STUDY ON AUGMENTED CONTEXT INTERACTION SYSTEM FOR VIRTUAL REALITY ANIMATION USING WEARABLE TECHNOLOGY

Jiyoung Kang[1] and Jongkuk Lim[2]

[1]Graduate School of Cinematic Content, Dankook University, Korea
[2]Department of Computer Engineering, Dankook University, Korea

ABSTRACT

Currently, the virtual reality market is receiving global attention, and VR animation entertainment content is attracting attention as a new content market. In this study, we successfully develop these kinds of VR animations by focusing on interactions, the largest feature of virtual reality content, and we present a novel interaction method for virtual reality animations. We created ".FLY," a virtual reality movie, using wearable technology, and designed augmented context interactions to verify the emotional empathy and emotional immersion effect of the active story-based interactions experienced by the user in a virtual reality environment. The wearable bands used in the interactions were trained through programming, starting at the movie development stage. The audience viewed the interactive virtual reality movie and actively interacted through gesture interactions using the wearable bands.

KEYWORDS

VR, Interactive Animation, Augmented Context Interaction, Wearable Technology, Emotional Immersion

1. INTRODUCTION

Currently, as the world heads toward a fourth industrial revolution, there is considerable interest in the field of virtual reality. Virtual reality technology began in 1940 with flight simulators and has been continuously studied since then. In recent years, as wearable virtual reality devices such as HMDs (Head Mounted Displays) have been commercialized owing to advances in related hardware and software, virtual reality is garnering attention as the content platform of the future. However, the virtual reality market, which was expected to grow rapidly, has actually experienced faltering growth. This is because companies have not yet been able to establish a virtual reality ecosystem capable of expanding the virtual reality market due to the imbalances of a market developed with a focus on only hardware and related technology. In particular, the lack of content for a rapidly growing market centered on virtual reality devices is emerging as a problem in the virtual reality market. Because of this, the world is attempting all possible to dominate the virtual reality content market. In particular, games, which are rapidly converging

Dhinaharan Nagamalai et al. (Eds) : ITCSE, WiMo, ICAIT, ICDIPV, CRYPIS - 2018
pp. 47–58, 2018. © CS & IT-CSCP 2018 DOI : 10.5121/csit.2018.80705

with the virtual reality content market, are already expanding the market through specialized game distribution platforms such as Steam, and high-quality virtual reality games are being released. In addition, virtual reality movies are being actively developed and introduced at international conferences, movie festivals, and in Hollywood, including at SIGGRAPH, the Sundance Film Festival, and the Cannes Film Festival. In 2017 Google's "Pearl" became the first VR animation to be nominated for an Academy Award [1], and Penrose Studios' "Arden's Wake" was awarded Best VR Film at the 74th Venice Film Festival [2], showing the full potential of VR animation.

VR animation makes users emerge in a virtual space due to a combination of factors such as physical and task environment [3]. Physical environment produces a state of immersion and the task environment produces a state of involvement [4]. These environments can consist of combined hardware platform, software systems and interactive scenarios. There have been some researches about affective and cognitive elements, or endogenous factors, that contribute to an immersive experience and there is much emerging support within the literature for a relationship between presence and immersive experiences [5] [6] [7]. Researchers have also looked at a sense of flow, or deep involvement, with a virtual environment and immersion [8] [9] [10].

With the development of sensor technology, content that utilizes the motion of two hands or the whole body in real-time interaction with HMDs(Head Mounted Display) [11] while using wearable technology have recently appeared, yet it is difficult to find a case that utilizes real-time interaction in animation films. Therefore, this study proposes augmented context interaction using a wearable band in VR animation films based on interactive narratives. This study is concerned with the creation of immersive VR animations, and it analyzes the difference in storytelling between existing traditional movies and virtual reality movies with a focus on augmented context interaction. We aim to study the emotional effects of augmented context interaction in virtual reality movies that use wearable technology.

2. RESEARCH SCOPE AND METHODS

Virtual reality is a technology in which the users can have real-time interaction in a virtual space created by computer systems. It is a convergent technology that allows users to feel immersed in this virtual space through the five senses of the human body and that provides a feeling of presence as if the user actually existed in that space [12]. There are a variety of types of virtual reality content including CAVE type, desktop type, and third person type, but in this study we focus on wearable display HMD-based virtual reality movies, which are the most appropriate for personal media and growing at the fastest speed [13].

Virtual reality games targeted at existing game users are quickly becoming plentiful among HMD-based game content. However, movies, which are one of the most popular forms of entertainment, have not been able to form a virtual reality movie market as quickly as games. This is partly because HMD devices are associated with several problems such as cost and motion sickness from the cognitive dissonance caused by resolution and lag speed; however, the most important reason why audiences do not seek virtual reality movies is the lack of compelling content. Because the characteristics of virtual reality movies differ from those of existing movies, there is a need for movie storytelling based on deep research into these characteristics. In this study, we aim to analyze these virtual reality movie characteristics and propose an augmented context interaction method that uses wearable technology suited to these characteristics. In addition, we perform experiments to verify the effect of the audience's emotional reaction on gesture interactions.

3. CURRENT VR FILM TECHNIQUE

3.1. Virtual Reality Movies and Presence

As with traditional 2D screen-based movies, virtual reality movies can largely be divided according to the methods by which they are made: live-action movies, graphics-based movies, and movies that were made by using a combination of these two methods. To date, research on these filmmaking methods made up most of the mainstream research on virtual reality movies, and research was not performed on methods of storytelling in virtual reality movies. This is because virtual reality movies are seen as a movie genre that is derived from the existing movie medium; thus, people have been unable to sense the need for a special grammar or type of storytelling that belongs only to virtual reality movies. However, when virtual reality movies borrow this kind of traditional movie storytelling style as-is it is difficult to properly show audiences the advantages held solely by virtual reality movies. Virtual reality movies that are produced with the same camera perspectives, etc. as the linear narratives of traditional movies can be called incomplete forms of virtual reality content that only add 360° imagery to traditional movies. However, we need a movie experience that is fundamentally different from existing traditional movies to enable audiences to feel the attraction of true virtual reality movies.

The ultimate goal of virtual reality is to allow the user to experience telepresence. Originally presence meant the sense of being within a certain environment, while telepresence can be called the experience of existing within a certain environment through a communications medium i.e., the mediated perception of an environment [14]. Feeling total presence through virtual reality movies is not possible simply through offering the point of view of an observer who turns their head around in a 360° space and sees a virtual space. That is, in order for the audience to feel this kind of presence in a virtual reality movie, we require communication beyond the passive communication of simply turning one's head. In short, we need mental and physical interactions between the audience and the medium called film. To do this, this paper proposes classifying this kind of communication as augmented context interaction.

3.2 Related Work

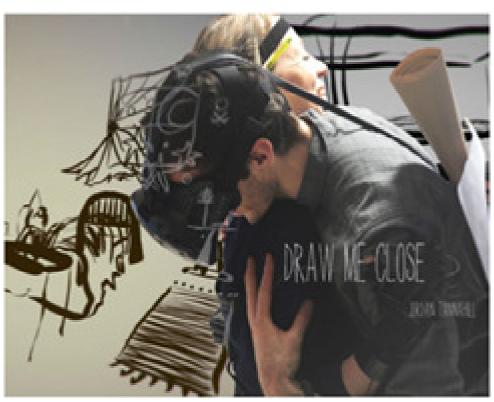

Figure 1: User interacting with an actor in the VR Film 'Draw Me Close'

Figure 2: Scene from the VR Film 'Draw Me Close'

In the spring of 2017, "Draw Me Close" [15] debuted at the Tribeca Film Festival, and it can be considered a typical example of active communication between an audience and a virtual reality movie (Fig. 1). The audience, equipped with an HMD and hand sensors, experiences 5-year-old Jordan's memories of his mother who passed away from cancer. People in the audience are equipped with an HMD and a motion capture system and experience something beyond direct participation in the space and time of the movie. The virtual space experienced through the HMD is shown in Fig.2, where we see Jordan's room and his mother expressed through line animation. In the virtual reality space, Jordan is lying down, and his mother pats him and hugs him as she puts him to sleep. There is active bodily interaction with the audience. Through this kind of embodied interaction, the audience feels immersion toward Jordan in the virtual space, and emotional storytelling takes place. However, in order to have this experience, the audience and the actors all have to wear special equipment to be motion captured, and there are problems with animation frame delays, etc. in the process of creating the real-time animation through the motion capture system. Moreover, a computer system with high specifications is needed to connect to the Vive HMD, and building the movie system is complicated.

Figure 3: Scene from VR Animation "Tree"

Active interaction in this kind of virtual reality animation can also be experienced in "Tree" [16], which debuted at the 2017 Sundance Film Festival and can be seen in Fig. 3. In this movie, members of the audience wear an Oculus Rift HMD and hold MIT-developed EMS (Electronic Muscle Stimulation) sensors in both hands and experience becoming a tree in the tropical rain forest. By bodily experiencing the whole cyclical process of the tree being born, growing up, and returning to the dirt, the audience has the mysterious experience of becoming part of nature. However, this movie does not provide the audience with emotional interactions that go beyond the intuitive experience of having the movement of their hands become the branches of the tree.

Conversely, due to the current limitations of technology, the disadvantages of such hand gesture inputs are: first, lack of comfort; second, every gesture can be interpreted by the system, whether or not it is intended, so the system must have well-defined means to detect the intention of the gesture; and finally, segmentation of hand gestures. In order to solve these, this study proposes augmented context interaction, which is suitable for VR animation stories experienced through wearable bands which the audience can easily put on. Through augmented context interaction, members of the audience, who easily put on the wearable bands without any special equipment and are in a free-hand state, can experience emotions and sympathy with the characters in a VR environment.

4. VIRTUAL REALITY ANIMATION USING WEARABLE TECHNOLOGY

4.1. Interactive VR Animation

Figure 4: Interactive VR Animation ".FLY"

In order to verify the emotional effects of these kinds of active communication in virtual spaces, this study created an interactive VR animation named ".FLY" [17]. ".FLY" is a virtual reality movie in which the audience interacts directly through gesture interactions using augmented content interaction, and it was shown by invitation in the Busan Film Festival's VR Theater in October 2017. Each member of the audience wore an HMD that incorporates a smartphone and wearable bands on the wrists that can recognize gestures to encounter a young girl within the movie. From a first person's view, the audience can either refuse the girl's hand as she asks for help or they can take her hand. In addition, they can console the girl when she collapses by patting her. Emotional storytelling occurs through these three different active bodily interactions.

4.2. Gesture Recognition Via Wearable Band

In order to recognize this kind of audience interaction in real-time as the movie plays, we created wearable bands that can be placed on people's wrists. According to the definition of the MIT Media Lab, a wearable device refers to anything that is attached to the body or worn and which can perform computing actions, including applications capable of performing some computing functions [18]. By placing wearable bands on the audience's wrists, we could immediately and naturally gather the movements made by their wrists as they watched the movie.

In order to recognize the audience's interactions as the movie plays in this way, we required the audience to put on wearable bands equipped with IMU (Inertial Measurement Unit) and watch the movie. IMU increases user convenience by combining a gyro and an accelerometer, and it contributes notably to raising the reliability and efficiency of the system [19]. Equipping the hand with a 6 degree-of-freedom sensor, we can map its position onto the position of a virtual hand in an immersive world and create the illusion that users can touch and grasp the girl in the VR animation using their own hands [20]. This allows users to interact naturally and intuitively in the immersive environment.

In order to provide gesture interactions to the audience in real-time as they watch the movie, it is necessary to follow a process of designing gesture interactions that are appropriate for the story and recognizing the designed interactions during the planning stage of the movie. To do this, we first gathered acceleration and gyroscope sensor data from the bands worn on the wrists, as in Fig. 5. Afterward, we selected features from the collected sensor data from which to distinguish actions. For example, distinguishing two kinds of actions requires us to calculate the speed, continuous motion values, etc. of the action as features. Subsequently, we implemented a methodology for distinguishing actions from the corresponding features through programming.

We implemented a rule-based gesture recognition system. Our interaction system design has the simplest form to minimize a computation burden. Accelerometer and gyroscope sensor are used in this system. First, we extract features as following, x, y, z, magnitude, minV, maxV. First three features, x, y, and z, are raw data from the sensor. Magnitude is defined as $\sqrt{x^2 + y^2 + z^2}$. Lastly, mixV and maxV are computed with each axis of sensor data. It is defined as $minV = [minX, minY, minZ]$, $maxV = [maxX, maxY, maxZ]$ and $minX = \min(minX, x_t)$.t is time and rest features are equivalently computed as minX. However, minV and maxV are initialized to 0 when the user attempts to start gesture interaction to avoid unintentional gesture. Features are computed with each sensor separately which generates 12 features. Every gesture can be divided into two steps First it waits until magnitude value exceeds threshold to recognize whether user has intention to make a gesture. Then it performs the actual recognition process. Recognizing holding a hand gesture exploits x, y, z values of accelerometer. It is watched if user holds the pre-defined angle for 3 seconds. Stroking hair exploits minV and maxV values of accelerometer and gyroscope. Thresholds of minimum and maximum are used to filter when the user tries to transit hand movement direction. Refusing a hand gesture can be thought of as a sequence of actions followed by holding and withdrawing a hand.

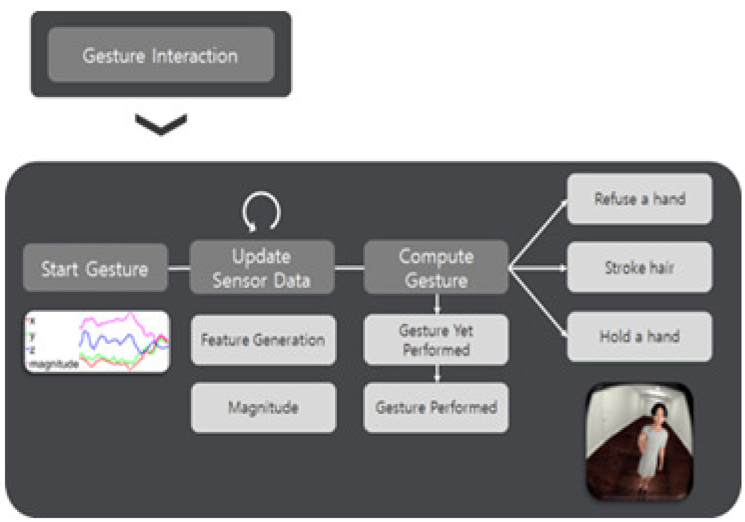

Figure 5: Process of gesture recognition

4.3. Augmented Context Interaction

Along with the recent growth of virtual reality content, continuing research has been focused on the interaction between users and content in order to provide the user with experiences optimized for interactions with content. In this research, there is interest in the extent to which the user can use the content in a natural way, and this is creating a turning point in the development of natural user interface (NUI) technology. In the most typical method of NUI, a voice is played, and currently this is used in many products and services. Along with voice, gestures are an often-used user interface. Gestures include not only the intentional actions a user undertakes to convey their meaning, but also the unconscious, meaningless actions that they make.

In gesture recognition technology, sensors or equipment can be broadly divided into the contact type, in which data are acquired as the user directly makes contact via their body, and the non-contact type, in which data are acquired using long-distance and short-distance sensors. In this study we used the contact type of gesture recognition, which utilizes wearable technology worn on the user's wrists. This is because this equipment is very well suited to virtual reality environments because it can detect detailed gestures by simply being worn, without any separate camera equipment or limits on the user's area of activity or space.

When designing the interactions for this virtual reality movie, we used a concept known as augmented context interaction to move the audience's hearts and naturally induce interactions from the audience. This means augmenting the context information contained in the story, characters, or scenarios of the movie for an interaction given to the audience within the movie such that the interaction can be performed without resistance. That is, augmented context interaction is provided such that interactions can be performed naturally such as when the audience naturally reaches out as the girl, who is in danger, reaches out and asks for help or when they stroke her hair after she collapses on the cold street.

The first definition of context-aware applications, provided by Schilit and Theimer [21], ranged from applications that are simply informed about context to applications that adapt themselves to context. Context-awareness has been described in various research efforts with the following language: adaptive [22], reactive [23], responsive [24], situated [25], context-sensitive [26], and environment directed [27]. Previous research also described context-aware computing in two main ways: using context and adapting to context. In this study we used adapting to context and proposed gesture interactions suitable to scenarios in which there is interaction with the story and characters of the animation. This enables the audience to have augmented bodily interactions with characters that were limited in the virtual reality animation environment. Even though these interactions are not felt directly through the skin as in the real world, the act of stroking the character in the virtual space with one's own hand or holding her hand is directly experienced; thus, it is possible to allow the audience to feel more immersion and emotional storytelling.

Performing this kind of augmented context interaction design requires us to make it possible for the audience to naturally recognize what kind of gestures they must perform based on their past experiences. Ulmer and Ishii call this "Expressive Representation (ER)," which is when the user already reads and interprets the circumstances during the interaction and devises behaviors and modifications [28]. That is, the audience understands and analyzes the situation within the movie through this kind of active interaction, enabling meaning to be created by each user.

Along with this, the movie introduces the audiences to virtual reality, makes a life exist within the movie, and causes direct bodily interactions. This is because the VR environment is the optimal environment for tracking the audience's potential movements, and it can provide richer interactions than other media [29]. According to previous research, the user has a deeper immersion effect toward content when performing bodily interactions in a virtual space than when viewing video images [30]. This means that if suitable bodily interactions are used in harmony with the story in a virtual reality movie as in the examples above, the audience's emotional immersion is raised, and a greater degree of immersion occurs.

Figure 6: Girl reaching out to the viewer

Along with this, the movie introduces the audiences to virtual reality, makes a life exist within the movie, and causes direct bodily interactions. This is because the VR environment is the optimal environment for tracking the audience's potential movements, and it can provide richer interactions than other media [29]. According to previous research, the user has a deeper immersion effect toward content when performing bodily interactions in a virtual space than when viewing video images [30]. This means that if suitable bodily interactions are used in harmony with the story in a virtual reality movie as in the examples above, the audience's emotional immersion is raised, and a greater degree of immersion occurs.

5. RESULTS

We conducted experiments to confirm that the active communication achieved through augmented context interaction caused emotional empathy by directly inviting the audience into the virtual reality space and providing an embodied experience. The users included 20 men and women between the ages of 20 and 40. In the tests, the users were divided into two groups of 10 people: user group A wore the wearable bands and had active interactions, and user group B watched the VR movie and did not have active interactions. After viewings, we evaluated the users' emotional immersion and presence through a survey.

Question 1 measures the emotion of sympathy felt by the audience when experiencing empathy toward the story as they watched the movie and encountered the girl from the main character's 1st-person point of view. Through this question, we aimed to evaluate the audience's emotional empathy effect. Questions 2 evaluated whether the audience felt as though they existed within the virtual space and the degree of presence they felt interacting with the girl character in the movie. We used these questions with the aim of evaluating the immersion effect.

1) If you felt sympathy or pity toward the girl as you watched the movie, to what extent did you have these feelings? (5 is the highest level of sympathy)

2) How much immersion did you feel toward the main character in the movie? (5 is the highest level of immersion)

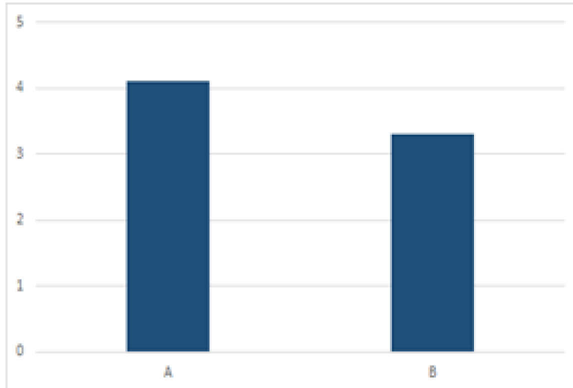

Figure 7: Result for empathy

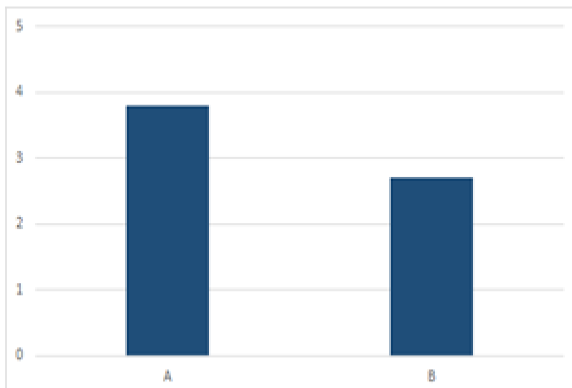

Figure 8: Result for immersion

Through the above survey we found that user group A, which wore the wearable bands and experienced active interaction, had more sympathetic emotions toward the girl in the movie at a score of 4.1 than user group B which did not experience gesture interactions and had a score of 3.3. Furthermore, group A's immersion as the main character of the movie was higher at 3.8 than group B's, which was 2.7. It can be seen that the users of group A, who wore the wearable bands and directly performed gesture interactions according to the given situation in the virtual reality environment, felt considerable emotional empathy and a sense of immersion compared to when gesture interactions were not performed during viewing. This means that when the audience is immersed in a virtual reality environment and views a movie, if augmented context interaction is properly used in a way that is suitable for the story and scenarios of the movie, it can strengthen the immersion and emotional storytelling of the virtual reality content.

6. CONCLUSIONS

We described research aimed at developing a method for virtual reality animation storytelling that combines a form of active communication known as augmented context interactions with virtual reality animation. We provided audiences with augmented context interaction by using wearable bands to provide gesture interactions that are emotionally and bodily suitable for the story and scenarios of a virtual reality movie. We also conducted experiments to confirm that the active communication achieved through augmented context interaction caused emotional empathy and immersion.

For the next step, we will perform user tests to verify that the degree of emotional immersion in an audience experiencing augmented context interactions compare to passive traditional interactions. Also, in this study, we performed research on augmented context interaction in virtual reality movies, and we must perform a variety of research on stories and dramatic methods suitable for this kind of interaction. Many tasks remain in the research and development of technology and content related to virtual reality movies to enable this kind of virtual reality animation to take a place in the mainstream cultural content industry

ACKNOWLEDGEMENTS

This research was conducted with the support of the 2017 National Research Foundation of Korea (2017R1D1A1B03027954), and we would like to extend our appreciation to the relevant government departments.

REFERENCES

[1] POLYGON. (2017). Watch VR’s first Oscar-nominated short film. https://www.polygon.com/2017/1/24/14370892/virtual-reality-first-oscar-nominated-short-film-pearl

[2] CHINA GLOBAL TELEVISION NETWORK. (2017). Animated film “Arden’s Wake” wins best virtual reality award in Venice. https://news.cgtn.com/news/326b444d34557a6333566d54/share_p.html

[3] Slater, M., & Usoh, M. (1993). Presence in Immersive Virtual Environments. IEEE Virtual Reality Annual International Symposium, 90-96.

[4] Chertoff, Dustin B., Brian Goldiez, and Joseph J. LaViola. (2010). "Virtual Experience Test: A virtual environment evaluation questionnaire." In Virtual Reality Conference (VR), 2010 IEEE, 103-110.

[5] Chertoff, D. B., Schatz, S. L., McDaniel, R., & Bowers, C. A. (2008). Improving presence theory through experiential design.Presence: Teleoperators and Virtual Environments , 17 (4), 405-413.

[6] Davide, F., & Walker, R. (2003). Engineering Presence: an Experimental Strategy. In G. Riva, F. Davide, & W. A. IJSselsteijn, Being There: Concepts, effects and measurement of user presence in synthetic environments. Amsterdam, The Netherlands: Ios Press. 41-57.

[7] IJsselsteijn, W. A. (2003). Presence in the Past: what can we learn from Media History? In G. Riva, F. Davide, & W. A. Ijsselsteijn, Being There: Concepts, effects and measurement of user presence in synthetic environments. Amsterdam, The Netherlands: Ios Press. 17-40.

[8] Chen, X. (2006). Flow in Games. University of South California.

[9] Fu, F.-L., Su, R.-C., & Yu, S.-C. (2009). EGameFlow: A scale to measure learner's enjoyment of e-learning games. Computers and Education , 52, 101-112.

[10] Sweetser, P., & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. Computers in Entertainment, 3 (3).

[11] Sutherland IE. (1968). A head-mounted three dimensional display. InProceedings of the December 9-11, fall joint computer conference, part I 1968 Dec 9. 757-764. ACM.

[12] Kim, Y. K., Yoon, Y. S., Oh, T. G., Hwangbo, Y. H., and Hwang, J. H. (2017). Real-time VR strategy chess game using motion recognition. J. Digital Contents Soc. 18, 2.

[13] Ullmer, B., and Ishii, H. (2000). Emerging frameworks for tangible user interfaces. IBM Syst. J. 39, 915-931.

[14] Gunawardena, C. N., & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. American journal of distance education, 11(3), 8-26.

[15] TRIBECA FILM FESTIVAL. (2017). Draw me close. https://tribecafilm.com/filmguide/draw-me-close-2017

[16] TREE OFFICIAL. (2017). The film. https://www.treeofficial.com/

[17] BUSAN INTERNATIONAL FILM FESTIVAL. (2017). 22nd International Film Festival, 12-21 October, 2017. http://www.biff.kr/

[18] MIT Media Lab. (2017). Relational AI. http://www.media.mit.edu

[19] Jacoby, R., Ferneau, M., Humphries, J. (1994). Gestural Interaction in a Virtual Environment. Stereoscopic Displays and Virtual Reality Systems, SPIE 2177, 355-364

[20] Poupyrev, I., Billinghurst, M., Weghorst, S. and Ichikawa, T. (1996), November. The go-go interaction technique: non-linear mapping for direct manipulation in VR. In Proceedings of the 9th annual ACM symposium on User interface software and technology. 79-80. ACM.

[21] Schilit, B. and Theimer, M. (1994). Disseminating active map information to mobile hosts. IEEE Netw. 8, 22–32.

[22] Brown, M. G. (1996). Supporting user mobility. Ed. by J. L. Encarnação and J. M. Rabaey. Mobile Communications, 69–77. Springer, Boston.

[23] Cooperstock, J. R., Tanikoshi, K., Beirne, G., Narine, T., and Buxton, W. A. (1995). Evolution of a reactive environment. In Proceedings of the SIGCHI Conference on Human Factors In Computing Systems, 170–177.

[24] Elrod, S., Hall, G., Costanza, R., Dixon, M., and Des Rivieres, J. (1993). Responsive office environments. Commun. ACM, 36, 84–85.

[25] Hull, R., Neaves, P., and Bedford-Roberts, J. (1997). Towards situated computing. In First International Symposium on Wearable Computers, 146–153.

[26] Rekimoto, J., Ayatsuka, Y., Hayashi, K. (1998). Augment-able Reality: Situated Communication through Physical and Digital Spaces. 2nd International Symposium on Wearable Computers. 68-75

[27] Fickas, S., Korteum, G., Segall, Z. (1997). Software Organization for Dynamic and Adaptable Wearable Systems. 1st International Symposium on Wearable Computers, 56-63

[28] Steuer, J. 1992. Defining Virtual Reality: Dimensions Determining Telepresence. J. Commun. 42, 73-93.

[29] Lanier, J. 2001. Virtually there. Scientific American, 66-75.

[30] Bailenson, J., Patel, K., Nielsen, A., Bajscy, R., Jung, S.H., and Kurillo, G. 2008. The effect of interactivity on learning physical actions in virtual reality. Media Psych. 11, 354-376.

AUTHORS

Jiyoung Kang
Associate Professor at Dankook University
Ph.D, Korea Advanced Institute of Science and Technology
M.P.S, New York University

Jongkuk Lim
Ph.D Candidate
Department of Computer Engineering,
Dankook University