Dynamic Focus Selection for First-Person Navigation with Head Mounted Displays

Thiago M. Porcino* Cristina N. Vasconcelos† Daniela Trevisan‡ Esteban Clua§

Fluminense Federal University, Institute of Computing, Brazil

Figure 1: Our plugin running in an virtual environment

ABSTRACT

The increase of Head Mounted Display systems in games is a natural tendency for the next generation of digital entertainment. While the immersion produced by those devices is high, discomfort is still a problem pointed by many users. In such context, this work proposes to reduce the discomfort caused by the lack of focus in these display interfaces. More specifically, we propose a dynamic selection of focus elements aiming to reduce visual discomfort in first-person navigation through immersive virtual environments (Figure 1). Behind the proposed effect, a heuristic model of visual attention is proposed sustaining a real-time selection of the on focus target. While most approaches use a simple point and focus selection, our solution creates a novel and more precise focus selection, considering different scene elements and attributes. A case of study of the model was made developing a component for a commercial game engine. It was tested by a set of users that explore a virtual scene using the Oculus Rift with and without the proposed visual effect. Positive qualitative results based on questionnaires confirmed discomfort reduction when using our solution.

Keywords: Head Mounted Display Systems, depth of field, first-person navigation, virtual reality, focus selection

1 INTRODUCTION

Head Mounted Displays (HMDs) are proving to be an important tool for an increase in 3D games immersion. Currently, a large and rapid growth in the emergence of several solutions can be observed and is being pointed as one of the most important game industry trends for the next years.

Latency was a serious problem in previous generation of HMDs, making it difficult to produce games for the mass market [1]. Modern HMDs assemblers state that they bypassed this issue, pointing that VR could now become the next generation gaming platform.

*e-mail: thiagomp@ic.uff.br
†e-mail: crisnv@ic.uff.br
‡e-mail: daniela@ic.uff.br
§e-mail: esteban@ic.uff.br

However, many users have reported discomfort due to the prolonged use of VR devices with other different causes, rather than latency [2]. According to [3] [11] [12] , the causes for visual discomfort regarding stereo vision devices can be listed as:

- Eyewear with image separation between eyes;
- Incorrect calibration or poor focus simulation;
- Convergence accommodation conflict.

This paper addresses the discomfort caused by the lack of adequate simulation of focus in stereoscopic Head Mounted Displays with head movement tracking. In humans, the focus produces blurriness effects according to the depth of field (DoF) and the range of distances of the objects in the visual area [13]. Due to the ocular convergence, objects outside this range, located behind or in front of the eyes, are perceived as blurred and unfocused. Human brains are used to interpret images with this natural effect. While 3D movies or cinematic productions solve this effect with parallax camera convergence, 3D games rarely apply this correction, due the real time interaction of the player and the lack of information on where exactly to apply the focus.

Different DoF simulation techniques were proposed in computer graphics, in order to generate realistic scenes [7]. Virtual reality and general entertainment applications often use such techniques to grab the viewers attention and enhance immersion [8].

Most techniques usually choose the focus target according with a game logic parameter or with a simple ray cast applied to the center of the image. This strategy may drastically reduce the discomfort on such environments, but they are often imprecise for choosing the exact point of focus desired by the user, especially when this point is a small object.

In this work we propose a novel and more robust heuristic for real time focus selection in HMD environments. Our strategy is based on different sets of parameters and possible relationships of the elements with the user.

Discomfort measurements (with and without visual effects) were evaluated in a virtual environment simulating tasks related to the observers different visual targets. In such environments, user immersion involved an Oculus Rift capturing stereo images. This study adopted pre-established criteria validated by various simulator sickness measuring experiments [6]. Minor changes to the original assessment procedure were necessary to contextualize our experiment, described in the results section. By comparing navigation tests with and without the use of a dynamic focus effect, this study demonstrates that HMD devices reduce discomfort when supported by our strategy.

2 Related Work

Kass et al. in [5] derived an algorithm for computing the DoF through an interactive graphics processing unit (GPU) diffusion simulation. According to them, filters using recursion on GPUs are problematic. They introduced a new DoF post-production model that uses heat diffusion formulation for precise real-time effects. Selgrad et al. in [10] proposed a DoF simulation algorithm that uses a composition of multilayered images. Their algorithm renders scenes using a stationary camera that calculates image layers and manages composition fragments using pixel lists; however, this can generate blurs in different pictures. Their data structure can generate real-time DoF but has greater computational expenses compared with other techniques (single layer, single-G-buffer). However, despite its computational load, this technique is able to produce all the information needed to generate the DoF effect, including transparent objects. While both contributions are not related to automatic focus point selection, they are efficient approaches for the real time DoF effect production, after choosing the desired point of interest.

In [4], Hillaire et al. developed a blur model in virtual world navigation for first-person cameras. Their study defines a static interest region (called auto focus zone), which is identified by a central rectangle in the image. While this work tries to solve a problem that is similar to that we want to solve with our approach, it is a naive and simple focus region selection, based on the direction of the virtual camera.

Recent related studies have pointed out the adverse and unnatural effects of HMDs viewing conditions. To enhance the viewing of natural images with HMDs, Liu and Hua [9] developed a hardware that produces a focus system using microlens cameras (fast liquid lens). In their study, focal planes can range from infinity to a distance of eight diopters. To test the concept, a comparison of effects was performed with two graphic rendering engines. The first method used a high-rate update frequency (f = 37.5 Hz), which produced unfocused images. The second method used low-frequency refresh rate (f = 18.75 Hz), which resulted in better focused images.

According to [9], the systems perception of stereoscopic depth is superior to that of conventional stereoscopic monitors with a single focal plane. Although their approach differs from our proposal, it proves the importance of creating dynamic and distinct focal planes.

Carnegie and Rhee‘s research is more closely related to our work [2]. They proposed the use of DoF simulation to decrease discomfort caused by HMD devices. Instead of using eye tracking systems, typically adopted to precisely calculate focus areas, they developed a dynamic real-time DoF using a GPU to maintain the screen center in focus. They used HMD devices that react to users neck movements. While focus is kept centered, a shift of 500ms focus delay is created to mitigate user discomfort caused by sudden focus changes. According to Carnegie and Rhee [2], the time to reorient focus depends on the users age and the lighting conditions at the scene. They assumed that for a real-time performance, users require 500 ms to refocus from an infinite distance to approximately 1 m. For the present study, they evaluated 20 participants with a simulator sickness questionnaire. For each of its 18 questions, participants verbally responded to symptoms using the Likert 5-point scale (ranging from none without symptoms to severe with traumatic symptoms). Accordingly, 30% first-time technology users were disturbed such that they were unable to use it for more than 30 min. Our results demonstrated that DoF rendering techniques can significantly reduce the discomfort caused by HMD devices. Although we initially used the adaption time suggested by their work, in our tests we also changed the focus adaption time based on empirical observations.

While Kass et al. in [5] and Selgrad et al. in [10] ignore users focus and aim at improving real-time blurring effect, whereas Hillaire et al. in [4] and Carnegie et al. in [2] always fix regions of interest (ROI) by uniting the users focus to the screen center. As an additional contribution to the aforementioned studies, we developed a model for selecting dynamic ROI by simulating a self-extracting mechanism of visual focus that isolates ROI in the visual field. This ROI is then used in real-time DoF effects calculation to decrease HMDs discomfort. In this study, ROI are dynamic and move in the 3D scene.

3 Our Model

Our study adopts a representative model of camera systems for HMD devices in which a pair of stereo cameras are positioned parallel to each other. This is a simplified model of epipolar geometry. We assume that two cameras (**CL** and **CR**) with optical centers (**OL** and **OR**) are positioned left and right to align both camera image planes (epipolar lines also coincide) and produce IL and IR images. The figure 2 illustrates our model, where a virtual camera (**CM**) is placed at the midpoint (**M**) of the line connecting the two projection centers (**OL** and **OR**). The normal vectors are parallel to both projection planes. **M** is associated with the normalized vector (**N**) defined by a direction such that **CM** aligns with the parallel geometry (Figure 2).

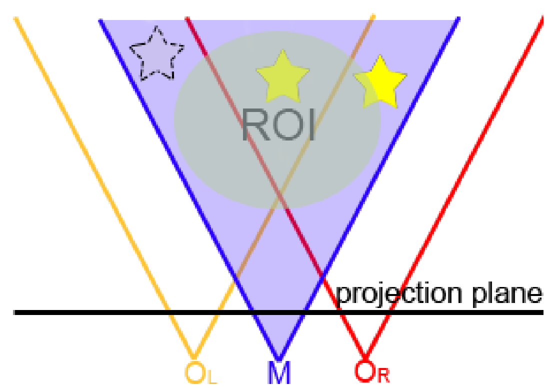

Figure 2: **OL**, **OR**, and their corresponding middle point **M**. ROI associated with the visual field pre-selects objects considered as candidates of visual attention (filled objects)

The heuristic behind the proposed visual focus is based on the geometry described to select a focus object. A ROI is initially defined for a 3D scene. This ROI facilitates the implementation of real-time heuristics in complex scenes by dynamic rendering by exclusion of objects located outside the observers visual field. The ROI of the visual field is thus defined along the optical axis **M-N** of camera **CM**.

After this first step, all objects still partially within the ROI become possible targets of the viewers visual focus and thus are to be viewed in focus. Objects are analyzed and selected by an importance metric function **I(o, C)**, where ”o” represents a given object within the ROI and **C** is a set of cameras. We introduce the concept of adopting several heuristics that may be derived from the proposed model to test usability issues or a specific interest of a given application type. This heuristics will search a best candidate inside the ROI to become the center of attention and temporally be the focus object. For this work, we propose a model inducing first-person virtual environment navigation.

3.1 Heuristics For First-Person Virtual Environment Navigation

For first-person exploration of virtual environments using HMDs, we propose the importance function (I):

$$I(o,C_1,C_2) = (PRM * RM(o,C_M)) + (PD * D(o,C_M)) + (PV * V(o)) \quad (1)$$

where **PRM + PD + PV** = 1 and represent pondering factors between metrics **RM**, **D**, and **V**.

The first visual focus evaluation parameter assigns greater importance to objects closer to the cameras centers. It is obtained by a series of rays within a cone centered at the midpoint between **C1** and **C2** (Figure 3). The cone interior is divided into concentric layers (**k**). For each layer, n rays are generated, called metric rays (**RM**). From their common origin (**M**), their uniform dispersion is then defined by the golden rule, resulting in k n rays (Figure 4) .

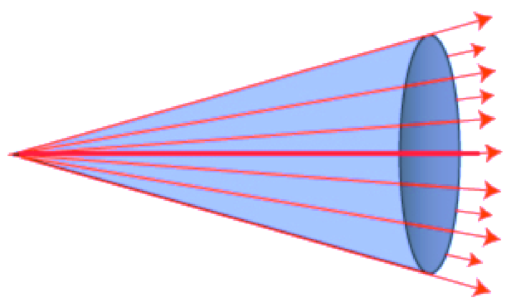

Figure 3: Metric Rays.

The number of metric rays hitting a scene object is evaluated by the proposed heuristic model. Each ray has an alpha weight depending on how close each layer is relative to the cones center, accounting for the importance of objects centered in the scene in the heuristic calculation.

$$RM(CM,o) = \sum_{i=1}^{k} \sum_{j=1}^{n} c(j,o) * \alpha(i) \quad (2)$$

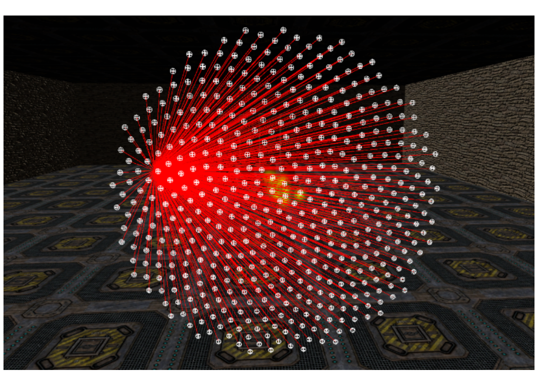

Figure 4: Points distribution by golden rule

The binary function c determines whether rays j collide with objects o. If a given radius collides with more than one scene object, only the one closest to the camera is considered. The second proposed parameter **D(CM,o)** uses depth relative to point M of a specific object o belonging to a ROI. This metric evaluates the proximity between the object and a viewer, assuming that closest elements tend to receive more attention, so become focused.

The third parameter is the added value of the object (**V**) and is incorporated to contextualize specific applications. For example, in games, dangerous or beneficial objects deserve greater attention from viewers compared with those that are merely decorative. Some objects can therefore be unconventionally focused. Added value is thus individually factored for each scene object by application models.

In a game, for instance, a dangerous or beneficial object deserves more observer attention than those that are purely decorative. Thus, depending on the context, objects can be focused in a non-conventional manner. The earned value technique should be individually attributed to each object on a scene by either an application modeler.

4 Experiment

In order to validate our proposal through a user‘s experiment, a virtual environment was developed using the Unity 3D game engine (Figure 5) and the Oculus Rift HMD. The user‘s navigation through the environment was captured solely by head movement analysis.

Figure 5: User experience session

The virtual environment has two interactive scenarios. The first scenario of the environment simulates a first-person virtual navigation in a scene that exhibits a set of eye-charts identified by numbers in a yellow region above each chart and disposed non-sequentially (Figure 6).

As illustrated in Figure 7, each eye-chart has ten numbered lines using successively smaller font sizes.

During the first phase of the experiment, an external evaluator informed, also in non-sequential order, the number identifying the next eye-chart and its corresponding line to be read by the user. Thus, the user had the task of seeking the corresponding chart by exploring the scenario and of reading one of its lines once found.

This task is completed once the user has found and read 12 eye-charts in the informed order. Table 1 details the order, line number, and the eye-chart content. Middle font sizes were chosen for the test as prior tests indicated that smaller font sizes were not read by some users with visual impairments, while larger font sizes were very easily read.

Once finished the first task, the user was informed to turn 180^0 in order to visit the second scenario of the virtual environment (Figure 8). In this scenario users were instructed to find another point in the scene containing several digital signboards similar to those found in airports. Users were then asked to read (spell) only one of the digital signboards. All digital signboards continuously switched

Table 1: Task description of scenario I: non-sequential order of the eye-chart visit, line to be read (from 5 to 7 in increasing level of difficulty) and corresponding content

Eye-chart	Line	Content
Eye-chart 3	5	IONIJD
Eye-chart 2	6	DCKLNXNG
Eye-chart 1	7	VPQOERERQK
Eye-chart 6	5	UYSDCK
Eye-chart 5	6	BMPQLGFE
Eye-chart 4	7	QANLHMBSQT
Eye-chart 9	5	LDUOKL
Eye-chart 8	6	OGKJJVJH
Eye-chart 7	7	SLIAQZXJPP
Eye-chart 12	5	DIJFEL
Eye-chart 11	6	JXKBXONQ
Eye-chart 10	7	JAQZBMABXQ

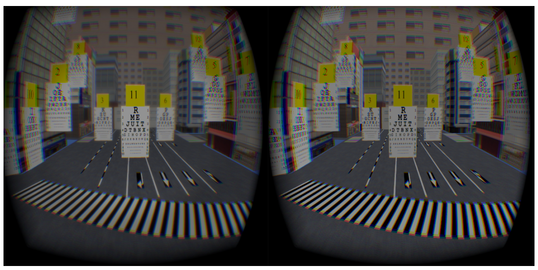

Figure 6: Virtual environment - scenario I : the user was supposed to read lines from a set of eye-charts disposed non-sequentially and identified by numbers in yellow regions (on left : plugin enabled, on right: plugin disabled).

between texts, and the users were instructed to focus on only one digital signboard. To ensure that the users were focusing on a specific digital signboard, they were asked to spell the three first letters while the texts were in motion for a period of 2 minutes. To make the task even more complex, demanding more user concentration, an evaluator turned on particle emission (simulating snow and flies) on the screen in the last 60 seconds of this task (Figure 9).

Details about the evaluation method applied to the experiment and its corresponding results are presented in the section 5.

5 Users Evaluation

To evaluate our model and application aimed at reducing HMD-device-generated discomfort, user evaluation was performed based on a simulator sickness questionnaire (SSQ) (Kennedy et al, 1993) with 16 symptoms of discomfort. The evaluation consists of the following tasks (Figure 10) in a sequential manner.

- Filling a profile questionnaire
- Filling an SSQ (Q1)
- Completion of the first test session (S1) using an HMD device
- Filling an SSQ (Q2)
- Completion of the second test session (S2) using an HMD device
- Filling an SSQ (Q3)

The profile questionnaire consisted of several items such as instruction degree, previous 3D movie screening, age, and previous use of the technology. 1 female and 23 male users, aged 18-50 years, underwent the process; of them, 37.5% had no experience in using the technology, and 62.5% had previously utilized this type of immersive technology.

In order to minimize the effect of discomfort related to the usage time the users were divided into two groups: group 1 attended S1 with the plugin disabled and S2 with the plugin enabled, and group 2 performed the tasks in a reverse order. Both sessions, S1 and S2, were defined by identical tasks as described in Section 4.

6 Results

In all cases, tests revealed that HMD-device-generated discomfort during the sessions when the plugin was enabled as lower than that during the sessions when the plugin was disabled.

6.1 First Analysis

The Symptoms in the Tables 2 and 3 are divided in four levels: 0 - None, 1 - Slight, 2 - Moderate, 3 - Severe.

In the first group the users began the experiment with the plugin

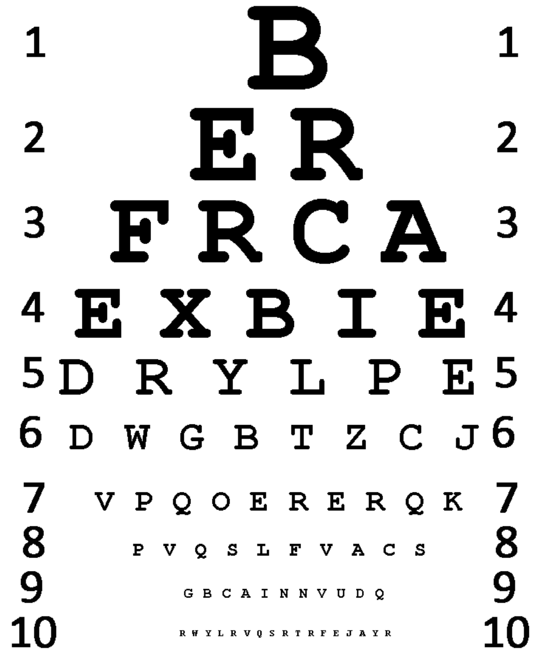

Figure 7: Eye Chart.

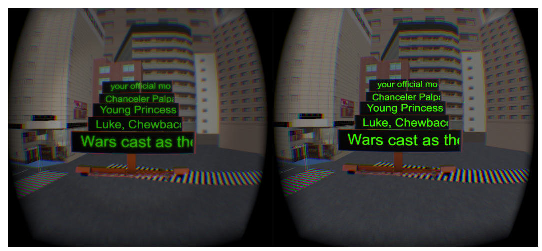

Figure 8: Virtual environment - scenario II: the user was supposed to spell the first 3 letters from the words in middle line of the the billboard during 2 minutes (on left : plugin enabled, on right: plugin disabled).

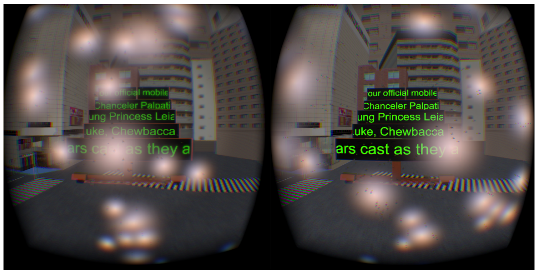

Figure 9: Virtual environment - Particles (on left : plugin enabled, on right: plugin disabled).

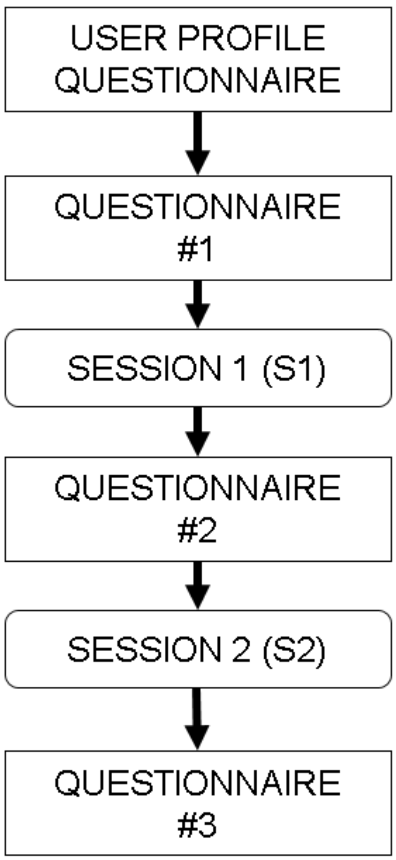

Figure 10: Tasks fluxogram

unactivated in the first session and activated on the second one. All discomfort occurrences was added . The first questionnaire (Q1) refers to results of user‘s discomfort before the sessions. Second and third questionnaires (Q2 and Q3) refers to results after sessions 1 and 2 (S1 and S2).

The results of group one questionnaires are shown in Table 2. According to the results, there was a significant increase of discomfort related cases from Q1 (before use) to Q2 (after session 1 from group 1 when plugin was unactivated).

Like the previous , the group 2 also fulfilled the three discomfort questionnaires. The results of questionnaires from Group 2 are seen in Table 3. After the group 2 first session (with depth of field plugin on), users completed the second questionnaire (Q2). Similar as occurred in group 1, an increased occurrences related to discomfort was occuried.

Then, the plugin was again disabled in session 2. According to the third questionnaire (Q3) responses, it was observed the high

Table 2: Results of questionnaires (Q1, Q2 and Q3) from group 1.

Group 1	**Q1**				**Q2**				**Q3**			
Symptoms	0	1	2	3	0	1	2	3	0	1	2	3
1. General discomfort	12	-	-	-	11	1	-	-	9	3	-	-
2. Fatigue	8	4	-	-	5	7	-	-	4	8	-	-
3. Boredom	8	4	-	-	9	3	-	-	8	4	-	-
4. Drowsiness	10	2	-	-	12	-	-	-	12	-	-	-
5. Headache	12	-	-	-	11	1	-	-	11	1	-	-
6. Sweating	11	1	-	-	12	-	-	-	12	-	-	-
7. Nausea	12	-	-	-	10	2	-	-	11	-	1	-
8. Difficulty concentrating	12	-	-	-	7	4	1	-	6	6	-	-
9. "Fullness of the head"	12	-	-	-	10	2	-	-	9	2	1	-
10. Blurred vision	12	-	-	-	7	5	-	-	10	2	-	-
11. Dizziness eyes	12	-	-	-	9	3	-	-	10	1	1	-
12. Vertigo	12	-	-	-	11	1	-	-	10	1	1	-
13. Visual flashbacks	12	-	-	-	12	-	-	-	12	-	-	-
14. Faintness	12	-	-	-	12	-	-	-	12	-	-	-
15. Stomach awareness	12	-	-	-	11	1	-	-	11	1	-	-
16. Other	12	-	-	-	12	-	-	-	12	-	-	-
Sum of Discomfort	**11**				**31**				**33**			

Table 3: Results of questionnaires (Q1, Q2 and Q3) from group 2.

Group 2	**Q1**				**Q2**				**Q3**			
Symptoms	0	1	2	3	0	1	2	3	0	1	2	3
1. General discomfort	12	-	-	-	12	-	-	-	11	1	-	-
2. Fatigue	9	-	3	-	8	4	-	-	5	5	2	-
3. Boredom	11	1	-	-	10	1	1	-	10	2	-	-
4. Drowsiness	5	5	1	1	9	1	1	1	8	2	1	1
5. Headache	12	-	-	-	10	2	-	-	10	1	1	-
6. Sweating	11	-	1	-	10	2	-	-	10	2	-	-
7. Nausea	12	-	-	-	12	-	-	-	10	1	1	-
8. Difficulty concentrating	9	3	-	-	6	6	-	-	4	7	-	1
9. "Fullness of the head"	11	1	-	-	10	2	-	-	8	4	-	-
10. Blurred vision	12	-	-	-	9	3	-	-	8	4	-	-
11. Dizziness eyes	12	-	-	-	10	2	-	-	8	4	-	-
12. Vertigo	12	-	-	-	12	-	-	-	10	2	-	-
13. Visual flashbacks	12	-	-	-	12	-	-	-	11	1	-	-
14. Faintness	12	-	-	-	12	-	-	-	12	-	-	-
15. Stomach awareness	12	-	-	-	12	-	-	-	10	2	-	-
16. Other	12	-	-	-	12	-	-	-	11	-	1	-
Sum of Discomfort	**16**				**26**				**46**			

increase in discomfort occurrences.

According to the results, there was a high increase of discomfort related cases when the plugin was unactivated in both groups.

6.2 Second Analysis

The discomfort values on the figures bellow were obtained from the derivative equation of each answered questionnaire, where "i" is the index of question:

- Results S1 [i] = (Q2[i]) - (Q1[i])
- Results S2 [i] = (Q3[i]) - (Q2[i])

The discomfort level is noticeably higher when the plugin was deactivated for the majority of users. In the graphs of Figures 11 and 12 we can notice a small level of discomfort during the sessions where plugin was enabled. However, discomfort level increase once the plugin was disabled. Fatigue, difficulty concentrating and blurred vision were the symptoms affected positively by the plugin use.

It was found that despite a constant increase in discomfort levels with the plugin enabled, increase in these levels was faster with the plugin disabled. These tests denote that discomfort levels increase proportionally to device usage time. However, the use of the developed plugin eases discomfort by reducing its growth speed.

7 Conclusion

This work is motivated by the widespread use of immersion devices such as HMD in virtual environments and the need for development of new strategies to reduce the level of visual discomfort caused by such devices.

Among several discomfort-causing factors, our study focused on simulation of human vision focus of attention. More specifically, our proposal presents a heuristic model for finding objects of interest dynamically and simulates attention by responding to focus in real time with depth of field effects.

It extends related studies that use the depth of field effects at fixed spots (mostly at the scene center). A more complex model is presented that is capable of inferring results from a relatively greater amount of application information and context, for instance according to a game strategy.

Furthermore, in order to validate it, a scenario of a virtual tour for case study was built. In it, the user is supposed to complete tasks related to visual attention. As another contribution, we adapted a well established simulator sickness questionnaire (SSQ) to the context of measuring HDM systems and used it to observe a group of 24 users (23 male, 1 female) on the case study tour. Results show that the plugin use reduced user discomfort when compared to the control group performing the same tasks without its use.

We believe that there is still space for the development of new heuristics for the object of interesting/focus selection specially considering particularity of distinct games or VR applications. In that sense, a public domain plugin was development for the Unity 3D game engine driven by the proposed model concepts (link removed for blind review issues). Such plugin allows developers to adjust their own heuristic as needed.

References

[1] Jonh Carmacks Delivers Some Home Truths On Latency. http://oculusrift-blog.com/john-carmacks-message-of-latency/682/. Accessed: 2016-06-01.

[2] K. Carnegie and T. Rhee. Reducing visual discomfort with hmds using dynamic depth of field. *Computer Graphics and Applications, IEEE*, 35(5):34–41, 2015.

[3] A. T. Duchowski, D. H. House, J. Gestring, R. I. Wang, K. Krejtz, I. Krejtz, R. Mantiuk, and B. Bazyluk. Reducing visual discomfort of 3d stereoscopic displays with gaze-contingent depth-of-field. In *Proceedings of the ACM Symposium on Applied Perception*, pages 39–46. ACM, 2014.

[4] S. Hillaire, A. Lécuyer, R. Cozot, and G. Casiez. Depth-of-field blur effects for first-person navigation in virtual environments. In *Proceedings of the 2007 ACM symposium on Virtual reality software and technology*, pages 203–206. ACM, 2007.

[5] M. Kass, A. Lefohn, and J. Owens. Interactive depth of field using simulated diffusion on a gpu. *Pixar Animation Studios Tech Report*, 2:1–8, 2006.

[6] 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.

[7] S. Kottravel, M. Falk, E. Sunden, and T. Ropinski. Coverage-based opacity estimation for interactive depth of field in molecular visualization. In *Visualization Symposium (PacificVis), 2015 IEEE Pacific*, pages 255–262. IEEE, 2015.

[8] J. Li, M. Barkowsky, J. Wang, and P. Le Callet. Study on visual discomfort induced by stimulus movement at fixed depth on stereoscopic displays using shutter glasses. In *Digital Signal Processing (DSP), 2011 17th International Conference on*, pages 1–8. IEEE, 2011.

[9] S. Liu and H. Hua. Time-multiplexed dual-focal plane head-mounted display with a liquid lens. *Optics letters*, 34(11):1642–1644, 2009.

[10] K. Selgrad, C. Reintges, D. Penk, P. Wagner, and M. Stamminger. Real-time depth of field using multi-layer filtering. In *Proceedings of the 19th Symposium on Interactive 3D Graphics and Games*, pages 121–127. ACM, 2015.

[11] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of vision*, 11(8):11–11, 2011.

[12] S. Yano, M. Emoto, and T. Mitsuhashi. Two factors in visual fatigue caused by stereoscopic hdtv images. *Displays*, 25(4):141–150, 2004.

[13] T. Zhang, H. T. Nefs, and I. Heynderickx. Human discrimination of depth of field in stereoscopic and nonstereoscopic photographs. *Perception*, 43(5):368–380, 2014.

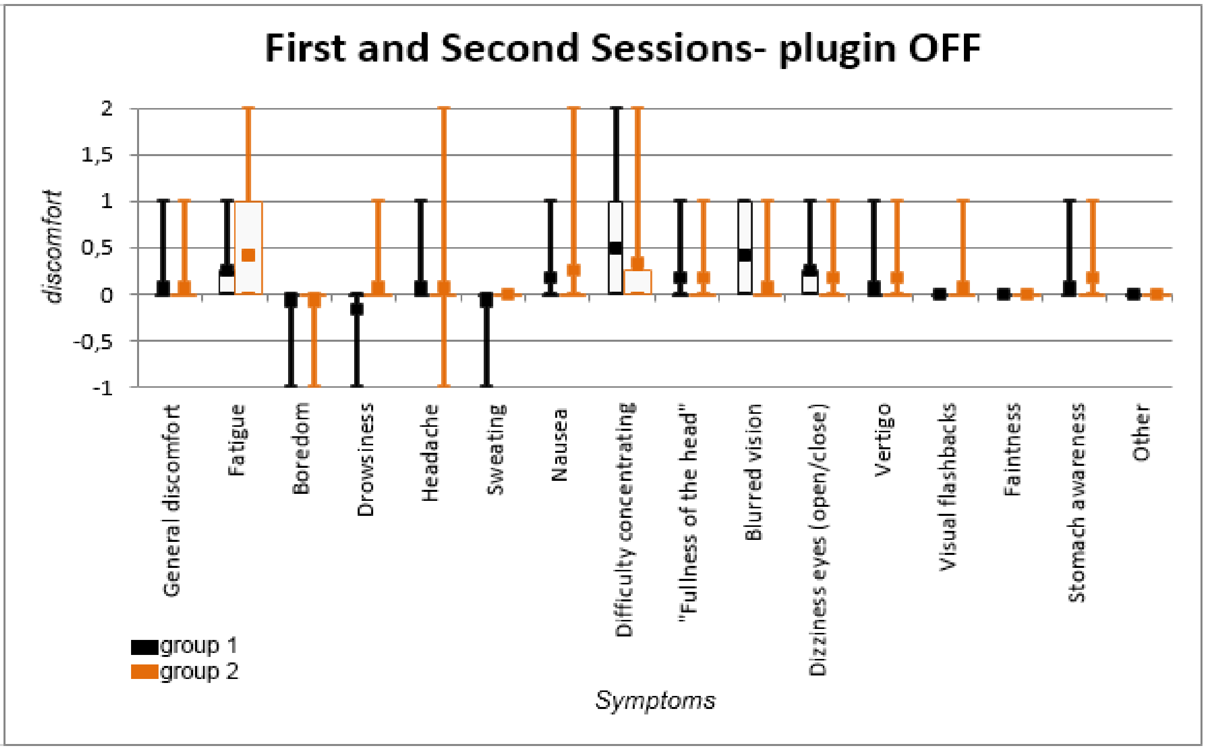

Figure 11: Discomfort results obtained from both groups when the plugin was disabled.

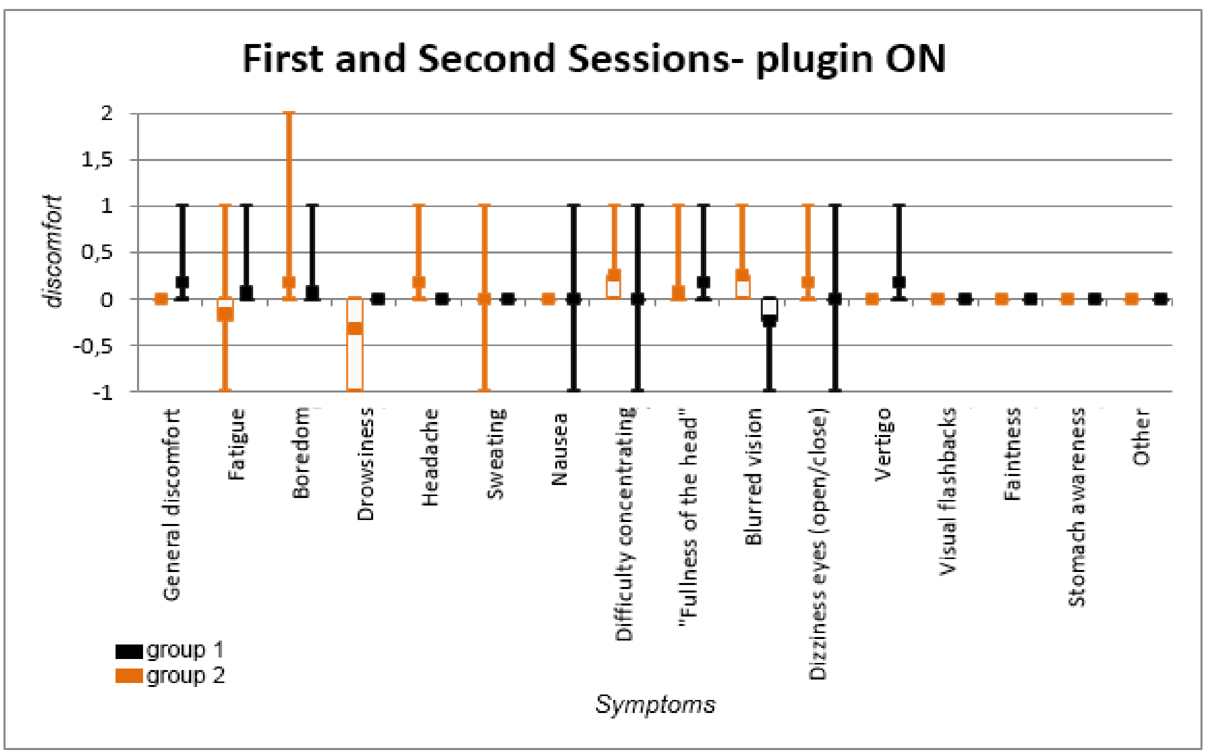

Figure 12: Discomfort results obtained from both groups when the plugin was activated.