Understanding User Attention In VR Using Gaze Controlled Games

Murtada Dohan
University of Northampton
Northampton, United Kingdom
murtada.dohan@northampton.ac.uk

Mu Mu
University of Northampton
Northampton, United Kingdom
mumu@ieee.org

ABSTRACT

Understanding user's intent has a pivotal role in developing immersive and personalised media applications. This paper introduces our recent research and user experiments towards interpreting user attention in virtual reality (VR). We designed a gaze-controlled Unity VR game for this study and implemented additional libraries to bridge raw eye-tracking data with game elements and mechanics. The experimental data show distinctive patterns of fixation spans which are paired with user interviews to help us explore characteristics of user attention.

CCS CONCEPTS

• **Human-centered computing** → **HCI theory, concepts and models**; **Empirical studies in HCI**.

KEYWORDS

user attention, virtual reality, eye tracking, gaze, serious games

ACM Reference Format:
Murtada Dohan and Mu Mu. 2019. Understanding User Attention In VR Using Gaze Controlled Games. In *ACM International Conference on Interactive Experiences for TV and Online Video (TVX '19), June 5–7, 2019, Salford (Manchester), United Kingdom.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3317697.3325118

TVX '19, June 5–7, 2019, Salford (Manchester), United Kingdom

ACM ISBN 978-1-4503-6017-3/19/06.
https://doi.org/10.1145/3317697.3325118

INTRODUCTION

The integration of user intent recognition is a crucial factor in enabling new immersive and personalised media experiences. For instance, digital artwork can react to user's attention and viewing habit, and elicit a unique experience for each viewer. Capturing user attention in VR is a non-trivial task. Eyes and head positions need to be accurately mapped into three-dimensional coordinates and associated with virtual objects in a dynamic virtual environment.

In order to facilitate the research, we conduct a user experiment using an specifically designed gaze-controlled VR game with built-in data gathering and analysis functions. Additional software libraries are implemented to bridge time-coded eye-tracking data with game elements and mechanics. The eye-tracking function of the VR headset is used as a gameplay mechanism as well as for eye fixation measurement. Test participants play through the game independently and the experimental data reveals distinctive patterns of how game players divide their attention. We also observe clusters of fixation spans that reflect different degrees and natures of user attention. Post-game interview provide us with insights in interpreting the eye fixation data.

In summary, the main contributions of the paper are:

- Design a gaze-controlled VR game as an eye-tracking research tool.
- Implement libraries (to be released as open-source) to facilitate the mapping of virtual objects with eye-tracking data.
- Design and conduct user experiments. Analyse fixation data and user feedback to understand different degrees and natures of user attention.

RELATED WORK

Eye tracking is one of the most expected features of VR applications. It provides the capabilities to precisely track users' gaze in a virtual environment while their eyes are usually not visible to onlookers. Recent developments have seen constant progression in improving latency, precision, cost and robustness [2]. Bekele et al. developed a system that monitors eye gaze in VR and physiological signals related to emotion identification to explore new efficient therapeutic paradigms [1]. Research on attention in VR is also valuable to help identify and mitigate the effect of mental disorders [5]. We are also witnessing an increasing support from headset manufactures. Besides FOVE 0 and many aftermarket VR eye-tracking solutions, HTC recently announced VIVE Pro Eye which will further strengthen the technological support for VR eye-tracking research and application development.

EXPERIMENTATION ENVIRONMENT

Eye-tracking Integration in Unity

FOVE 0 VR headset has built-in eye-tracking capabilities using infra-red cameras and a position tracking camera. We update the standard FOVE Unity plugins and consider Unity Colliders as the main object for attention detection and user's three-dimensional fixation point is snapped to the corresponding game object in order to identify which object is being looked at. Our enhanced eye detection also combines independent measurements from each eye in order to accommodate extreme viewing angles and eye conditions (e.g., strabismus). The Unity libraries implemented for our experiment are shared on GitHub[1].

[1] https://git.io/fjqO1

Game Environment and Mechanics

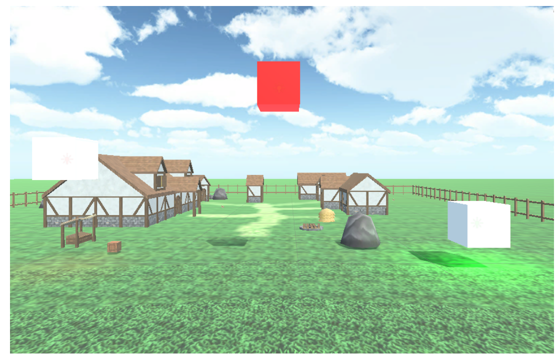


Figure 1: Game environment

The VR game environment consists of four components: static scene, camera, targets, and non-player character (NPC) (Figure 1). The scene mimics a farm with multiple farm houses and cottages within a partially open wooden fence enclosure. The player's view is set so that all points of interests in the game are in front of the players within their 180°view (i.e., hemisphere). The targets of the game are three white cubes that change their colour to red, green and blue respectively when being looked at by a player. The cubes float in the air and move randomly in different speed and directions. NPCs include animated flying birds and farm animals. The purpose of NPC is to study how players attention can be affected by non-target objects.

The storyline is around shooting at three aliens attackers with eye-controlled missiles. We simplified the targets' visual representation to minimise the impact from unrelated factors. To play the game, players keep their eyes on any of the three cubes for as long as they could in order to gain credits. The overall score is calculated as the sum of credits on all three cubes.

User Attention

We capture user attention as fixation spans on each target. A fixation span is a number of continuous frames that a player looks at an object. Unlike a reading task where users have sudden changes of fixation points (saccade [3]), our players follow the cubes through smooth pursuit using both eyes and head movements.

The eye detection is conducted 30 times (frames) per second, therefore the granularity of the measured fixation spans is around 33 milliseconds. Short eye blinks are eliminated so that a long gaze with blinks is recognised as one continuous fixation span. We also record the colour and distance of the cube for each fixation span captured.

USER EXPERIMENT

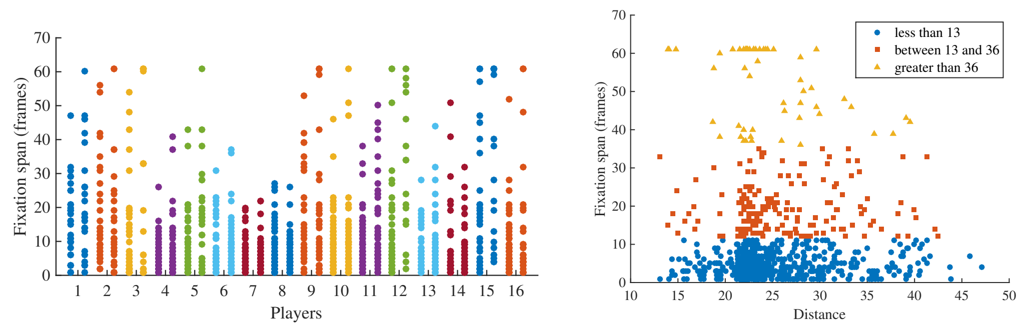
Test procedure

The experiment took place on multiple study areas on a University campus. Overall, 20 participants (mostly University students) took part in the test. 5 participants are female and the age of our participants ranges between 16 and 40.

The test started with a short instruction which 1) helps participants to get familiar with the games environment, and 2) calibrates eye tracking tools for each player. Three participants did not pass the calibration phase due to their eyes not correctly captured by the eye-tracking cameras. One participant felt dizzy and withdrawn from the test. Each of the remaining sixteen participants played through the game twice (two attempts). After each test a short interview was conducted to collect user feedback.

Results

Figure 2(a) shows the in-game fixation spans on all three cubes from each player. The figure includes both attempts in the order that they were taken. In general, players achieved longer attention spans in the second attempts due to their increased familiarity and confidence with the gaze-controlled game play mechanism. The results also manifest themselves in clusters, especially in the second attempts. We believe that the clusters reflect different degrees of user attention in a VR environment. We use *Ckmeans* [6] (a variation of *k*-means clustering) to cluster numeric fixation spans in groups. *Ckmeans* is developed as a dynamic programming approach to group data with the least within-group sum-of-squared-deviations (optimally homogenous within groups). With a three group configuration, our experimental data is split with two cut points: 12 and 36 frames. We believe that the groups are associated with very short fixations when scanning objects; medium fixations to observe specific items; and longer gazes where intense focus is required for a task.



(a) Fixation span of each player

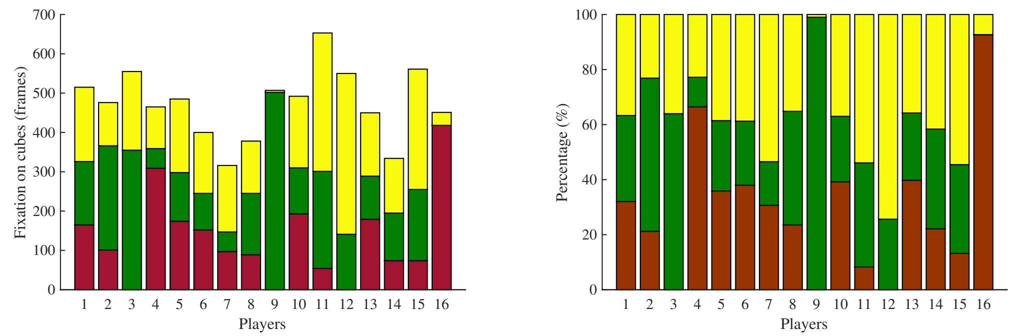
(b) Fixation span vs cube distance

Figure 2: Fixation span

We hypothesised a negative correlation between the length of fixation span and the distance of the corresponding cube from the player. The experimental data does not exhibit any significant correlation on any of the three attention clusters as Figure 2(b) shows. However, as the distance to target increases, players struggled to achieve long gaze. This can be attributed to the cubes becoming smaller as they move away or the interruption from the players' head movements while the cubes move out of the visual field.

The players' choices of game strategy also led to unique observations. Many players (e.g., Player 1, 2, and 5) chose to switch their attention to the nearest cube or the slowest moving cube (Figure 3). Some other players (e.g., Player 9 and 16) decided to consistently follow one cube no matter how far it is or how fast it travels. Favouring one cube did help some players to gain relatively high score. However, this did not lead to significantly higher fixation spans (Figure 2(a)).

All sixteen players responded positively in the post-experiment interview about their game experience. They were generally intrigued by the gaze-control and "can't wait to play the next version of the game". Some players commented that gazing at a cube in bright colour "becomes a bit uncomfortable after a few seconds" and they couldn't help "looking away from the cube briefly". Such brief discontinuities in fixation can be several frames in duration and therefore exceed the limits set by our eye blink filter. This could be one of the reasons that we did not observe high fixation spans on players who clearly had continuous attention on one cube through the game. Based on this observation we



(a) (b)

Figure 3: Players' game strategies

believe that the measurement and the use of attention span should be more tolerant to accommodate brief fixation discontinuities in an VR environment.

When asked about the game strategy, one player responded "I tried to look at the coloured boxes as much as possible while keeping the other boxes in peripheral vision...". Other players also mentioned that they noticed “a bird passing by” while looking at the cubes. The comments have triggered our interest in further study on the use of peripheral vision in VR which is believed to be twice as sensitive as central vision to a moving object.

CONCLUSIONS AND FUTURE WORK

Our VR game and experimental results provide some interesting insights on this research topic. We plan to integrate more comprehensive game elements to further study user attention in different scenarios such as virtual classrooms and mechanical workshops. Working with colleagues from our psychology department, we will look into using eye-tracking and user behaviour data to infer user intent and personality. We are also working with an expert in fine art to experiment with responsive VR paintings that can react to how it is viewed as a pathway to better connect artists with audience.

ACKNOWLEDGEMENT

This work is supported by UK Research and Innovation (UKRI) under EPSRC Grant EP/P033202/1 (SDCN).

REFERENCES

[1] E. Bekele, Z. Zheng, A. Swanson, J. Crittendon, Z. Warren, and N. Sarkar. 2013. Understanding How Adolescents with Autism Respond to Facial Expressions in Virtual Reality Environments. *IEEE Transactions on Visualization and Computer Graphics* 19, 4 (April 2013), 711–720. https://doi.org/10.1109/TVCG.2013.42

[2] David D. Bohn. 2015. Integrated eye tracking and display system. *US8998414B2* (2015).

[3] David Fitzpatrick Lawrence C Katz Anthony-Samuel LaMantia James O McNamara Dale Purves, George J Augustine and S Mark Williams. (Eds.). 2001. *Neuroscience*. Sinauer Associates Inc. https://www.amazon.com/Neuroscience-Book-CD-ROM-Dale-Purves/dp/0878937420?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=0878937420

[4] FOVE. [n. d.]. FOVE 0 EYE TRACKING VR DEVKIT. *https://www.getfove.com/* ([n. d.]).

[5] C. Mei, B. T. Zahed, L. Mason, and J. Ouarles. 2018. Towards Joint Attention Training for Children with ASD - a VR Game Approach and Eye Gaze Exploration. In *Proc. IEEE Conf. Virtual Reality and 3D User Interfaces (VR)*. 289–296. https://doi.org/10.1109/VR.2018.8446242

[6] Haizhou Wang and Mingzhou Song. 2011. Ckmeans. 1d. dp: optimal k-means clustering in one dimension by dynamic programming. *The R journal* 3, 2 (2011), 29.