157978, -0.049215   Right GazePoint : -0.188047, -0.0309652
Left GazePoint : -0.146941, -0.0536574   Right GazePoint : -0.18673, -0.0310806
Left GazePoint : -0.139837, -0.0640658   Right GazePoint : -0.173338, -0.0328131
Left GazePoint : -0.126273, -0.0657657   Right GazePoint : -0.16004, -0.0365138
Left GazePoint : -0.124284, -0.065812   Right GazePoint : -0.153547, -0.0384996
Left GazePoint : -0.118663, -0.0657708   Right GazePoint : -0.149477, -0.0396637
Left GazePoint : -0.116178, -0.065813   Right GazePoint : -0.146023, -0.0406012
Left GazePoint : -0.110641, -0.0666827   Right GazePoint : -0.141825, -0.0412119

```

Figure 4: Output of the C++ program for gaze vector coordinates

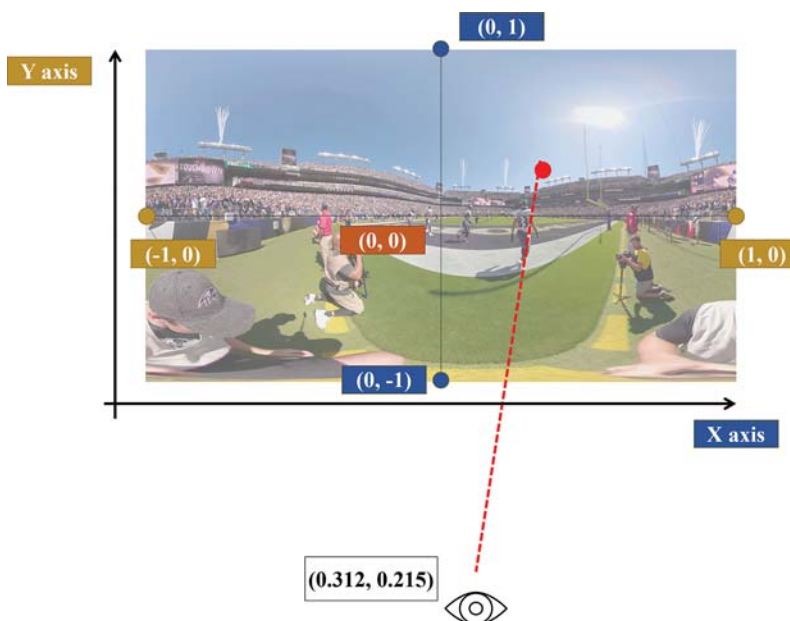


Figure 5: Example of the FOVE screen gaze vector

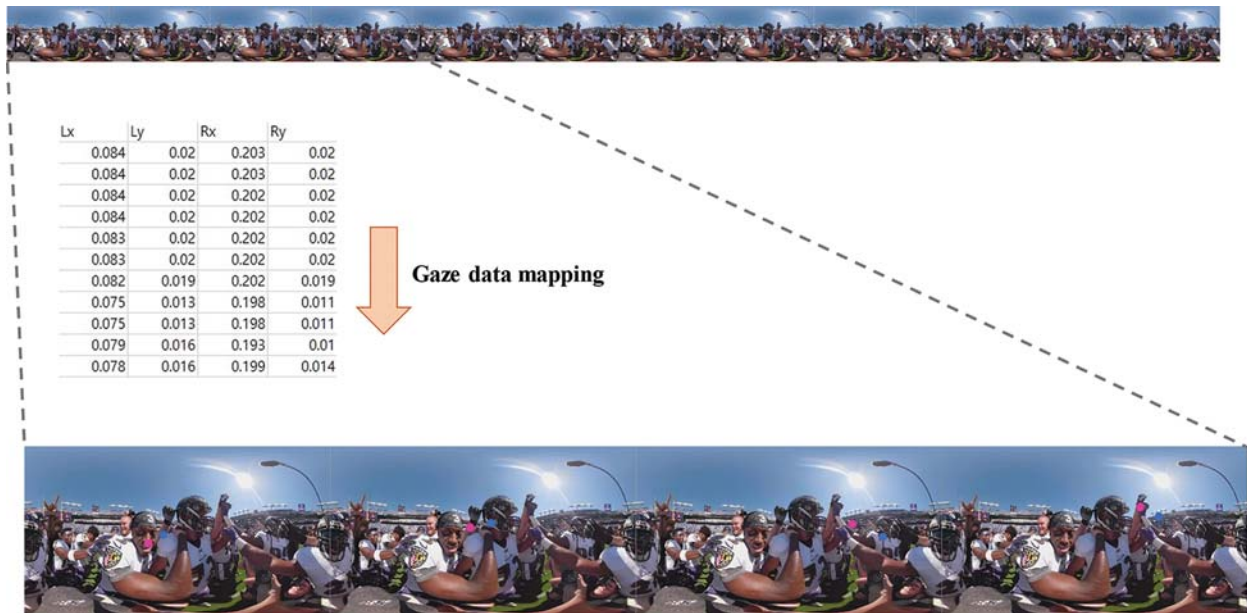


Figure 6: Gaze data and video-mapping process

Table 5: CSV example output of the C++ program

Timestamp	Lx	Ly	Rx	Ry
00:00:00	0.084	0.02	0.203	0.02
00:00:01	0.084	0.02	0.203	0.02
00:00:02	0.084	0.02	0.202	0.02
00:00:03	0.084	0.02	0.202	0.02
00:00:04	0.083	0.02	0.202	0.02
00:00:05	0.083	0.02	0.202	0.02
00:00:06	0.082	0.019	0.202	0.019
00:00:07	0.075	0.013	0.198	0.011
00:00:08	0.075	0.013	0.198	0.011
00:00:09	0.079	0.016	0.193	0.01
00:00:10	0.078	0.016	0.199	0.014

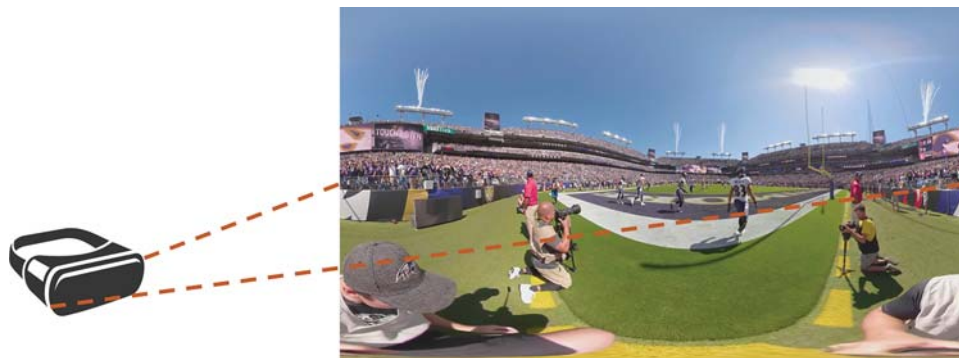


Figure 7: Example of the HMD's perspective

4 Experiment Result

Analyzing the collected data and looking at the positions of the participants' left and right eyes by content, we can derive graphs as shown in Fig. 9 below.

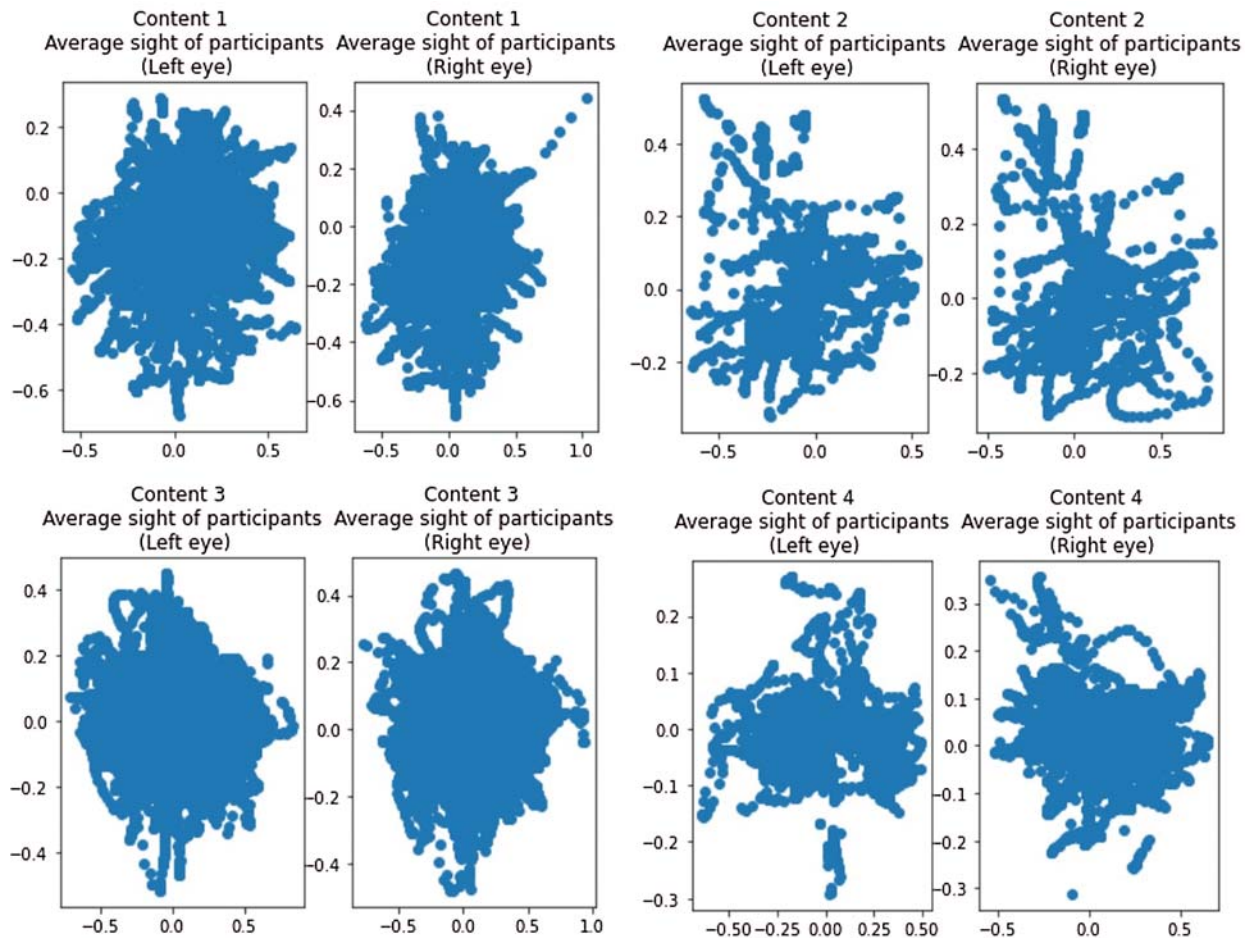


Figure 8: Distribution for each content

Tab. 6 shows the results of participants' questionnaires, which were conducted immediately after each content (D: Disagree, N: Neutral, A: Agree). The categorization of Tab. 6 has the following meaning: (E: Emotion, Int: Interest, Imm: Immersion, S: Satisfaction/P: Positive, N: Negative).

Number 1, the representative question of the questionnaire, is related to emotional transfer, an important factor in the degree of immersion. The average response is that Contents 1 and 3 that included sound were emotionally immersive; however, Contents 2 and 4 for which the sound had been removed were not well-transmitted.

Number 4 is a question that checks how much the participant assimilated into the content based on the degree of immersion. As an average response, Contents 1 and 3 responded that they wanted to move, and Content 2 and 4 showed neutral and negative responses, respectively. This question shows that the user's reaction may be different depending on the contents' characteristic.

In addition, in the experiment of this study, males have a greater proportion of males and females. With this factor, we could understand the result that Content 2 was neutral.

Table 6: Questionnaire results and averages

Questionnaire number	Content 1	Content 2	Content 3	Content 4	Categorization
1	D	N	D	A	E/N
2	A	D	A	D	Int/P
3	D	D	D	A	Int/N
4	A	N	A	D	Imm/P
5	A	A	A	A	S/P
6	A	D	A	D	S/P
7	D	N	D	A	S/N
8	A	D	A	D	Imm/P
9	D	A	D	A	Imm/N
10	N	N	A	D	Imm/P
11	D	A	D	N	Imm/N
12	D	A	D	A	Imm/N
13	A	D	A	D	Int/P
14	N	A	D	A	Imm/N
15	A	N	A	D	Int/P
16	A	N	A	N	Imm/P
17	N	N	N	D	Imm/P
18	A	D	A	N	Imm/P
19	D	A	D	A	Imm/N
20	7.71	4.14	8.5	7.14	Imm/S

Number 13 shows how strongly the user was interested in the content. Due to the nature of video content, unlike games that are directly manipulated, participants experience events in order without knowing what they will experience later. In the case of Contents 1 and 3, participants were interested in the content by being blocked from the outside on average.

Number 16 directly asks about the participant's immersion. Participants expressed that while experiencing the content on average, they seem to experience events directly within the content, rather than being in the real world. However, an important point is that Contents 2 and 4 had more neutral responses than disagreement. It can be seen that experiencing VR content through an HMD can give users a sense of immersion.

Number 18 asks how deeply the participant experienced the content. Contents 1 and 3 were each able to enable complete immersion in participants on average. However, negative responses were average for Contents 2 and 4, suggesting that whether the content includes sound greatly affects the level of immersion.

We can confirm some facts from [Tab. 6](#) and [Fig. 9](#). Although the questionnaire is a subjective indicator, relatively few participants were confirmed as neutral (N), calculated as average values. Through this, it can be confirmed that participants have relatively similar opinions on the content. What can be inferred from the immersion score for each piece of content for participants is that in the sports category, audio data is very influential on immersion. This is because, in the

case of sports content, not only visual elements such as commentary, shouting, cheering, etc. but also immersion through audio can be specified. However, in the case of travel, there are some differences but we can see that this has very little effect. In travel contents, enhanced immersion can be expected through the guide's audio data but it is understood that the influence of audio data on immersion is relatively weak because traveling alone can be viewed as sufficiently common. We can deduce one more fact here. The element that contributes to immersion in the subject of content is more effective if that element is not intended by the participant. For example, audio data in the sports category includes commentary, cheering, and shouting, as mentioned earlier. These elements can be heard sufficiently even if participants do not intend them, so this element can bring a sense of presence. However, in the case of travel, the only content that can be received as audio data is guide data or ambient white noise. This is because audio data has a significantly lower weighting than that of sports subjects, so it has the amount of change as shown in Fig. 9.

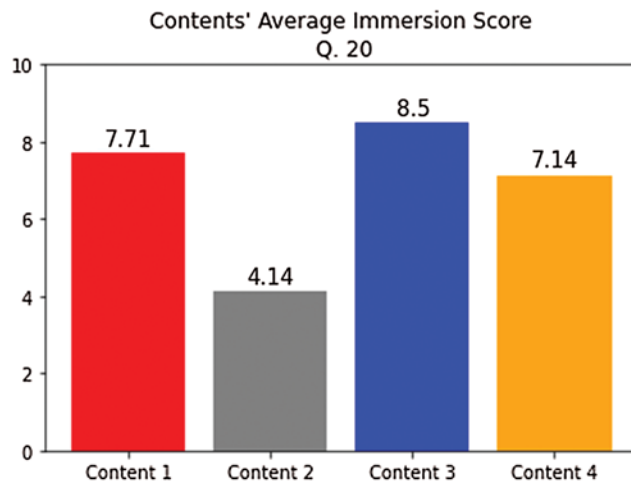


Figure 9: Averaged immersion scores for the questionnaire

5 Discussion and Conclusion

Smart cities are still in the process of development with IoT technology, future-generation networking services, and other technological innovations still being in the pipeline. In this context, we identified a common feature of studies in this domain: an online system with various devices. We focused on the word “immersion” in the keywords and had some confidence in the growth of the online education market. Improving the quality of online educational services such as lectures, conferences, and presentations requires increasing participants’ immersion in online environments. We found that the best way to achieve this is by means of a VR environment with an HMD device, which is commonly used when experiencing VR content. To enhance VR immersion, we intensively analyzed several factors that have been extensively researched in the past. Based on eye-tracking coordinates and audio data that can best be grasped in VR environments, we investigated what characteristics can be identified when a user is immersed and when they are not immersed.

The biggest factor that presents a problem in VR content is motion sickness. Until now, research to relieve VR motion sickness has been steadily progressing [25–28]. VR motion sickness induces symptoms such as nausea, vomiting, and dizziness, and Reference [29] said that through an experiment, there was a big difference in the degree of motion sickness according to movement

in the genre. In addition, although it depends on the user's environment and gender, the experience of VR does not have a significant effect.

When experiencing content in both 2D and in a VR environment, sound has a great influence on user immersion. First, in Fig. 8, you can see different gaze distributions between Contents 1 and 2 in the same category. In the case of Content 1, this content is reproduced with sound in the sports content and has a distribution concentrated in the center. In the case of Content 2, it is possible to check the distribution in a form that spreads in all directions by removing the sound.

Unlike Contents 1 and 2 that have the theme of sports, Contents 3 and 4 on the theme of travel show different sights. Both contents seem relatively less influenced by sound when compared to sports-themed contents. The theme of travel can give a better view of the objects you want to see by moving the position of the HMD device itself (the position of the head) to show what the user wants to see. As can be seen in Tab. 6, the immersion ratings of Contents 3 and 4 are not significantly different on average compared to Contents 1 and 2, which have the sports theme. This is an analysis of both objective indicators (Gaze data) and subjective indicators (Questionnaire) shown through Reference [21].

However, a limitation of this study is that, due to the nature of VR, which necessitates direct experimentation using HMD, it was not possible to recruit many experimental participants due to social distancing and non-face-to-face activities caused by the COVID-19 pandemic. If the experiment were conducted based on the sample data received from more participants, more accurate results could be obtained.

The lack of 360° video content is also a limitation. It is expected that more detailed results could be obtained if the contents of various subjects other than travel and sports conducted in this study were secured. In addition, motion sickness in VR can cause a big problem in immersion. In fact, in the case of the experimental participant with >70% of the missing values as a result of the experiment, it is said that only the first half of all contents were processed. When motion sickness arose, they were waiting for the contents to end. Since it was difficult for participants to judge that the content experience was valid, there was a situation where they gave up and were replaced with another person. In VR, motion sickness usually occurs when the frame rate of the content actually viewed by the user differs from the frame rate of the VR environment viewed through the HMD.

Through this study, we were able to approve the effect of sound on immersion in a VR environment and that understanding of immersion through gaze data was related to the subject of the content. With experiments in this paper, we were able to confirm the following:

- 1) Gaze data can be a good indicator to determine whether users are immersed but more accurate indicators should be made for gaze patterns by category. This experiment was conducted on contents with themes of sports and travel, but there are various other contents in addition to VR contents.

- 2) Sound data, like 2D content, has a great influence on content immersion. However, in the case of the travel category in the experiment in this study, it was confirmed that even without sound data, participants can still fully immerse themselves. Through this confirmation we were able to plan the future experiment for the category of the VR content.

- 3) Due to the use of an HMD device, which can be viewed as the difference between 2D and 3D contents, it was found that the gaze patterns for grasping immersion have different patterns.

For example, frequently mentioning the content topic derived from the content or continuously showing related objects is used to check whether the user was immersed in the 2D content. Therefore, to grasp the degree of immersion through the gaze pattern in 2D content, other factors (audio data contents, recognition of objects exposed to the content, etc.) must be identified. However, in the case of 3D content, when the user was immersed, the gaze pattern tended to be concentrated in the center. Conversely, when the user could not concentrate, the dispersive gaze pattern could be confirmed.

The study was conducted with a relatively small number of participants due to the COVID-19 pandemic. For future research, we will check the distribution of gazes in various contents such as fashion, education, and advertisements other than sports and travel, the similarities or differences between various contents, and the object recognition method in the VR environment to measure the immersion. It is necessary to proceed with this experiment in detail based on contents of various categories by collecting more content that was insufficient in this study. If further research proceeds, we plan to conduct more accurate data analysis by recruiting participants with more diverse topics and more diverse samples.

Furthermore, object recognition in VR content becomes a very important index for immersion. VR content that users experience with more choices, rather than showing only a fixed and limited screen as in 2D content, has limited capacity to grasp objects only with simple coordinates. Currently, it is widely used to grasp the location of an HMD's Gaze Vector Point through Unity and check the ID of the Asset there. However, object recognition using Unity's assets is impossible for 360° video, which is a type of VR content. To solve this problem, future research should check the scene facing the HMD in 360° video and conduct a study on object recognition methods through You Only Look Once (YOLO).

Funding Statement: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1A2C2011966).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] Y. S. Jeong and J. H. Park, "IoT and smart city technology: Challenges, opportunities, and solutions," *Journal of Information Processing Systems*, vol. 15, no. 2, pp. 233–238, 2019.
- [2] H. Alshammari, S. A. El-Ghany and A. Shehab, "Big IoT healthcare data analytics framework based on fog and cloud computing," *Journal of Information Processing Systems*, vol. 16, no. 6, pp. 1238–1249, 2020.
- [3] J. S. Park and J. H. Park, "Future trends of IoT, 5G mobile networks, and AI: Challenges, opportunities, and solutions," *Journal of Information Processing Systems*, vol. 16, no. 4, pp. 743–749, 2020.
- [4] J. Z. Lim, Mountstephens and J. Teo, "Emotion recognition using eye-tracking: Taxonomy, review and current challenges," *Sensors*, vol. 20, no. 8, pp. 2384, 2020.
- [5] R. Pathan, R. Rajendran and S. Murthy, "Mechanism to capture learner's interaction in VR-based learning environment: Design and application," *Smart Learning Environments*, vol. 7, no. 1, pp. 1–15, 2020.
- [6] M. Wang, X. Q. Lyu, Y. J. Li and F. L. Zhang, "VR content creation and exploration with deep learning: A survey," *Computational Visual Media*, vol. 6, no. 1, pp. 3–28, 2020.

- [7] C. Imperatori, A. Dakanalis, B. Farina, F. Pallavicini, F. Colmegna *et al.*, “Global storm of stress-related psychopathological symptoms: A brief overview on the usefulness of virtual reality in facing the mental health impact of COVID-19,” *Cyberpsychology Behavior, and Social Networking*, vol. 23, no. 11, pp. 782–788, 2020.
- [8] R. P. Singh, M. Javaid, R. Kataria, M. Tyagi, A. Haleem *et al.*, “Significant applications of virtual reality for COVID-19 pandemic,” *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, no. 4, pp. 661–664, 2020.
- [9] H. Song and N. Moon, “User modeling based on smart media eye tracking depending on the type of interior space,” *Advances in Computer Science and Ubiquitous Computing*, vol. 474, pp. 772–776, 2017.
- [10] A. Sara, D. Najima and A. Rachida, “A semantic recommendation system for learning personalization in massive open online courses,” *International Journal of Recent Contributions from Engineering, Science & IT*, vol. 8, no. 1, pp. 71–80, 2020.
- [11] F. Q. Chen, Y. F. Leng, J. F. Ge, D. W. Wang, C. Li *et al.*, “Effectiveness of virtual reality in nursing education: Meta-analysis,” *Journal of Medical Internet Research*, vol. 22, no. 9, pp. e18290, 2020.
- [12] L. Peng, Y. Yen and I. Siswanto, “Virtual reality teaching material-virtual reality game with education,” *Journal of Physics: Conference Series*, vol. 1456, no. 1, pp. 12039, 2020.
- [13] K. Ahir, K. Govani, R. Gajera and M. Shah, “Application on virtual reality for enhanced education learning, military training and sports,” *Augmented Human Research*, vol. 5, no. 1, pp. 1, 2020.
- [14] J. Radianti, T. A. Majchrzak, J. Fromm and I. Wohlgenannt, “A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda,” *Computers & Education*, vol. 147, pp. 103778, 2020.
- [15] V. Clay, P. König and S. U. König, “Eye tracking in virtual reality,” *Journal of Eye Movement Research*, vol. 12, no. 1, pp. 1–18, 2019.
- [16] A. Yaqoob, T. Bi and G. M. Muntean, “A Survey on Adaptive 360° video streaming: Solutions, challenges and opportunities,” *IEEE Communications Survey & Tutorials*, vol. 22, no. 4, pp. 2801–2838, 2020.
- [17] J. Kopf, “360 video stabilization,” *ACM Transactions on Graphics*, vol. 35, no. 6, pp. 1–9, 2016.
- [18] Y. C. Lin, Y. J. Chang, H. N. Hu, H. T. Cheng, C. W. Huang *et al.*, “Tell me where to look: Investigating ways for assisting focus in 360° video,” in *Proc. of the 2017 CHI Conf. on Human Factors in Computing Systems*, Denver, Colorado, USA, pp. 2535–2545, 2017.
- [19] R. Moratto, “Original paper a preliminary review of eye tracking research in interpreting studies: Retrospect and prospects,” *Studies in Linguistics and Literature*, vol. 4, no. 2, pp. 19–32, 2020.
- [20] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in *ICLR 6th Int. Conf. on Learning Representations*, Vancouver, British Columbia, Canada, pp. 1–14, 2014.
- [21] K. Sharma, M. Giannakos and P. Dillenbourg, “Eye-tracking and artificial intelligence to enhance motivation and learning,” *Smart Learning Environments*, vol. 7, no. 1, pp. 1–19, 2020.
- [22] H. Lim, “Exploring the validity evidence of a high-stake, second language reading test: An eye-tracking study,” *Language Testing in Asia*, vol. 10, no. 1, pp. 1–29, 2020.
- [23] T. T. Brunye, T. Drew, D. L. Weaver and J. G. Elmore, “A review of eye tracking for understanding and improving diagnostic interpretation,” *Cognitive Research: Principles and Implications*, vol. 4, no. 1, pp. 1–16, 2019.
- [24] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps *et al.*, “Measuring and defining the experience of immersion in games,” *International Journal of Human-Computer Studies*, vol. 66, no. 9, pp. 641–661, 2008.
- [25] X. Li, C. Zhu, C. Xu, J. Zhu, Y. Li *et al.*, “VR motion sickness recognition by using EEG rhythm energy ratio based on wavelet packet transform,” *Computer Methods and Programs in Biomedicine*, vol. 188, no. 4, pp. 105266, 2020.
- [26] E. Chang, H. T. Kim and B. Yoo, “Virtual reality sickness: A review of causes and measurements,” *International Journal of Human-Computer Interaction*, vol. 36, no. 17, pp. 1658–1682, 2020.

- [27] H. Kim, J. Park, Y. Choi, M. Choe, “Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment,” *Applied Ergonomics*, vol. 69, pp. 66–73, 2018.
- [28] L. Tychsen and P. Foeller, “Effects of immersive virtual reality headset viewing on young children: Visuomotor function, postural stability, and motion sickness,” *American Journal of Ophthalmology*, vol. 209, no. 24, pp. 151–159, 2020.
- [29] U. A. Chattha, U. I. Janjua, F. Anwar, T. M. Madni, M. F. Cheema *et al.*, “Motion sickness in virtual reality: An empirical evaluation,” *IEEE Access*, vol. 8, pp. 130486–130499, 2020.

Appendix A. Immersion Questionnaire List for Contents

Answers to check:

(SD: Strongly Disagree, D: Disagree, N: Neutral, A: Agree, SA: Strongly Agree)

- (1) I did not feel any emotional attachment to the content.
- (2) I was interested in seeing how the content’s events would progress.
- (3) It did not interest me to know what would happen next in the content.
- (4) I was so focused on the content that I wanted to make a movement related to the content topic.
- (5) I enjoyed the graphics and images of the content.
- (6) I enjoyed the content.
- (7) The content was not fun at all.
- (8) I felt myself to be directly travelling through the content according to my own will.
- (9) I did not feel as if I was moving through the content according to my own will.
- (10) I was unaware of what was happening around me.
- (11) I was aware of surroundings.
- (12) I still felt attached to the real world.
- (13) At the time, the content was my only concern.
- (14) Everyday thoughts and concerns were still very much on my mind.
- (15) I did not feel the urge at any point to stop playing and see what was going on around me.
- (16) I did not feel like I was in the real world but the content world.
- (17) I still felt like I was in the real world while watching it.
- (18) To me, it felt like only a very short time had passed.
- (19) The play time of the content appeared to go by very slowly.
- (20) How immersed did you feel? (10: Very immersed, 0: Not immersed).