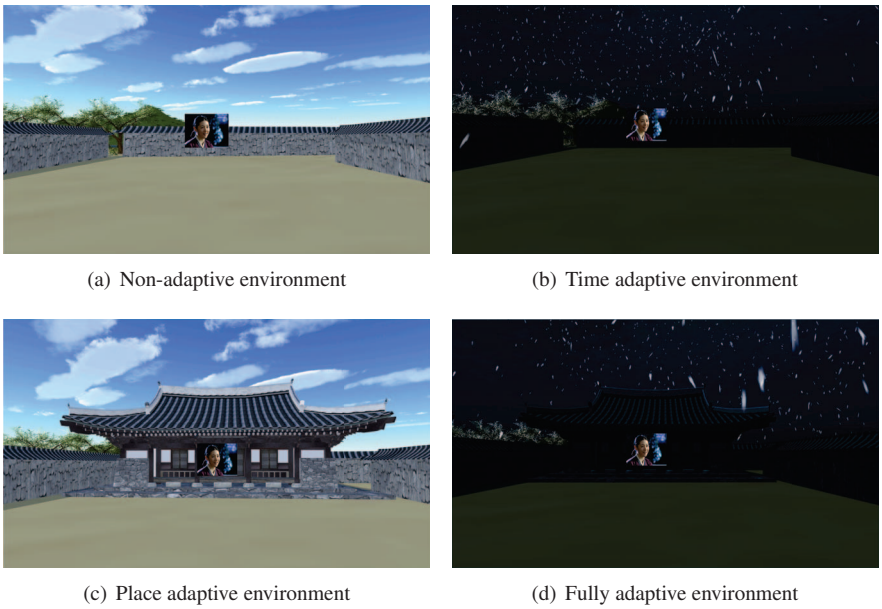
Video Content Representation to Support the Hyper-reality Experience in Virtual Reality

Hyerim Park*
KAIST UVR Lab

Woontack Woo†
KAIST UVR Lab



(a) Non-adaptive environment (b) Time adaptive environment

(c) Place adaptive environment (d) Fully adaptive environment

Figure 1: Four virtual environment conditions for video content.

Abstract

Most research on providing location-based content in 3D interactive virtual reality has been limited to social media content. Few studies have suggested how to represent the video clip of movies or TV shows in virtual reality. This paper investigates a video content representation method to provide a hyper-reality experience of the narrative world in virtual reality. We reflect the time and place settings of the video content in virtual reality and have participants watch the video in four different virtual reality environments. We reveal that reflecting the story's environment settings to the virtual reality environment significantly improves the spatial presence and narratives engagement. We also confirm a positive correlation between spatial presence and narrative engagement, including subscales such as emotional engagement and narrative presence. Based on the study results, we discuss how to provide the hyper-reality experience in content-adaptive virtual reality.

Keywords: Cinematic virtual reality, hyper-reality, video representation, narrative engagement, spatial presence

Index Terms: Human-centered computing—Human Computer Interaction (HCI)—Empirical studies in HCI—; Human-centered computing—Interaction design—Empirical studies in interaction design

*e-mail: hyerimpark@kaist.ac.kr
†Corresponding author. e-mail: wwoo@kaist.ac.kr

1 Introduction

Virtual reality technology has made various attempts to support the transportation into the narrative world beyond the limits of existing media such as text, image, and video [6, 22]. However, most research on providing multimedia content in 3D interactive virtual reality has been limited to social media content [2, 35]. Few studies proposed the video representation method in the filming locations or the related background place. Although Park et al. provided a video clip of a TV show that introduced the site in the reconstructed virtual heritage site, it only showed the video clips regardless of the context of content and environment [40]. There was also a proposal to combine various location-based photos, music, and storytelling videos with 3D virtual reality to create an informative and interactive tourist destination [26]. However, they did not design the specific content representation method.

To effectively represent video clips of movies or TV shows based on location in virtual reality, it is necessary to understand the film tourism experiences called hyper-reality. Hyper-reality describes "the condition whereby imitations or reproductions of reality acquire more legitimacy, value, and power than the originals themselves" [4]. Film tourists recognize and experience the filming location as a place depicted in the film regardless of the location's authenticity [17, 28]. The more visitors fully experienced the hyper-reality in film tourism, the higher they were satisfied with the tour experience [11]. In contrast, poor conditions or lack of film set remain caused dissatisfaction with the tourism experience. Thus, we aimed to investigate a video content representation method to support the user's hyper-reality experience in a virtual reality environment.

This paper examines the effect of content-adaptive virtual reality

2021 IEEE Virtual Reality and 3D User Interfaces (VR) | 978-1-6654-1838-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/VR50410.2021.00064

2642-5254/21/$31.00 ©2021 IEEE
DOI 10.1109/VR50410.2021.00064

on the hyper-reality experience. We expected that reflecting the time and place settings of video content in the virtual reality environment would enhance the hyper-reality experience. For the user study, we chose one of the most popular TV dramas that have led Korea's film-induced tourism as a stimulus [33]. To efficiently investigate the effect of reflecting the narrative setting, we selected a video scene with a time and place setting that is most different from a virtual environment that reproduces a real background location. We reconstructed the actual background of the video content into a 3D virtual reality and applied the spatial and temporal elements. We had participants watch the same video clip in four different virtual environments: the non-adaptive environment, the time adaptive environment, the place adaptive environment, and the fully adaptive environment. We evaluated the hyper-reality experience in each condition using the spatial presence and the narratives engagement questionnaires. We also analyzed the correlation between spatial presence and narrative engagement.

As a result, we revealed that reflecting the time and place setting of the video's narratives to the virtual reality environment significantly improved the hyper-reality experience. In particular, in our stimulus condition, reflecting time factors such as night and snow in the virtual reality environment had a more significant effect on the hyper-reality experience than the place factor such as a building. We also demonstrated a significant correlation between spatial presence and narrative engagement, including sub-scales such as emotional engagement and narrative presence. The content-adaptive virtual environment allowed participants to intuitively understand the story's spatial background and feel the presence of the narrative world. In addition, it enhanced users' sympathy and empathy for the characters of the video. We discussed how to provide a hyper-reality experience in content-adaptive virtual reality based on the experiment results.

The main novel contributions of the paper are:

- Proposal of a video content representation method reflecting the narrative setting of movies or TV shows in 3D Virtual Reality.
- Findings from a user study that demonstrates the effect of the proposed method on the user's hyper-reality experience.
- Discussions to provide hyper-reality experience in content-adaptive virtual reality.

2 Related Work

Film tourism stands for visitation to sites where movies and TV programs have been filmed or tour to related places such as production studios, film-related theme parks, and actual background places [5]. In film tourism, hyper-reality is a critical concept as tourists' expectations and experiences are created and perceived by cinematic narratives [13,18,31,41]. Various studies have been conducted on the hyper-reality experience and its effects on tourism satisfaction. Kim et al. showed that film tourists enhance their touristic experiences by re-enacting cinematic scenes [31]. Waysdorf et al. revealed that movie tourists imagine the filming location as a narrative world [55]. They also found that tourists perceive the environment around the filming location as a new story setting. Carl et al. revealed that the lack of preservation of the film set reduces the satisfaction of the visitors [11]. However, few studies have examined how to reproduce and preserve the filming site to give a higher hyper-reality experience.

Several technologies and designs have emerged for importing and reusing geo-tagged content in virtual reality. Social Street View [1] is an immersive geospatial social media system for virtual reality environments [16]. It renders the social media in its natural geospatial context provided by immersive maps, such as Google Street View. However, as shown in the demo video[2], the visualization of some images was rather crude and unnatural because of the different temporal and spatial context between the geo-tagged content and the virtual backdrop. Also, there have been attempts to integrate social media content into an immersive street-level 3D virtual city environment [2, 35]. Ruofei et al. proposed various virtual representation methods of social media in 3D VR environments [15]. However, research on the remediation of existing content in virtual reality has mainly focused on searching and visualizing social media content based on the location. Also, few studies have investigated context factors that should be considered in the virtual content representation.

In a narrative, there are three forms of involvement: spatial, temporal, and emotional. In particular, spatial involvement responds to the setting, delivering the sense of place [49]. The setting is a literary element used in novels, short stories, plays, and films, and indicates the time and place of the story [53]. Settings include the story's environment, which consists of geographic and physical location, physical environment, time period, or social and cultural backgrounds [3]. There are many ways in which time and place represent settings. Time covers the character's time of life, the time of day, time of year, time period such as the past, present, and future. Place covers many areas, such as room, building, city, country, beach, mountain, and various transportation modes. The setting is presented through the unfolding of the story. It helps readers or viewers understand the characters in the plot and creates an atmosphere of the narrative world. However, there have not been attempts to increase spatial involvement by reflecting the story setting to the virtual environment in providing video clips.

The new media experience for narrative video using an immersive display is described as cinematic virtual reality (CVR) [44]. CVR covers various concepts from passive 360 videos to interactive narrative videos that allow the viewer to affect the story. The new attempts have expanded the experience of video from the monoscopic, passive, and fixed-viewpoint manner to the immersive [20, 29], interactive [37], and dynamic-viewpoint manner [1]. However, most cinematic virtual reality research focuses on producing and experiencing new content rather than reusing and remediating existing content. Few studies provide an immersive experience about the existing drama or movie based on the background places. Thus, in this paper, we investigated the effect of matching the virtual environment with the story setting of the video content on users' hyper-reality experience.

3 User Study

We assumed that a hyper-reality experience based on a 2D video in 3D virtual reality is included in the CVR experience. Thus, we adopted two variables, spatial presence and narrative engagement, used as the CVR experience evaluation method to measure the hyper-reality experience [36]. We expected that reflecting time and place elements, which are narrative settings of video content, would significantly impact users' spatial presence and narrative engagement in a virtual reality environment. Also, we supposed that there would be a positive correlation between the spatial presence and narrative engagement. Our research questions were as follows:

RQ1. How do reflecting time and place elements of the story setting in virtual reality affect the user's spatial presence or narrative engagement?
RQ2. Is there a significant correlation between spatial presence and narrative engagement?

We set four experimental conditions to investigate the above questions as follows:

[1] https://socialstreetview.com/

[2] https://vimeo.com/221188014

[3] https://literaryterms.net/setting/

- Non-adaptive environment: virtual reality reflecting real world conditions
- Time adaptive environment: virtual reality reflecting temporal settings of the video scene
- Place adaptive environment: virtual reality reflecting spatial settings of the video scene
- Fully Adaptive environment: virtual reality reflecting both temporal and spatial settings of the video scene

The non-adaptive environment is a virtual reality environment that reflects the real world and is a basic content experience environment. In most cases, the background place or filming location of a movie or drama does not preserve the story's time and place setting. For example, in the case of Ben-Hur[4], users cannot see the Colosseum they saw in the movie in a virtual reality environment that is reconstructed based on its current appearance like a digital twin [19]. The time adaptive environment is a virtual environment in which the time setting of the video is applied. The time element includes the time of day, time of year, and time period. Therefore, when the video's temporal background is night, the virtual reality environment is adjusted to the night. Environmental factors representing a particular season or period also apply to the virtual environment. The place adaptive environment applies a place element to the non-adaptive environment. If the background building is not in the actual background place or its appearance is different due to CG or props, the building is reproduced as in the video. The fully adaptive environment completely reproduces the environmental settings shown in the video by applying time and place elements to the non-adaptive environment.

We chose a within-subject test design. CVR research has used a within-subject experiment to effectively compare and evaluate various visual [45–48] and auditory [25] stimuli on presence and engagement. Thus, we designed a within-subject experiment to allow participants to analyze and assess the impact of each environmental condition on the hyper-reality experience. For removing order effects, we used the complete counterbalancing method. All possible orders were used, and each subject was given a different random order of conditions.

3.1 Study Setup

For user study, we chose a Korean television series, called Dae Jang Geum [5] known as "Jewel in the Palace" in English. It has been one of the most popular TV dramas that have affected the film-induced tourism of Korea [32]. The story is based on the true story of Jang Geum, the first female royal physician of the Joseon Dynasty. As a stimulus, we selected the scene where the setting of time and place is the most different from the virtual environment replicating the real world. In the scene, two main characters sat on the floor and talked about their love and life on a snowy night. Depending on the general range of narrative settings, the temporal elements of the scene were night and snow that represent the time of day and time of year. The spatial element was the building in which the characters were sitting in the video. The scene was filmed on an outdoor set, but the actual setting of the story is the historical place called Jeju-mok government office[6]. As it is a historic site, tourists can visit it for a limited time, and only the site of the pharmacy building where the main characters was sitting remains. Therefore, tourists cannot experience the story environment in the real world or the virtual world that realistically reconstructs the site.

The prototype was developed in Unity 2018.2.13f1, and we used the Google Daydream SDK for the virtual reality function. For reproducing the spatial background of the narrative, we realistically reconstructed Jeju-mok government office using Autodesk MAYA. We also reconstructed the pharmacy building, the main place of the story, referencing the video scene. The total faces of polygons were 405,055. Every UV map was warped and redesigned in MAYA UV editor. Location and proportion were set according to the satellite map provided by Google. Every building model was exported as FBX format into Unity. The background place of the video scene was implemented as shown in Figure 2. For the temporal setting, we imported 'cloudy night' and 'clear noon' from the Sky-box Pack [7], and 'cinematic blizzard' from Snow VFX [8]. The video clip was first edited by automatic shot detection using the video editing tool Wondershare Filmora[9], and then manually merged into a scene based on the location. The duration of a video clip was 3 minutes and 47 seconds.

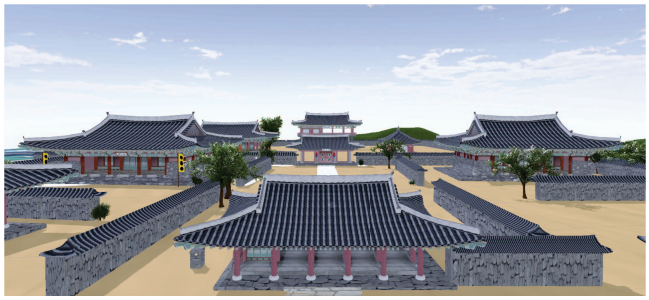


Figure 2: The virtual background place of the video content.

As shown in Figure 1(a), in the non-adaptive environment, the user watched the video in a virtual environment that did not reflect any of the environmental settings of the story and reproduced the real world. As shown in Figure 1(b), in the time adaptive environment, the user watched the video in Jeju-mok government office on a snowy night, but there was no background building. As shown in Figure 1(c), in the place adaptive environment, the user watched the video in front of the pharmacy building, but the time setting was not synchronized with the video scene. Finally, as shown in Figure 1(d), in the fully adaptive environment, the user watched the video in a virtual environment reflecting both spatial and temporal settings of the story.

3.2 Measures

We adopted the spatial presence and the narrative engagement questionnaire to measure the hyper-reality experience. For evaluating the spatial presence, we adopted the MEC Spatial Presence Questionnaire [54]. Spatial presence has often been referred to as experience of being there in a mediated environment [24]. We evaluated the presence with total 24 statements. All items rated on 5-point Likert scale. In adopting the questionnaire, we replaced the 'medium' to 'drama'.

For the narrative engagement, we used the Measuring Narrative Engagement Questionnaire (MNEQ) [9]. This scale examines how the participant engages with narrative across four subscales: narrative understanding, attentional focus, narrative presence, emotional engagement. The score for each subscale is computed by averaging the three items associated with the subscale, and an overall narrative engagement score is computed by calculating the mean for all 12 items. All items rated on 7-point Likert scale.

[4] https://en.wikipedia.org/wiki/Ben-Hur_(1959_film)
[5] https://en.wikipedia.org/wiki/Dae_Jang_Geum
[6] www.jeju.go.kr/mokkwana/index.htm
[7] https://assetstore.unity.com/packages/2d/textures-materials/sky/skybox-pack-1-14-skyboxes-19778
[8] https://assetstore.unity.com/packages/vfx/particles/environment/snow-vfx-58508
[9] https://filmora.wondershare.com

3.3 Procedure

This study and all procedures were approved by the Korea Advanced Institute of Science and Technology (KAIST) Institutional Review Board (KH2019-45). We recruited the 24 participants through the school bulletin board according to the complete counterbalancing method. When participants arrived, a research assistant explained the study, time commitment, and compensation. All participants were provided a written and informed consent before beginning the study. They completed questionnaires on a laptop using the online Google Forms. Before watching the video clip in VR, participants completed two questionnaires, the Visual Spatial Imagery [54] and the IRI Fantasy scale [14], to assess their spatial and empathetic ability. We used the result for removing outliers of the self-assessment result about the spatial presence or narrative engagement.

Figure 3: A participant with Google Daydream View and a paired controller.

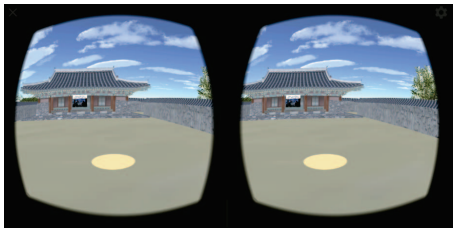


Figure 4: Video viewing point.

Participants wore a Daydream view with a built-in smartphone. The phone was a Samsung Galaxy S9 with the within VR app, and the ear headphones were plugged into the audio jack on the phone. The user could change their position and orientation by tilting and changing the direction of the head, using positional tracking features of the Google Daydream SDK. Also, they could play and stop the video clips with the controller. The research assistant introduced how to move in the virtual environment and select content using the HMD and the controller. During the experiment, to help the participant wear and remove the equipment, the research assistant stayed near the participant.

When the experiment started, participants were sitting on a chair as shown in Figure 3. They could look around the virtual Jeju-mok government office from the starting position. Then they moved to the yellow circle point marked on the floor, as shown in Figure 4. The reason for the move was to reenact the film tourism situation. The yellow point was set to view the video clip and the background place at once. Participants could move forward by tilting their heads slightly and stop moving by raising their heads. Although there was no limit on the time to move, participants could see the yellow point at a glance from the starting point, and the moving speed was the same. Thus, all participants moved to the viewing point within a similar movement pattern and time duration. The participant played the video clip with the controller and watched the same video clip in four virtual environments.

After watching the video clip in each condition, participants completed the questionnaires of spatial presence and narrative engagement on the laptop. Every participant experienced all conditions in a different order. The process took about 30 minutes. After finishing the experiment, they completed the Simulator Sickness Questionnaire(SSQ) [30] and the demographic information questionnaire. The simulator sickness was measured to confirm that simulator sickness differences between participants did not affect the experiment result. Lastly, we conducted an open-ended interview for about 10 minutes to understand their hyper-reality experience in each experimental condition. The total experiment time was approximately 60 minutes. After the experiment, participants received about $ 10 as compensation.

3.4 Participants

Of 24 participants, 17 identified as male and 7 as female, and 14 were undergraduate students, and 10 were graduate students. The ages ranged from 20 to 32 years (M=24, SD=2.851), and in nationality 23 participants were Korean and one was Chinese. The majority of the participants was quite familiar with VR head mounted display. 21 of them had VR experience, and 3 of them had not experienced VR at all. No participant had visited Jeju-mok government office before, which is the actual place of the virtual reality environment. The average score of the prior knowledge about the content, Dae Jang Geum, was less than three on the 5-point Likert scale (M = 2.37, SD = 1.469). In interviews, most participants said they were too young when the original series aired on TV, so they didn't remember the story in detail. However, most participants were well aware of the title and main actors because the series is famous. Before watching the video clip, participants completed two measures, the Visual Spatial Imagery (M=3.27, SD=.599) and the IRI Fantasy scale (M=4.90, SD=1.02). Also, after the experiment, they completed the Simulator Sickness Questionnaire (M=10.20 SD=1.07). The results of the Visual Spatial Imagery, the IRI Fantasy, and the Simulator Sickness uncovered no outlier among the participants. Thus, the final sample consisted of 24 participants.

4 Results

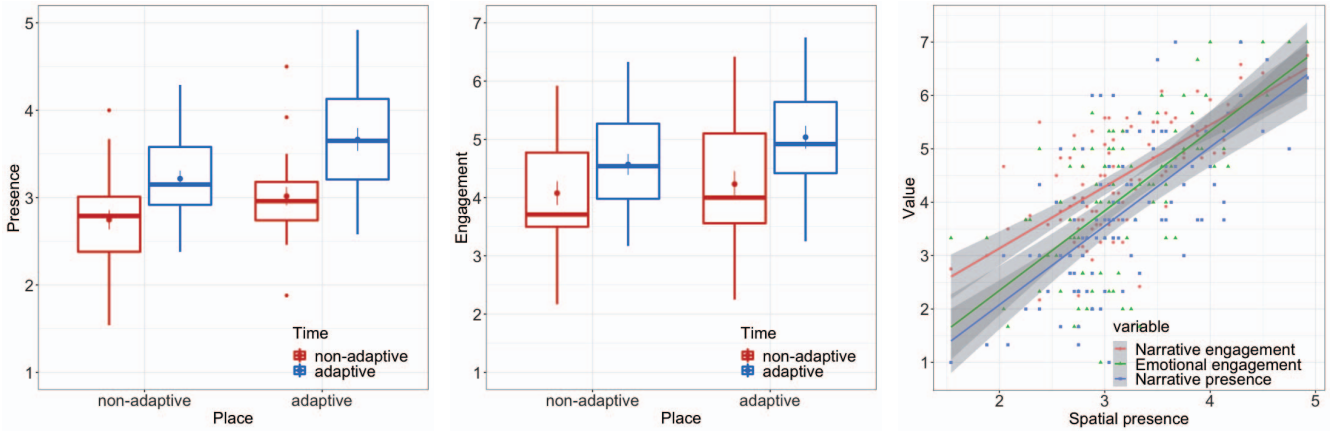
This section reports the analysis result of spatial presence, narrative engagement, correlations, and interview content.

4.1 Quantitative Analysis

We analyzed our data using repeated measures ANOVA. In a repeated-measures ANOVA, sphericity is an essential assumption as it refers to the condition where the variances of the differences between all possible pairs of within-subject conditions (i.e., levels of the independent variable) are equal. As the result of Mauchly's Test of Sphericity did not support the assumption, we did a conservative test with Greenhouse-Geisser, and Huynh-Feldt. To find the correlation between the spatial presence and narrative engagement, we used the Pearson correlation.

4.1.1 Spatial presence

As shown in Figure 5(a), the analysis revealed that the time factors had a significant main effect on the spatial presence with a large effect size; $F(1,23)$= 31.098, p <.05, η_p^2= .575. The spatial presence was significantly higher when time factors were applied to the virtual environment(M= 3.44, SD= .09) than when time factors were not reflected in the virtual environment (M= 2.88, SD= .10). The place factor also had significant main effect on the spatial presence with a large effect size; $F(1,23)$= 32.956, p <.05, η_p^2=.589. The spatial presence was significantly higher when the place factor was applied to the virtual environment(M= 3.34, SD= .09) than when



(a) Spatial presence (b) Narrative engagement (c) Correlation of spatial presence and narrative engagement

Figure 5: Spatial presence, narrative engagement, and correlation in four virtual environment conditions.

the place factor was not reflected in the virtual environment (M= 2.98, SD= .08). When comparing the main effects of each factor on spatial presence, the spatial presence was more improved when the temporal element was reflected in virtual reality than when the spatial element was reflected in virtual reality. However, we could not find any significant influence of interaction effects between time and place factors on the spatial presence; $F(1,23)$= 1.911, $p > .05$. Finally, the spatial presence was the highest in a fully adaptive virtual environment(M= 3.66, SD= .64) reflecting both time and place elements.

4.1.2 Narrative engagement

As shown in Figure 5(b), the time factors had a significant main effect on the narrative engagement with a large effect size; $F(1,23)$= 17.248, $p <.05$, η_p^2= .429. The narrative engagement was significantly higher when time factors were applied to the virtual environment(M= 4.80, SD= .16) than when time factors were not reflected in the virtual environment (M= 4.15, SD= .19). The place factor had significant main effect on the narrative engagement with a large effect size; F= 7.184, $p <.05$, η_p^2= .238. The narrative engagement was significantly higher when the place factor was applied to the virtual environment(M= 4.63, SD= .18) than when the place factor was not reflected in the virtual environment (M= 4.32, SD= .16). When comparing the main effects of each factor on narrative engagement, the narrative engagement was more improved when the temporal element was reflected in virtual reality than when the spatial element was reflected in virtual reality. However, we could not find any significant influence of interaction effects between time and place factors on the narrative engagement; $F(1,23)$= 1.865, $p > .05$. Finally, the narrative engagement was the highest in a fully adaptive virtual environment(M= 5.03, SD= .96) reflecting both time and place elements.

4.1.3 Correlation of spatial presence and narrative engagement

We analyzed a correlation of hyper-reality measurement variables. As shown in Figure 5(c), there was a positive correlation between the spatial presence and narrative engagement; r = .702, n = 96, p = .000. We also analyzed the correlation between the spatial presence and other sub-scales of narrative engagement. As a result, there was a positive correlation between the spatial presence and emotional engagement; r = .657, n = 96, p = .000. There was also a positive correlation between the spatial presence and narrative presence; r = .656, n = 96, p = .000.

4.2 Qualitative Analysis

We collected qualitative feedback through open-ended interviews to gain deeper insights into the proposed method for video representation. We adopted an inductive and grounded approach [12] and conducted a thematic analysis of our data [8]. The author who performed the interview aggregated interviewed content and converted it to Excel format to enable document analysis. The coder classified data according to each experimental condition. When a sentence contained opinions about several experimental conditions, we divided the sentence by conditions. In the process of data classification, organization, and interpretation, two authors and another VR researcher shared a note of comments involved in disagreements and different opinions. Through this process, we have secured 95 meaningful sentences for result analysis. We organized the classification results into themes and sub-themes as follows.

4.2.1 The effect of reflecting time setting of the story in virtual reality

When comparing the time and place settings, participants felt that the time factors applied to the virtual reality provided a similar atmosphere to the drama, enhancing their hyper-reality experience.

Similar atmosphere Participants felt night and snow created a similar atmosphere with the drama in the virtual environment. They thought it improved their spatial presence and narrative engagement. P4: *"The same atmosphere was created by the night and the snowfall, and that significantly influenced the sense of presence."* P6: *"When it was snowing at night, I felt that a similar atmosphere to the drama was created. It made me more immersive and empathetic to the drama."* P9: *"When elements that deepened the atmosphere of the drama, such as snow and night, were equally implemented in virtual reality, it seemed to be more immersive."*

Higher presence Participants felt they were in the fictional world of the TV show. P1: *"When time matched, I felt closer to the narrative world.* P13: *"When I watched the video in the night environment, I felt more that I was in the drama scene."* P15: *"When time coincided with the drama, I could feel more immersed and in a drama situation."*

Higher engagement Participants said they were better focused and engaged on the video in the time adaptive environment. P7: *"When virtual reality was in the same time zone as the video, I relatively well concentrated on the drama's story and felt the virtual environment more realistic."* P16: *"The attention was greater when the time coincided than when the place coincided with the drama."* P21: *"I felt that temporal differences had far more impact on engagement than spatial differences. When virtual reality was at night, whether the space is similar to a drama or not, I could focus on the drama emotionally from the start."*

4.2.2 The effect of reflecting place setting of the story in virtual reality

Participants stated that merely providing the background building where the main characters were sitting did not significantly affect their presence and engagement. The reason was that the temporal background and perspective were different. Besides, as their attention to the video decreased, participants negatively assessed virtual space realism.

Different atmosphere Participants said the different time and weather created the different atmosphere of virtual reality. It reduced their presence and engagement to the video content regardless of the background building. P8: *"The time background in the video was night, so I didn't feel that immersed in the daytime."* P19: *"My concentration on the drama decreased because of the different weather and atmosphere."* P24: *"The background building was similar to the video, but it was a bright day, so it felt like I was playing the video separately in a sunny place."*

Different viewpoint Participants felt that the presence and engagement decreased when they recognized that the perspective of the video was different from theirs. P11: *"Because the viewpoint kept changing depending on the camera position in the drama, It felt like the engagement was broken. I thought the place was where the drama scene was filmed, rather than feeling that the situation was happening at that location."* P16: *"Because I was standing in front of the building, my viewpoint was toward the characters' front side. So, I felt less concentration on the drama content when I saw the back of the characters in the video. If the characters were only shown in the front, it would have felt much more realistic."* P17: *"I didn't feel that the video and virtual reality environment had been matched. When the video's camera viewpoint looked at characters from inside of the building, I thought it would have been better if I could see them in the same direction."*

Negative realism assessment As participants' engagement and attention on the video were distributed to the surroundings, they negatively assessed virtual space realism. P9: *"The tree was shaking so hard that it bothered me. The building didn't play a significant role in engagement because it didn't seem to match the space in the drama."* P22: *"It seemed a little flawed to see dots(pixels) on display due to viewing the phone from a close distance."* P23: *"Because of the limitations of virtual reality technology to reproduce the world of the drama, I was not satisfied enough with the surrounding environment."*

4.2.3 The effect of reflecting time and place setting of the story in virtual reality

Participants stated that they felt the highest spatial presence and narrative engagement in the fully adaptive environment. In particular, they had a specific mental image of the world in which the character lived, and empathized with them.

Highest presence and engagement Participants said that they felt the most immersive hyper-reality experience in the fully adaptive environment. They also said that they could focus more on the TV show and better remember the story. P5: *"When space and time were all in sync, I felt like I was in the scene of the drama."* P15: *"The situation in which time and space were all matched with the drama was the most immersive."* P19: *"I think I was able to focus more on the video. Also, the contents of the drama remained in my memory for a longer time."*

Concrete mental image of the fictional world Participants said that the fully adaptive environment helped them understand the drama's spatial surroundings and improved the presence. P1: *"By the implementation of the surrounding environment, I felt like I was actually in the drama. Also, I could understand the background environment of the drama quickly because of the virtual space."* P23: *"When both time and place coincided, the virtual reality presented a concrete surrounding of the fictional world. It was a great help in understanding the world vividly in which the characters of the video live."*

Empathy to the characters Participants said the fully adaptive environment increased the overall narrative engagement and the emotional engagement to the characters. P10: *"The spatial and temporal match allowed me to empathize with the characters' situation and how they felt."* P21: *"I felt completely immersed in the drama in the virtual reality that applied all time and place settings of the drama. Beyond the feeling of being in the background of the drama, I felt like I was in the situation of the main characters as one of them."*

5 Discussion

According to the analysis of the quantitative and qualitative results, we confirmed that reflecting the time and place setting of the video content in virtual reality significantly improved the spatial presence and narrative engagement. In particular, we found that applying time settings was more effective in enhancing the hyper-reality experience than place settings. We also found a positive correlation between spatial presence and narrative engagement. In this section, we analyze the reason that time and place factors had a different degree of effect on the hyper-reality experience in virtual reality. We interpret the meaning of the positive correlation between the presence and narrative engagement. We propose design implications for effectively providing hyper-reality experiences in virtual reality.

5.1 Analysis on the Study Results

The previous studies have demonstrated that an emotional virtual environment offers a more engaging, immersive, and realistic experience to the user than a neutral environment. For example, Baños et al. compared two virtual environments, one involving emotional content and the other involving neutral content [3]. Riva et al. compared two emotional virtual environments and one neutral virtual environment [43]. Both studies demonstrated that affective content had an impact on presence. In these studies, they manipulated sound, music, shadow, lights, and textures to provide different emotional conditions in the same virtual environment. In particular, they produced the emotional environment and induced different moods by adjusting the light parameters such as tone, direction, and brightness. Our study also created the time adaptive environment by changing light parameters and adding weather effects in the basic virtual environment. Thus, it can say that the temporal element created an emotional environment for the drama's narrative. Participants felt a higher sense of presence by experiencing a higher emotional state to the drama story in the time adaptive environment.

Participants also felt greater attention and narrative engagement in the time adaptive environment. These findings are consistent with previous studies showing that attention is biased toward the emotional stimuli [38, 42]. Attention focuses mostly on an emotionally arousing stimulus, resulting in better engagement and memory for the stimuli [21, 23]. Therefore, we can say that the time adaptive environment served as emotional stimuli, thereby improving attention, narrative engagement, even memory about video content.

On the other hand, the place adaptive environment did not adjust the factors that affect the overall atmosphere of the virtual reality environment. It only added one building behind the video. The place adaptive environment was close to a neutral condition compared to the time adaptive environment. Previous studies have shown that when virtual reality does not provide an emotional environment, the presence decreases, and the negative realism evaluation increases. In our study, participants also felt less presence or engagement in the place adaptive environment. As attention to the video decreased, they negatively evaluated the virtual reality environment around the video and display quality. Therefore, the experimental results show that simply visualizing the building where the character was sitting did not provide an emotional virtual reality related to the story.

Furthermore, the building emphasized different viewpoints that looked at characters and space between the video and participants. When the video began, a participant was looking at the main characters in front of the building in our stimulus. It was the same as the camera's point of view. However, as the stories progressed, the camera took the characters from the inside and next to the building. On the contrary, the user's perspective was fixed. Thus, participants could be more aware of the different perspectives of seeing characters between video and them. As a result, the background building did not increase the presence of the story world as much as we expected. It only gave participants the perception that they watched the video in front of the place where the story happened.

Finally, we found a positive correlation between spatial presence and narrative engagement, including emotional engagement and narrative presence. These results are consistent with previous studies showing presence has positive correlations with narrative engagement [7] and empathy [50] in CVR. Through the result, we confirmed that the spatial presence and narrative engagement significantly interact with each other not only when experiencing new CVR content but also when viewing 2D videos of existing dramas and movies in 3D virtual reality. Specifically, when participants felt a higher spatial presence, they perceived the narrative space more concretely and vividly and experienced a higher emotional engagement with the characters.

5.2 Design Implications

Based on the analysis of our experimental results, we propose design implications to implement the content-adaptive virtual reality for video content and give an immersive hyper-reality experience as follows. The video 'scene' unit indicates a series of situations or events occurring at the same time and place [27]. As the scene unit can divide and deliver video content semantically based on the location, we consider scene-based video clips are suitable to provide location-based content service in virtual reality. In the meantime, as the metadata of broadcasting and movie content has been generated manually using authoring tools, it has been difficult to build a large database, and errors have occurred. However, recently, various technologies, such as content recognition, automatic tagging, and topic model learning, have automatically generated metadata of scenes or topic [34]. Along with various text data such as scripts, captions, and call sheets[10], these technologies can provide information about the elements that make up the video scene, such as time, location, characters, props, special effects, and background music. Therefore, it is possible to implement advanced content-adaptive virtual reality by simultaneously providing video and related multimedia content based on scene metadata.

The content-adaptive virtual reality should automatically adjust the light parameters and weather assets based on the scene metadata when playing video clips. Through this, viewers can experience the same atmosphere as the narrative and feel a higher spatial presence and narrative engagement to the fictional world. Also, viewers should be able to see characters and spaces from the camera's position. This allows viewers to experience a more immersive narrative world based on the place elements. In particular, for switching viewpoints, it is necessary to apply diegetic guidance [10], such as moving objects related to the story, rather than the forced rotation technique [39, 52]. The reason is that in most movies and dramas, cut transitions take place in seconds. If the user's viewpoint is automatically switched according to the camera's position, it can result in motion sickness, and viewers may feel uncomfortable or strenuous [51]. The narrative setting of video can be expanded to interactive 3D virtual reality through the proposed design implications. The user can experience the more immersive hyper-reality through it.

[10] A daily schedule of filming for a movie or television show

5.3 Limitations and Future Work

As the limitation, this paper used one video content as the stimulus to examine the content-adaptive virtual environment's effect on the narrative engagement or spatial presence. However, it is necessary to divide the degree of difference between the video's time and place setting and the virtual reality environment. Through this, we should investigate the different effects of reflecting environmental settings of the narrative to the virtual reality. In future work, we will find the most optimized video content representation method considering the various difference between the story's environment and the virtual reality environment. We will also investigate suitable diegetic guidance methods to guide users to the point of view for watching the 2D video.

6 Conclusion

This paper investigated a video content representation method to support the hyper-reality experience in virtual reality. We proposed the content-adaptive virtual reality that reflects the time and place setting of video content. We revealed that reflecting the environment setting of the video content to the virtual reality significantly improves the spatial presence and narrative engagement. We also demonstrated a significant correlation between spatial presence and narrative engagement. Based on the result, we analyzed how time and place elements influenced spatial presence and narrative engagement. We discussed design implications for providing a fully adaptive environment and an immersive hyper-reality experience based on time and place elements in virtual reality.

Acknowledgments

The authors thank Junki Kim and Taesoo Jang for helping to implement the system for the user study. This work was supported by Institute of Information communications Technology Planning Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2019-0-01270, WISE AR UI/UX Platform Development for Smartglasses).

References

[1] T. Aitamurto, S. Zhou, S. Sakshuwong, J. Saldivar, Y. Sadeghi, and A. Tran. Sense of presence, attitude change, perspective-taking and usability in first-person split-sphere 360 video. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 545. ACM, 2018.

[2] M. Badri, M. Pakanen, P. Alavesa, H. Kukka, and T. Ojala. Design, development, and usability evaluation of a system for adding and editing social media banners in the immersive street-level 3d virtual city. In *Virtual Worlds and Games for Serious Applications (VS-Games), 2017 9th International Conference on*, pp. 102–108. IEEE, 2017.

[3] R. M. Baños, C. Botella, M. Alcañiz, V. Liaño, B. Guerrero, and B. Rey. Immersion and emotion: their impact on the sense of presence. *Cyberpsychology & behavior*, 7(6):734–741, 2004.

[4] J. Baudrillard. *Simulacra and simulation*. University of Michigan press, 1994.

[5] S. Beeton. *Film-induced tourism*. Channel view publications, 2016.

[6] M. Billinghurst, H. Kato, and I. Poupyrev. The magicbook-moving seamlessly between reality and virtuality. *IEEE Computer Graphics and applications*, 21(3):6–8, 2001.

[7] S. W. Bindman, L. M. Castaneda, M. Scanlon, and A. Cechony. Am i a bunny?: The impact of high and low immersion platforms and viewers' perceptions of role on presence, narrative engagement, and empathy during an animated 360° video. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, p. 457. ACM, 2018.

[8] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2):77–101, 2006.

[9] R. Busselle and H. Bilandzic. Measuring narrative engagement. *Media Psychology*, 12(4):321–347, 2009.

[10] C. Cao, Z. Shi, and M. Yu. Automatic generation of diegetic guidance in cinematic virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 600–607. IEEE, 2020.

[11] D. Carl, S. Kindon, and K. Smith. Tourists' experiences of film locations: New zealand as 'middle-earth'. *Tourism Geographies*, 9(1):49–63, 2007.

[12] J. Corbin and A. Strauss. *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage publications, 2014.

[13] N. Couldry. The view from inside the'simulacrum ': visitors' tales from the set of coronation street. *Leisure Studies*, 17(2):94–107, 1998.

[14] M. H. Davis. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology*, 44(1):113, 1983.

[15] R. Du, D. Li, and A. Varshney. Geollery: a mixed reality social media platform. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI), CHI*, p. 13, 2019.

[16] R. Du and A. Varshney. Social street view: blending immersive street views with geo-tagged social media. In *Web3D*, pp. 77–85, 2016.

[17] U. Eco. *Travels in hyper reality: Essays*. Houghton Mifflin Harcourt, 1990.

[18] T. Edensor. Performing tourism, staging tourism: (re) producing tourist space and practice. *Tourist studies*, 1(1):59–81, 2001.

[19] A. El Saddik. Digital twins: The convergence of multimedia technologies. *IEEE multimedia*, 25(2):87–92, 2018.

[20] A. Elmezeny, N. Edenhofer, and J. Wimmer. Immersive storytelling in 360-degree videos: An analysis of interplay between narrative and technical immersion. *Journal For Virtual Worlds Research*, 11(1), 2018.

[21] P. A. Gable and E. Harmon-Jones. Approach-motivated positive affect reduces breadth of attention. *Psychological Science*, 19(5):476–482, 2008.

[22] K. Gil, J. Rhim, T. Ha, Y. Y. Doh, and W. Woo. Ar petite theater: Augmented reality storybook for supporting children's empathy behavior. In *Mixed and Augmented Reality-Media, Art, Social Science, Humanities and Design (ISMAR-MASH'D), 2014 IEEE International Symposium on*, pp. 13–20. IEEE, 2014.

[23] E. Harmon-Jones and P. A. Gable. Neural activity underlying the effect of approach-motivated positive affect on narrowed attention. *Psychological Science*, 20(4):406–409, 2009.

[24] C. Heeter. Being there: The subjective experience of presence. *Presence: Teleoperators & Virtual Environments*, 1(2):262–271, 1992.

[25] J. Y. Hong, B. Lam, Z.-T. Ong, K. Ooi, W.-S. Gan, J. Kang, J. Feng, and S.-T. Tan. Quality assessment of acoustic environment reproduction methods for cinematic virtual reality in soundscape applications. *Building and environment*, 149:1–14, 2019.

[26] Y. C. Huang, K. F. Backman, S. J. Backman, and L. L. Chang. Exploring the implications of virtual reality technology in tourism marketing: An integrated research framework. *International Journal of Tourism Research*, 18(2):116–128, 2016.

[27] J. Hunter and L. Armstrong. A comparison of schemas for video metadata representation. *Computer Networks*, 31(11-16):1431–1451, 1999.

[28] B. Jewell and S. McKinnon. Movie tourism—a new form of cultural landscape? *Journal of Travel & Tourism Marketing*, 24(2-3):153–162, 2008.

[29] B. R. Jones, H. Benko, E. Ofek, and A. D. Wilson. Illumiroom: peripheral projected illusions for interactive experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 869–878, 2013.

[30] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993.

[31] S. Kim. Extraordinary experience: Re-enacting and photographing at screen tourism locations. *Tourism and Hospitality Planning & Development*, 7(1):59–75, 2010.

[32] S. Kim. A cross-cultural study of on-site film-tourism experiences among chinese, japanese, taiwanese and thai visitors to the daejanggeum theme park, south korea. *Current Issues in Tourism*, 15(8):759–776, 2012.

[33] S. Kim and H. Wang. From television to the film set: Korean drama daejanggeum drives chinese, taiwanese, japanese and thai audiences to screen-tourism. *International Communication Gazette*, 74(5):423–442, 2012.

[34] S.-H. Kim, D. Jung, S.-Y. Lee, and S.-J. Kim. Scene-based metadata generation and open api provisioning method for smart broadcast service. In *2018 20th International Conference on Advanced Communication Technology (ICACT)*, pp. 664–668. IEEE, 2018.

[35] H. Kukka, M. Pakanen, M. Badri, and T. Ojala. Immersive street-level social media in the 3d virtual city: Anticipated user experience and conceptual development. In *CSCW*, pp. 2422–2435, 2017.

[36] A. MacQuarrie and A. Steed. Cinematic virtual reality: Evaluating the effect of display type on the viewing experience for panoramic video. In *Virtual Reality (VR), 2017 IEEE*, pp. 45–54. IEEE, 2017.

[37] L. T. Nielsen, M. B. Møller, S. D. Hartmeyer, T. Ljung, N. C. Nilsson, R. Nordahl, and S. Serafin. Missing the point: an exploration of how to guide users' attention during cinematic virtual reality. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, pp. 229–232. ACM, 2016.

[38] A. Öhman, A. Flykt, and F. Esteves. Emotion drives attention: detecting the snake in the grass. *Journal of experimental psychology: general*, 130(3):466, 2001.

[39] Y. Onuki and I. Kumazawa. Reorient the gazed scene towards the center: Novel virtual turning using head and gaze motions and blink. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1864–1871. IEEE, 2019.

[40] H. Park, J. Kim, S. Bang, and W. Woo. The effect of applying film-induced tourism to virtual reality tours of cultural heritage sites. In *2018 3rd Digital Heritage International Congress (DigitalHERITAGE) held jointly with 2018 24th International Conference on Virtual Systems & Multimedia (VSMM 2018)*, pp. 1–4. IEEE, 2018.

[41] A. Portegies. Places on my mind: Exploring contextuality in film in between the global and the local. *Tourism and Hospitality Planning & Development*, 7(1):47–58, 2010.

[42] L. Riggs, D. A. McQuiggan, N. Farb, A. K. Anderson, and J. D. Ryan. The role of overt attention in emotion-modulated memory. *Emotion*, 11(4):776, 2011.

[43] G. Riva, F. Mantovani, C. S. Capideville, A. Preziosa, F. Morganti, D. Villani, A. Gaggioli, C. Botella, and M. Alcañiz. Affective interactions using virtual reality: the link between presence and emotions. *CyberPsychology & Behavior*, 10(1):45–56, 2007.

[44] S. Rothe, D. Buschek, and H. Hußmann. Guidance in cinematic virtual reality-taxonomy, research status and challenges. *Multimodal Technologies and Interaction*, 3(1):19, 2019.

[45] S. Rothe and H. Hußmann. Guiding the viewer in cinematic virtual reality by diegetic cues. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*, pp. 101–117. Springer, 2018.

[46] S. Rothe, H. Hußmann, and M. Allary. Diegetic cues for guiding the viewer in cinematic virtual reality. In *Proceedings of the 23rd ACM symposium on virtual reality software and technology*, pp. 1–2, 2017.

[47] S. Rothe, B. Kegeles, and H. Hussmann. Camera heights in cinematic virtual reality: How viewers perceive mismatches between camera and eye height. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, pp. 25–34, 2019.

[48] S. Rothe, A. Schmidt, M. Montagud, D. Buschek, and H. Hußmann.

Social viewing in cinematic virtual reality: a design space for social movie applications. *Virtual Reality*, pp. 1–18, 2020.
[49] M.-L. Ryan. *Narrative as virtual reality 2: Revisiting immersion and interactivity in literature and electronic media*, vol. 2. JHU Press, 2015.
[50] D. Shin. Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior*, 78:64–73, 2018.
[51] M. Speicher, C. Rosenberg, D. Degraen, F. Daiber, and A. Krúger. Exploring visual guidance in 360-degree videos. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*, pp. 1–12, 2019.
[52] T. Stebbins and E. D. Ragan. Redirecting view rotation in immersive movies with washout filters. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 377–385. IEEE, 2019.
[53] M. Toolan. *Narrative: A critical linguistic introduction*. Routledge, 2012.
[54] P. Vorderer, W. Wirth, F. R. Gouveia, F. Biocca, T. Saari, L. Jäncke, S. Böcking, H. Schramm, A. Gysbers, T. Hartmann, et al. Mec spatial presence questionnaire. *Retrieved Sept*, 18:2015, 2004.
[55] A. Waysdorf and S. Reijnders. The role of imagination in the film tourist experience: The case of game of thrones. *Journal of Audience and Reception Studies*, 14(1):170–191, 2017.