Under the Movement of Head: Evaluating Visual Attention in Immersive Virtual Reality Environment

Honglei Han
School of Animation and Digital Arts
Communication University of China
Beijing, China
Department of Computer Science
UNC Charlotte
Charlotte, NC 28223, USA
hanhonglei@cuc.edu.cn

Aidong Lu Unique Wells
Department of Computer Science
University of North Carolina at Charlotte
Charlotte, NC 28223, USA
aidong.lu, uwells@uncc.edu

***Abstract*—A method to measure what and how deep the user can perceive in immersive virtual reality environments is proposed. A preliminary user study was carried out to verify that user gaze behaviors have some specific differences in immersive virtual reality environments compared with that in traditional non-immersive virtual reality environments base on 2D monitors and interactive hardware. Analyzed from the user study result, the user gaze behavior in immersive virtual reality environments is more likely to move their head to let interested object locates in the center of the view, while in non-immersive virtual reality environments the user tends to move their own eyes and only move the avatars head when necessary. Base on this finding, a quantitative equation is proposed to measure the users attention in immersive virtual reality environments. It can be used into a quality evaluate system to help designers find out design issues in the scene that reduce the effectiveness of the narrative.**

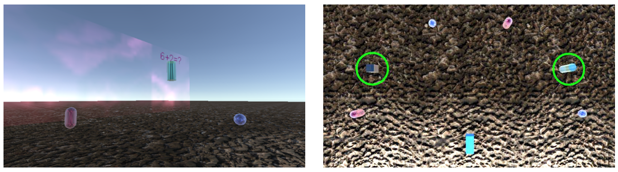
***Keywords*-view guide; visual attention; virtual reality**

I. Introduction

With the evolution of modern three dimensional rendering engines and Head Mounted Displays (HMDs), users can vividly experience immersion, which has been pursued in the computer graphics field for decades.

In immersive VR, users can explore using natural interaction methods, which are much easier to utilize excessively. Therefore, it is vital for VR film directors to be able to evaluate the user's attention. This assists the directors with understanding their users experiences, and then helps give ideas on how to improve them, to make sure the user focuses mostly on significant objects not straying too much to observe background objects.

Very recently, Vincent *et al.* addressed the problem of how users explore virtual environments and what constitutes saliency in immersive applications[1]. The same result with ours is derived in this paper: head movement can be a valuable tool to analyze the approximate regions that users attend to in a scene without the need for additional eye-tracking hardware. Contrasting with them, the interactive immersive VR scenes, which are more popular in modern VR applications, are the main study objects in our paper instead of still panoramas used in [1]. A virtual object, which



(a) From user's view (b) Top view

Figure 1. Shots captured during the user study scene (a), and its top view (b). Green circled objects are task-related

is the basic unit in the VR designing process, based visual attention is obtained in our paper rather than conventional region based saliency map obtained in [1].

In the immersive video field, Thomas *et al.* presented a visualization framework for analyzing head movements and gaze data for immersive 360° videos in the paper [2]. It can further be used as an evaluation tool to detect whether the attention guidance of an immersive video works as expected. However, viewing directions used in that paper cannot efficiently express what particular semantic virtual objects are focused on by the participants.

We propose a hypothesis about users having similar patterns in exploration with immersive VR environments. A preliminary user study experiment has been established to verify it. We sought to find any particular user gaze patterns in immersive VR, and used non-immersive VR environment patterns as a comparison scale.

II. User study

There are five user study scenes developed with Unity 3D game engine. Each of them contains one or two objects with three addition problems, appearing sequentially, between two numbers less than ten for participants to solve (as seen from Figure 1(a)). These objects with math problems are task-related, and the others as background objects, since it has been proven that task-related gaze behavior can dominate over saliency [3]. The scenes all have the same scene arrangement, as shown in Figure 1(b). Objects circled in green may have math problems on top of them, and a camera is located at the bottom middle (see the blue box as the representation in Figure 1(b)). Scene $1static$ contains only one task-related object, while scene $2static$ contains two. In order to find out if participants' view direction is strongly correlated to task-related objects, we set the objects to be active in the last three scenes. In $1move$ scene, there is only one move task-related object moving horizontally, back and forth. The $2moveS$ scene contains two in up-and-down and left-and-right trajectories, respectively. The final scene $2move$ contains two moving task-related objects in triangular and rectangular trajectories, which are more complex compared to previous scenes.

We invited twenty-five participants, who are all students and faculty members in our university, aging from 19 to 48 with fifteen males and ten females. After fully understanding the purpose and control method of the user study application, they were asked to play the scenes in both immersive and non-immersive modes. The gaze behavior we are mostly concerned with is the included angle between view direction and the vector from the camera to the object. It indicates how centered the object is in the camera. We believe the included angle is a strong hint of the objects' visual attention degree from the user.

It can be easily seen from the captured included angle data, in immersive mode, participants tend to focus on the task-related objects exclusively when they were trying to solve the math problems upon the objects. In non-immersive mode, however, they tend to only keep the task-related objects inside the screen and use their own eyes to trace the moving objects instead of trying to rotate the virtual camera to trace them. A probable reason is that in immersive environments people act like they are in real world, in which they tend to keep the goal right in the center of their vision.

III. Object-based attention quantitative equation

We can conclude from previous section that objects receiving the most user attention continuously appear in the center of the screen in immersive VR mode, as the included angles of task-related objects can always be very small.

We found out that the Logistic function fairly meets our requirements to serve as a math model to reflect the gaze pattern in immersive VR scenes, as it emphasizes the center effect. Hence, the object-based visual attention value A can be calculated with Equation 1, a revised Logistic function,

$$A(O) = \sum \frac{1}{1 + e^{k(deg-d)}} dt, \qquad (1)$$

where k denotes the steepness of the curve; here we use 0.5. deg denotes the included angle degree between view direction and the vector pointing from the camera to the object O. d denotes the sigmoid curve's midpoint, which means that for the threshold of gazing included angles, where we use 15° as the default value. Included angles below 10° can be seen as focusing in immersive VR scenes. We add 5° as a tolerance. The attention value is accumulated along with time t to prove the statement that the more that subjects have interest in an object, the longer they will focus on it[4].

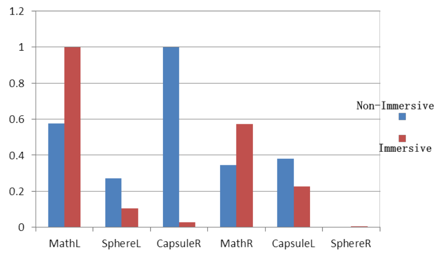


Figure 2. A user's normalized attention degree calculated by Equation1 in scene $2moveS$ using non-immersive and immersive modes respectively

From Figure 2, the apparent correlation between saliency levels of objects and their attention degrees calculated by Equation 1 are easily seen. In immersive VR mode, the task-related objects received the highest attention values. This means that the participant always put them in the center of his or her view. While in non-immersive VR mode, the participant kept the view unmoving so much that background object $CapsuleR$ was put in the center of his or her view unintentionally, resulting in its highest attention value.

IV. Conclusion

A quantitative object-based attention evaluation method was proposed in this paper. Based on the results of the user study to diagnose the gaze pattern for task-related objects in VR scenes, we found that head movement is vital, especially in immersive VR environments. In accordance with this finding, we borrowed a Logistic function to evaluate the visual attention degree, with included angles between view directions and the vectors from the camera to the object. Our method can be easily embedded into current immersive VR applications without adding complex calculations or equipment costs.

References

[1] V. Sitzmann, A. Serrano, A. Pavel, M. Agrawala, D. Gutierrez, and G. Wetzstein, "Saliency in vr: How do people explore virtual environments?" *arXiv preprint arXiv:1612.04335*, 2016.

[2] T. Löwe, M. Stengel, E.-C. Förster, S. Grogorick, and M. Magnor, "Visualization and analysis of head movement and gaze data for immersive video in head-mounted displays," in *Proceedings of the Workshop on Eye Tracking and Visualization (ETVIS)*, vol. 1, 2015.

[3] M. Land, N. Mennie, and J. Rusted, "The roles of vision and eye movements in the control of activities of daily living," *Perception*, vol. 28, no. 11, pp. 1311–1328, 1999.

[4] F. Alt, A. S. Shirazi, A. Schmidt, and J. Mennenöh, "Increasing the user's attention on the web: using implicit interaction based on gaze behavior to tailor content," in *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*. ACM, 2012, pp. 544–553.