# A Novel Way of Estimating a User's Focus of Attention in a Virtual Environment

Xuanchao He and Zhejun Liu[✉]

College of Arts and Media, Tongji University, Shanghai 201804, China
winggeddreamer@tongji.edu.cn

Abstract. Results from prior experiments suggested that measuring immersion objectively (using eye trackers) can be a very important supplement to subjective tests (with questionnaires). But, traditional eye trackers are not usable together with VR HMDs (Head Mounted Displays) because they cannot "see" an audience's eyes occluded by helmets. The eye trackers compatible with HMDs are not easily accessible to students, researchers and developers in small studios because of the high prices. This paper explores a novel way of estimating a user's focus of attention in a virtual environment. An experiment measuring the relationship between subject's head movement and eyesight was conducted to investigate whether eye movement can be closely approximated by head rotation. The findings suggested that people's eyesight tended to remain in the central area of the HMD when playing a VR game and the HMD orientation data was very close to the eyesight direction. And therefore, this novel way that employs no other equipment than HMDs themselves can hopefully be used to estimate a user's focus of attention in a much more economic and convenient manner.

Keywords: Head movement · Evaluation · VR games · Focus of attention
HMD

1 Background of the Experiment

1.1 Introduction

Virtual reality (VR) is a computer-generated scenario that simulates a realistic experience. The immersive environment can be similar to the real world in order to create a lifelike experience grounded in reality or sci-fi [1]. When most people think of virtual reality, or VR, they probably think of hovering cars and time machines as well. While we are still away from traveling through time like Marty McFly, VR is knocking on our door. Technologies like the Samsung Gear VR and Google Cardboard are readily available and affordable for most consumers, as people are trying in more ways than ever to incorporate this innovative technology into their daily lives [2]. Immersive VR systems allow users to experience where they are, whom they are with, and what they are doing as if it was a real experience. In this context, the concept of presence refers to a phenomenon where users act and feel as if they are "really there" in a virtual world created by computer displays [3–5].

Many people think that the VR technology is an emerging technology. In fact, that was not the case. Since 1960s, research institutions have begun to simulate dynamic shapes and sounds using computers, which is followed by the budding of VR. In the next ten to twenty years, early concepts and theories of VR were formed [6]. However, the new wave of VR came in 2016. In the decades since Sega's initial experiments, VR technology has come a long way, and has made significant progress on its early shortcomings [7]. It all comes down to the development of VR headsets. The head-mounted display is the core component of the VR experience. The uniquely immersive experience of VR relies on the head-mounted display. A head-mounted display shows the computer graphics (CG) stereoscopically in front of users' eyes. It is more like a virtual camera through whose lens a user can see a 360-degree virtual environment. It provides wide angle displays covering the normal range of a human being's field of view: at present, Oculus Rift DK2 has field of view of 100° and Oculus Rift CV1 even has field of view of 110°, close to the that of human eyes which is about 120°. This feature makes people forget that they are immersed in the virtual world with head-mounted devices, which basically also allows you to interact with the virtual world with natural human vision. Today, consumers and developers can choose from three basic variations on the theme of stereo-3-D headsets for VR: systems with a dedicated internal screen; headgear that holds a smartphone as the source of the VR input; and augmented reality, in which the headset superimposes 3-D images and data on the user's view of the real world [7].

Although the second type, smartphone-based headset, is much cheaper, the user experience is not as good as the first type which is symbolized by Oculus Rift. In a headset like Oculus Rift, internal optics focus the user's vision onto a screen that is typically only a few centimeters away from his/her eyes. Sensors detect the position and orientation of the head, while the headset uses that information to calculate the images displayed on the screens. More importantly, although the processors, sensors and small high-resolution screens used in dedicated headsets are of the same types used in smartphones, in dedicated headsets they are all optimized for VR [7]. The dedicated headsets like Oculus Rift are connected to high-performance computers or game consoles that generate stereo pairs of images, generally at a rate of 90 times per second.

Meanwhile, because of the revival of VR, more and more game companies start to develop VR games for PC, PlayStation, Xbox One, smart phone or other platforms. There are 2651 VR games for HTC Vive and 1646 for Oculus Rift found on STEAM, one of the biggest online game store [8]. The number of PlayStation VR games is approximately 343 according to Google and Wikipedia [9]. In addition, the advent of VR-ready game engines has greatly simplified the production, not only the CG content creation but also the programming, of VR game, greatly reducing the development costs. As a result, most people would agree that there's never been a better time for VR games than now.

1.2   Related Works

Non-VR digital games have been enjoyed by millions of people around the world for a pretty long time. Modern games often have huge virtual environments for people to explore. Controls are more sophisticated, allowing people to carry out a wider variety

of maneuvers in a game. Through the use of the internet people can even play against opponents thousands of miles away. Charlene Jennett mentioned that the success of a computer game depends on many factors. Despite the differences in game design and appearance, successful computer games all have one important element in common: they have the ability to draw people in [10]. Concerning VR gaming, players are more than gamers, they are looking forward to the fulfilled potential of VR and the feeling of "losing" themselves in the VR game world. Such experience is referred to as "immersion", a term often used by gamers and reviewers. Immersion is often viewed as critical to game enjoyment, and is usually the outcome of a good gaming experience. When measuring and defining the experience of immersion in games, Charlene Jennett described it as "the psychology of sub-optimal experience, which clearly has links to the notion of flow (flow is described as the process of optimal experience, the state in which individuals are so involved in an activity that nothing else seems to matter [11]) and CA (Cognitive Absorption, as a state of deep involvement with software [12])". He also mentioned that immersion rather is the prosaic experience of engaging with a videogame [10].

The outcome of immersion may be divorced from the actual outcome of the game: people do not always play games because they want to get immersed, it is just something that happens. It does seem though from previous work that immersion is key to a good gaming experience [10]. The team explored immersion further by investigating whether immersion can be defined quantitatively. In one of the experiments, researchers investigated whether there were changes in participants' eye movements during an immersive task. They use an eye tracker to record participants' eye movements while they were engaged with a task/game. Overall the findings suggested that measuring immersion objectively (using eye trackers) can be a very important supplement to subjective tests (with questionnaires) [10].

1.3   Limitations of Eye Tracker in HMD

Based on these findings, we deduce that immersion of VR games can also be measured objectively (such as task completion time, eye movements). Using questionnaires only is not a precise and reliable enough way to measure immersion of VR games. But, the traditional eye trackers are not usable together with VR HMDs because they need to "see" an audience's eyes, but they are occluded by helmets (Fig. 1).

Eye tracking for HMDs is a natural next step and gained much attention in the research and development sector (e.g., FOVE Inc., Arrington Research, ASL Eye-Track, SR Research, or Sensor Motoric Instruments (SMI)). Even though first attempts started in the year 2000 [13], current inside-helmet eye tracker prototypes are still far from being consumer-ready: the SMI's eye tracker in the Oculus Rift is priced up to USD 15,000, and as for Tobii Pro, USD 28,800 approximately. Obviously, they are not easily accessible to students, researchers and developers in small studios. Moreover, with the ever-changing hardware technology, head-mounted devices may upgrade very quickly. Once installed into an HMD, the expensive eye trackers cannot be detached from it and reinstalled into a new one, which make it more unaffordable and uneconomic for those with a tight budget.

Fig. 1. The typical scenario of using a traditional eye tracker

2 Hypotheses

According to my personal experience, when playing a VR game with an HMD headset on, the eyeball rotation is limited. It is largely complemented by head rotation, allowing the player to look into different directions in the virtual environment. Then, several VR game players were closely observed, and it was found that their heads and bodies almost always moved and rotated when they want to look into different directions in the virtual environment. Such phenomenon can be seen more clearly in a demonstration video offered by SMI, a provider of inside-helmet eye tracking solutions.

In Fig. 2, there are two pictures are capture from a YouTube vide uploaded by SMI (https://www.youtube.com/watch?v=Qq09BTmjzRs). It's not difficult to conclude from the video that in most time the user's focus of attention, visualized as white circles, is near the center of the screen.

Because of this, a hypothesis was brought forward that people's eyesight tends to remain in the central area of HMD screens when they play a VR game. They prefer to move their heads to look at what they want to see, instead of moving eyeballs only. In

Fig. 2. Foveated rendering at 250 Hz from SMI

this way, eye movement can be closely approximated by the orientation of the head, which also means that the data from HMD rotation sensors can be recorded and analyzed to evaluate peoples' focus of attention. Since it can be further used as a low-cost data source for the evaluation of VR games, this approach, which employs no other equipment than HMDs themselves, is hopefully much more economic and convenient for most developers.

3   Method

Though where the subject's eyes were looking at could not be known while wearing a HMD headset, the purpose of the experiment was to evaluate the relationship between subject's head movement and his/her focus of attention by comparing the subject's eyesight direction with the orientation of the HMD. The orientation of the HMD was easy to measure, because it was directly sent by the device as rotation values to the game engine. However, the direction of eyesight was relatively difficult to measure because where the subject's eyes were looking at could not be known directly without an eye tracker. Therefore, a game was designed with the purpose to effectively guide players' vision.

The research team developed an original VR game called "Clock" with the Unity engine, and its functions were as follow. After the experiment began, an alarm clock appeared at a random location in the VR world and the player would immediately hear the ringing it. What the players needed to do was to find out where the clock was and to read out aloud the time shown on its surface. A very simple background was used to ensure no distraction so that the possibility of the player's being attracted to the background was greatly reduced. If the clock stayed in the view for more than 5 s, which meant that the player had found it, it would disappear. Then, the second clock would be generated in a new random position and the procedures described above repeated. There were 5 clocks in total, and the players were encouraged to find them and read out the time on the clock surface as fast as possible. The purpose was to make sure that the players looked at the clock carefully so that the position of the clock could be deemed as the focus of attention.

The angle between two eyesight rays, one going through the center of the HMD view, obtained from the HMD's orientation data and the other going through the center of the clock, was recorded from the beginning, when the first clock appeared, to the end of the test, when all 5 clocks were found, at a rate of 5 Hz. This angle data was saved into a text file. Big angle numbers meant that the alarm clock was far from the center of the screen while small numbers meant that the HMD was pointed accurately at the clock (Fig. 3).
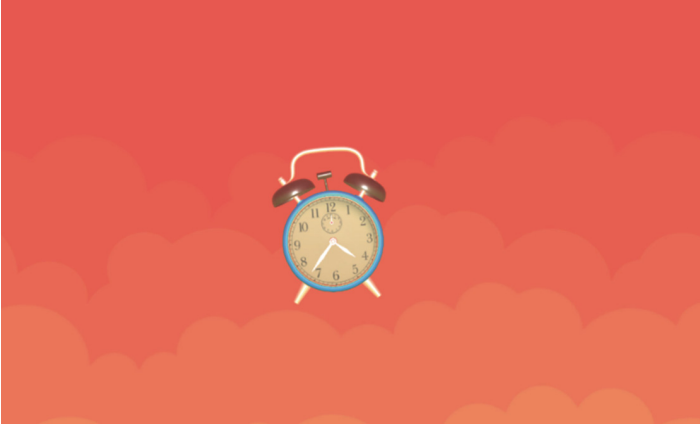
Fig. 3. Screenshot of the game "Clock"

3.1 Participants

20 participants took part in the experiment, all of whom were recruited from Tongji University. The average age was 22.85 (SD = 2.412), ranging from 20 to 26. Ten were male and ten were female. All the participants were willing and healthy. Since this game was intuitive and easy to play, prior experience of VR gaming was not concerned.

3.2 Equipment

Oculus Rift CV1 was used in the experiment as the HMD device. It was connected to a DELL T5810 workstation with an Intel Xeon E5-1660 CPU and an NVIDIA GeForce GTX 1080 graphic card. The high-performance computer was completely capable of running the "Clock" program at a framerate over 120 fps, guaranteeing that there was no bias caused by performance limitations.

The Oculus Rift CV1 had a built-in headset, so no external speakers were used. The "Audio Listener" in the Unity engine received inputs from every "Audio Source" in the scene and played sounds through the headset. For most applications it makes the most sense to attach the listener to the "Main Camera" [15], where the player was located. In this way, the engine could dynamically calculate the relationship between the audio sources and the player, at the same position of the "Audio Listener", and dynamically generated the sound effects correctly. As a result, the player could judge the location of the clock by listening to its ringing via the headset.

3.3 Procedure

Participants took part in the experiment one at a time, and the total duration of one session was about 5–10 min, depending on how much time a participant spent in finding clocks.

An experimenter firstly explained the rules of the game:

1. What a player needed to do was to find the clock and read out aloud the time shown on its surface;
2. Each clock would disappear after having been found for more than 5 s;
3. There were 5 clocks in total, and the player was encouraged to find all of them and read the time on the clock as quickly as possible.

Then, when the participant fully understood the rules, the experimenter helped the participant to put on the HMD device and started the game (Fig. 4).
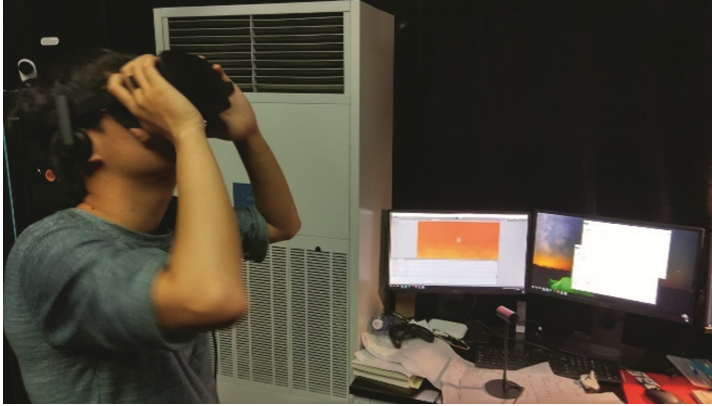

Fig. 4. Participant trying to find the clock

The game had no time limit: it would not end until the participant found all the 5 clocks.

Finally, when the game ended each participant was asked whether the clock remained at the central of the screen after he/she found the clock. The subject answer to this question served as an auxiliary confirmation to ensure that the participant was indeed looking at the clock.

3.4   Results

19 out of all the 20 participants replied "yes" to the question mentioned in the last paragraph of 3.3. There was only one exception: one of the player spent quite some time searching for the clock with no success and began to feel boring, so after he found it, he did not keep focusing on it for long enough. Considering the game design and the players' answers, the ray from the virtual camera to the clock could be thought of as the direction of the eyesight.

The participants' performances were shown in Figs. 5 and 6. The X axis stands for time and the Y axis stands for the angle between the eyesight and the central line of the HMD. The charts are not arranged according to the order of experiments because the subjects are independent to each other and the order was not relevant.
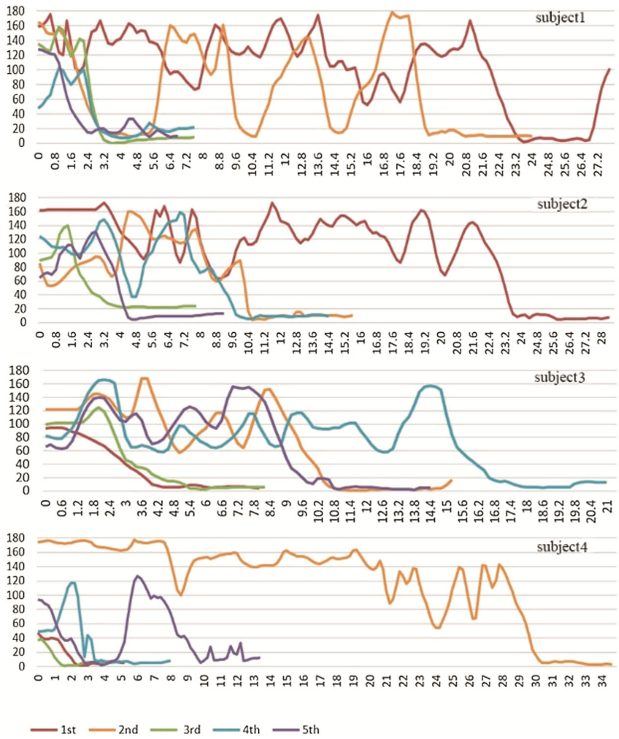
Fig. 5. Player's performance chart (part 1)

Every participant found the clock for five times, so there are 5 lines in each chart, each standing for one search-and-find processes. The left wavy parts of the curves indicate the period of time when the subject rotated his head searching for the clock.

One participant might spend different amount of time searching for each clock, so the lengths of the wavy parts are not exactly the same. For instance, subject 4 spent less than 3 s in finding the clock for the first, third, fourth time. But the second attempt took him as long as 30 s. No correlation was found between the total time used to find a clock and which attempt in order it was. The assumption that one subject could make progress through practice and thus shorten the searching time was not supported.

For the right parts of the curves, the Y value stabilized and approached to 0 for a period of time, which means that the subject found the clock and kept looking at it to read the time on its surface.

The parts whose Y value stay below 30° for more than 4 s are considered as "stable parts", while the rest are considered as the "wavy parts". As time went by the position of the clock on the screen became increasingly steady and close to the center.

Figure 7 shows the situation clearly. Since the alarm clock itself occupies a certain span of angles, we define it as "close to the center" when the angle difference between the eyesight and the HMD central line is below 20°, and "very close to the center" when below 10°. Among all the 100 polylines in the charts in Figs. 5 and 6, Y values of the
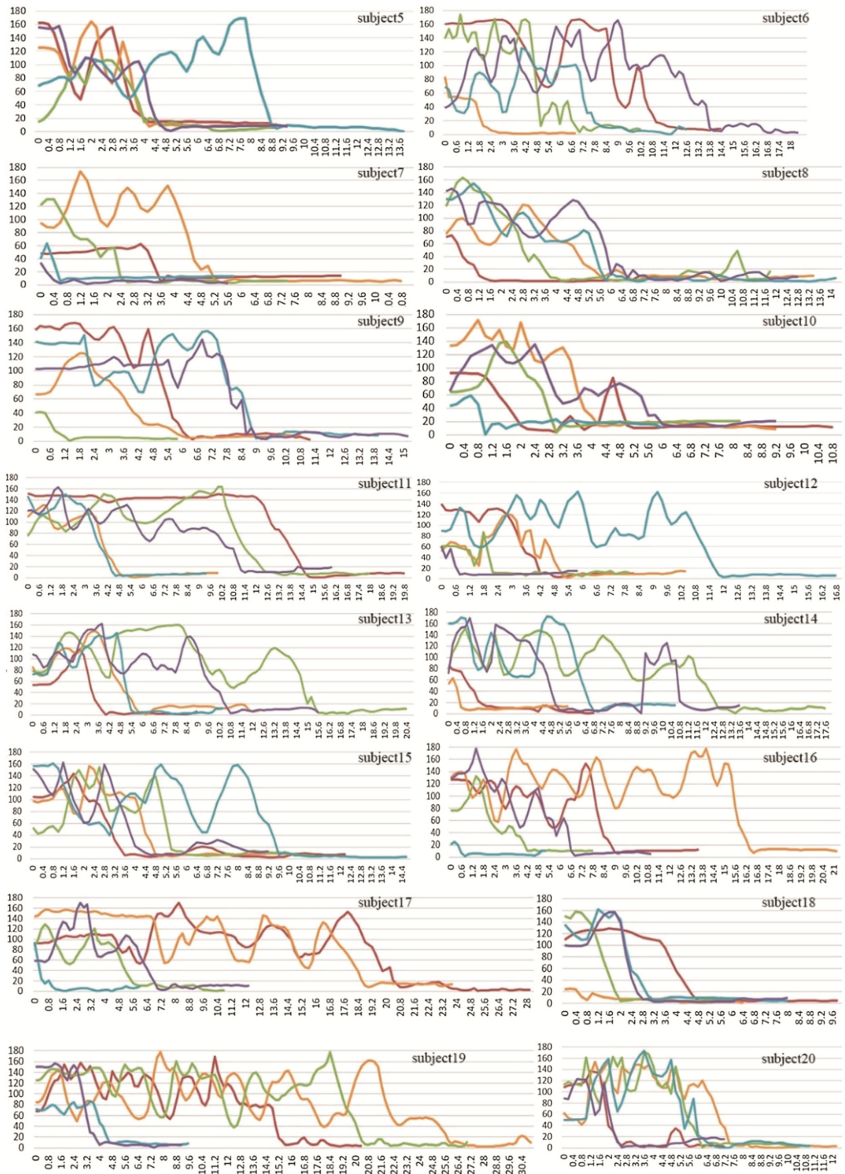
Fig. 6. Player's performance chart (part 2)

stable parts were "close to the center" (<20°) in 88 out of 100 cases (88%), and were "very close the center" (<10°) in 65 out of 100 cases (65%).
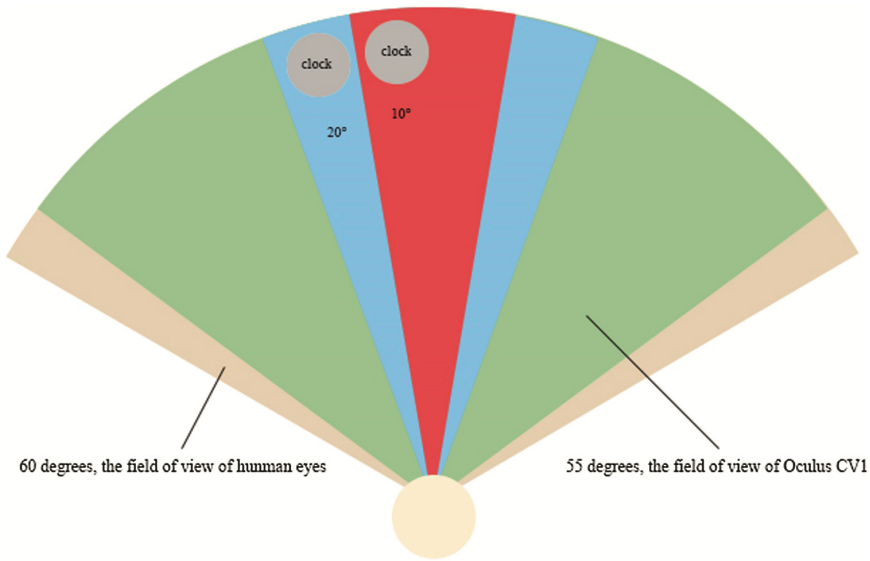
Fig. 7. Simulation of looking at the clock

4 Conclusion

From the experiment results, we found that people's eyesight tends to remain in the central area of the HMD screen when playing a VR game, and the hypothesis was partially supported. As a result, despite the deviations, the direction of the HMD can be used to roughly evaluate a player's focus of attention and to take over the role of an eye tracker to some extent.

Moreover, our work may contribute to the evaluation and optimization of VR games. What has been tested and verified in this paper means that the data directly from HMD rotation sensors can be used to evaluate a user's focus of attention, and maybe further used to evaluation the immersion of the game. This approach employs no equipment other than the HMDs themselves is much more economic and convenient when compared with other methods dependent on expensive devices like in-helmet eye trackers. In the future, the research team also plan to investigate the possibility of using this approach to evaluation the immersion level of VR games based on Charlene Jennett's prior research and findings.

Finally, this the limitations of this experiment are as follow, and future researchers on this topic may need to pay more attention.

- The sample size was not big enough. 20 subjects were not adequate to bring about a universal conclusion.
- Even though people at age 20–26 are the main target users of VR devices, this age range cannot stand for all users of VR games or programs. Future researchers are suggested to expand the age range.

- An in-helmet eye tracker, if affordable, may act as an objective standard reference. Instead of assuming the eyesight focuses at the center of the clock, as we did in this research, it may accurately tell where the focus is, and thus makes the result more accurate.

References

1. Wikipedia: Virtual reality. https://en.wikipedia.org/wiki/Virtual_reality
2. Danaj, M.: Virtual reality. Greater Lansing Bus. Mon. **30**(10), 14–15 (2017)
3. Slater, M.: Presence and the sixth sense. Presence-Teleoperators Virtual Environ. **11**(4), 435–439 (2002). https://doi.org/10.1162/105474602760204327
4. Sanchez-Vives, M.V., Slater, M.: From presence to consciousness through virtual reality. Nat. Rev. Neurosci. **6**, 332–339 (2005). https://doi.org/10.1038/nrn1651
5. Slater, M., Sanchez-Vives, M.V.: Enhancing our lives with immersive virtual reality. Front. Robot. AI **3**, 74 (2016). https://doi.org/10.3389/frobt.2016.00074
6. Chengwei, W.: VR technology and the harmonized HCI environment. J. Comput. Res. Dev. (1997)
7. Hecht, J.: Optical dreams, virtual reality. Opt. Photonics News **27**(6), 24–31 (2016). https://doi.org/10.1364/opn.27.6.000024
8. The official website of STEAM. http://store.steampowered.com/vr/
9. Wikipedia: List of PlayStation VR games. https://en.wikipedia.org/wiki/List_of_PlayStation_VR_games
10. Jennett, C., Cox, A.C., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., Walton, A.: Measuring and defining the experience of immersion in games. Int. J. Hum.-Comput. Stud. **66**(9), 641–661 (2008). https://doi.org/10.1016/j.ijhcs.2008.04.004
11. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper and Row, New York (1990)
12. Agarwal, R., Karahana, E.: Time flies when you're having fun: cognitive absorption and beliefs about information technology usage. MIS Q. **24**(4), 665–694 (2000)
13. Duchowski, A.T., Shivashankaraiah, V., Rawls, T., Gramopadhye, A.K., Melloy, B.J., Kanki, B.: Binocular eye tracking in virtual reality for inspection training. Int. J. Ind. Eng.: Theor. Appl. Pract. **9**(2), 89–96 (2000). https://doi.org/10.1145/355017.355031
14. Stengel, M., Grogorick, S.: An affordable solution for binocular eye tracking and calibration in head-mounted displays. In: Proceeding: MM 2015 Proceedings of the 23rd ACM International Conference on Multimedia, pp. 15–24 (2015). https://doi.org/10.1145/2733373.2806265
15. Unity - Manual: Audio listener. https://docs.unity3d.com/Manual/class-AudioListener.html