

**REP. OF TURKEY
TED UNIVERSITY
GRADUATE SCHOOL
INTERACTIVE COMPUTING AND INFORMATION SYSTEMS**

**INVESTIGATING PHYSIOLOGICAL AND AFFECTIVE
RESPONSES IN VIRTUAL REALITY SCENES AND
ENHANCING USER EXPERIENCE WITH A
GAZE-DIRECTED METHOD**

BERK CEBECİ

ANKARA, DECEMBER 2020

INVESTIGATING PHYSIOLOGICAL AND AFFECTIVE RESPONSES IN
VIRTUAL REALITY SCENES AND ENHANCING USER EXPERIENCE
WITH A GAZE-DIRECTED METHOD

A Thesis Submitted To
The Graduate School
Of
TED University

by

Berk Cebeci

In Partial Fulfillment Of The Requirements
For
The Degree Of Master Of Science
In
Interactive Computing and Information Systems

Ankara, December 2020

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Berk Cebeci

Signature :

ABSTRACT

INVESTIGATING PHYSIOLOGICAL AND AFFECTIVE RESPONSES IN VIRTUAL REALITY SCENES AND ENHANCING USER EXPERIENCE WITH A GAZE-DIRECTED METHOD

BERK CEBECİ

M.S., Interactive Computing and Information Systems

Supervisor : Prof. Dr. Tolga. K. Çapın

Co-Supervisor : Assist. Prof. Dr. Ufuk Çelikcan

December 2020

Virtual Reality (VR) is becoming more common and growing its application areas in many fields. Aside from professional usage, recent releases of commercially available head-mounted displays (HMDs) increased the popularity of VR among consumers. However, VR can cause unwanted effects and fixed stereoscopic camera parameters cause discomfort and shallow depth feeling. Thus, achieving comfortable and high quality stereoscopic 3D view became more crucial. In this thesis, we aim to investigate the effects of VR in detail and develop a method to improve the quality of VR experience. In the first part of the thesis, we investigate the physiological and emotional changes caused by VR. We also take account the cybersickness and analyze these measures jointly. For the second part of the thesis, we propose a method which modifies stereoscopic camera parameters of HMD based on gaze-tracking in real time. We conduct a user study that demonstrates improvements of our method in terms of visual quality, visual comfort, perceived depth and overall experience of the VE. The results show that our method significantly enhances VR experience compared to fixed default parameters.

Keywords: Virtual Reality, Eye Tracking, Stereoscopic 3D, User Experience, Physiological and Affective Responses

ÖZET

SANAL GERÇEKLİK SAHNELERİNDEKİ FİZYOLOJİK VE DUYGUSAL TEPKİLERİN ARAŞTIRILMASI VE BAKIŞ GÜDÜMLÜ YÖNTEM İLE KULLANICI KALİTESİNİN İYİLEŞTİRİLMESİ

BERK CEBECİ

M.S., İnteraktif Bilişim Sistemleri

Tez Yöneticisi : Prof. Dr. Tolga. K. Çapın

Ortak Tez Yöneticisi : Assist. Prof. Dr. Ufuk Çelikcan

Aralık 2020

Günümüzde Sanal Gerçeklik (VR) yaygınlaşmakta ve uygulama alanlarını genişletmektedir. Profesyonel kullanım dışında, piyasada satışa çıkarılan 3 boyutlu (3B) gözlükler VR popülerliğini kullanıcılar arasında da arttırmıştır. Ancak, VR kullanımı istenmeyen etkilere sebep olurken sabit çiftli (stereo) kamera parametreleri rahatsızlığa sebep olup derinlik hissini azaltmaktadır. Bu nedenle, konforlu ve yüksek kaliteli stereo 3B görüntü elde edilmesi önem kazanmıştır. Bu tezde, VR etkilerini detaylı olarak inceleyip ve VR tecrübesinin kalitesini arttıracak bir yöntem geliştirmeyi amaçladık. Tezin ilk bölümünde, VR tarafından ortaya çıkan fizyolojik ve duygusal değişimleri araştırdık. Sanal gerçeklik tutmasını da dikkate alarak sonuçların analizini yaptık. Tezin ikinci bölümü için, VR gözlüğünün stereo kamera parametrelerini gerçek zamanlı göz takibine göre değiştiren yöntemimizi sunduk. Yöntemimizi kullanıcı testleriyle test etmiş ve görsel kalite, görsel konfor, hissedilen derinlik ve genel VR tecrübesindeki iyileştirmelerini göstermiş bulunmaktayız. Bu sonuçlar, metodumuzun sabit kamera parametreleriyle karşılaştırıldığında VR tecrübesini önemli ölçüde iyileştirdiğini göstermiştir.

Keywords: Sanal Gerçeklik, Göz Takibi, Stereo 3D, İzleyici Konforu, Fizyolojik ve Duygusal Tepkiler

To my family

ACKNOWLEDGMENTS

First of all, I would like to express my gratitude to my advisor Prof. Dr. Tolga K. apın who has been my mentor even before I started this work. I feel very lucky to have found the chance to work with him. Without his guidance and positive attitude, it would not be possible to complete this thesis. I am also thankful to my co-advisor Assist. Prof. Dr. Ufuk elikcan for his valuable feedback and assistance. His support in all aspects of this work is invaluable.

I am thankful to Assist. Prof. Dr. Abdullah Blbl and Assist. Prof. Dr. Ulař Gle for accepting to be in my thesis committee, and for their valuable feedback and suggestions.

This work is part of the “Enhancing the User Experience of 3D Displayed Virtual Scenes” project which was supported by TUBITAK (Scientific and Technological Research Council of Turkey) 1001 program (Grant No: 116E280). I am thankful to professors Assist. Prof. Dr. Ufuk elikcan, Prof. Dr. Hařmet Gray and Prof. Dr. Tolga K. apın for conducting this project. I also want to thank to TUBITAK 1001 program members Emre Avan and Dilara Albayrak for their help.

I express my sincere thanks to my friends for giving me encouragement and moral support.

Last but not least, I am grateful to my family for their continuous support and unparalleled love throughout my years of study and my whole life. This accomplishment would not have been possible without them.

TABLE OF CONTENTS

PLAGIARISM	iv
ABSTRACT	v
ÖZET	vi
DEDICATION	vii
ACKNOWLEDGMENTS	viii
TABLE OF CONTENTS	ix
LIST OF TABLES	xii
LIST OF FIGURES	xiii
LIST OF SYMBOLS AND ABBREVIATIONS	xvi
CHAPTERS	
1 INTRODUCTION	1
1.1 Scope of the Thesis	3
1.2 Outline of the Thesis	4
2 PHYSIOLOGICAL AND AFFECTIVE RESPONSES CAUSED BY DYNAMIC VIRTUAL ENVIRONMENTS	5
2.1 Related Works	7
2.2 Experimental Evaluation	9
2.2.1 Equipment	9

2.2.2	Virtual Environments	9
2.2.3	Subjects	10
2.2.4	Procedure	11
2.3	Information Gathered	13
2.4	Results	14
2.5	Threats to Validity	21
2.6	Discussion	22
2.6.1	Related Publications	23
3	OPTIMIZING STEREOSCOPIC PARAMETERS IN REAL TIME USING A GAZE-DIRECTED METHOD	24
3.1	Background	25
3.1.1	Stereo Geometry	25
3.1.2	Stereoscopic Scene Rendering Method	28
3.1.3	Accommodation and Convergence Conflict	28
3.2	Related Work On Improving Stereo Viewing Comfort	28
3.3	Approach	31
3.3.1	Gaze-Directed Stereoscopic Camera Control	31
3.3.2	Stereopsis Limits Calculation	32
3.3.3	Camera Convergence Distance Modification	32
3.3.4	Optimizing the Interaxial Camera Distance	34
3.3.5	Projection Matrix Modification	36
3.4	Experimental Evaluation	37
3.4.1	Equipment	37

3.4.2	Virtual Environments	37
3.4.3	Subjects	37
3.4.4	Procedure	38
3.5	Results	40
3.5.1	Frame Rate	40
3.5.2	Evaluation Scores	41
3.5.3	Statistical Analysis	41
3.5.4	Comparing With Saliency Guided Camera Control Method	44
3.6	Threats to Validity	44
3.7	Discussion	45
4	CONCLUSION	46
4.1	Future Work	47
	REFERENCES	48
	APPENDICES	
A	VAS QUESTIONNAIRE	54
B	SIMULATOR SICKNESS QUESTIONNAIRE	55
C	SCENE EVALUATION QUESTIONNAIRE	56
D	ETIK KURUL KARARI	57

LIST OF TABLES

TABLES

Table 2.1	Differences in physiological measurements and questionnaire outcomes between the VEs. Both groups are in descending significance order.	16
Table 2.2	SSQ scores and VAS ratings before and after the VEs.	18
Table 2.3	Correlations between physiological measures and questionnaire scores for each VE.	21
Table 3.1	Averages of participants' ratings on 7-point Likert scale (a) for CR (b) for SS.	43

LIST OF FIGURES

FIGURES

Figure 2.1	Sample frames from (a) Campfire VE (b) Hospital VE (c) Roller Coaster VE	10
Figure 2.2	User study flow.	11
Figure 2.3	Average pupil diameter length for each brightness step. The error bars represent the corresponding ± 1 standard deviation.	13
Figure 2.4	Mean curves of blinks per 5 seconds, fixations per 5 seconds, momentary heart rate and momentary pupil dilation for the three VEs. (a) Campfire VE (b) Hospital VE (c) Roller Coaster VE. The horizontal axis represents the time passed since the start of a VE in seconds.	15
Figure 2.5	Color bars indicate means for physiological response measures and questionnaire outcomes in the three VEs: (a) Pupil dilation (b) Heart rate (c) Blink count (d) Saccade mean speed (e) Saccade rate (f) Fixation duration (g) Change in total SSQ score (h) Change in happiness feeling (i) Change in sadness feeling (j) Change in comfort feeling (k) Change in anxiety feeling. The error bars represent the corresponding ± 1 standard error of the mean.	17
Figure 3.1	Accommodation and convergence conflict	26
Figure 3.2	Different stereoscopic scene rendering methods	27
Figure 3.3	Flow diagram of gaze-directed method	31
Figure 3.4	Stereopsis limits calculation steps	33

Figure 3.5	Sample frames from (a) CR and (b) SS	38
Figure 3.6	Eye tracker calibration points	39
Figure 3.7	Ranges of frame rates of sessions are shown with blue bars. Red dots mark the averages.	40
Figure 3.8	Ranges of scores given by the user study participants are shown with blue bars for (a) image quality, (b) perceived depth, (c) visual comfort, and (d) overall experience. Red dots mark the respective averages.	42

LIST OF SYMBOLS AND ABBREVIATIONS

Symbols

d_{max}	Maximum Disparity
d_{min}	Minimum Disparity
f_c	Focal Length of the Cameras
t_c	Interaxial Separation
Z_c	Convergence distance

Abbreviations

3D	Three-dimensional
CF	Campfire Environment
CR	Classroom Environment
DOF	Depth of Field
ECG	Electrocardiogram
EEG	Electroencephalogram
EGG	Electrogastrogram
EOG	Electrooculogram
GSR	Galvanic Skin Response
HH	Hospital Environment
HMD	Head-mounted Display
MSSQ	Motion Sickness Susceptibility Questionnaire

PANAS	Positive and Negative Affect Schedule
PPG	Photoplethysmogram
RC	Roller Coaster
SS	Suburban Street
SSQ	Simulator Sickness Questionnaire
VAS	Visual Analogue Scale
VE	Virtual Environment
VR	Virtual Reality

CHAPTER 1

INTRODUCTION

Rapid technological advances in recent years led Virtual Reality (VR) to become more common in many different professional fields such as education, health, archaeology and entertainment [1]. Aside from professional fields, major technology companies released commercially available head-mounted displays (HMDs) such as Oculus Rift, HTC Vive and Sony PlayStation VR. Due to the increase in available commercial HMDs and their compatibility with widely used game engines such as Unity 3D, led developers to create VR specific interactive applications and games which resulted in a rapid increase in VR's popularity. Achieving a high quality and comfortable stereoscopic 3D view has gained more importance in real time VR applications. However, there are still unresolved issues with the VR technology which prevent it to reach its full potential. Since VR is becoming more accessible and its application areas are growing, solving these problems are crucial for achieving high quality experience in HMDs.

Due to immersive nature of HMDs, virtual environments (VEs) are more engaging and believable to users. VEs can be used to help people overcome their phobias [2] or positively affect post-traumatic stress disorder patients [3].

Apart from positive effects, VEs can cause unwanted physiological and affective responses on humans, such as causing anxiety [4]. Considering VEs effects on the people's emotional states, it is very important to monitor people's feelings accurately so that developers can modify the VEs according to this information, thus improve the general quality of the scenes.

A widely use method to measure a user's emotional state is using questionnaires [5]

[6]. In addition to that, the user's physiological measurements such as heart rate and pupil dilation also can be used for this purpose [7, 8, 9].

VEs do not only cause emotional changes. A phenomenon called cybersickness causes unpleasant symptoms [10] after using HMDs for a period of time, thus decreasing the quality of VR experience greatly. Similar to emotional state measurements, cybersickness can be detected using questionnaires [11] and physiological measurements [12, 13].

Another issue with commercial HMDs is the fixed stereoscopic parameters. There are two important parameters while creating a stereoscopic vision in VR: interaxial separation, which is the distance between two virtual cameras, and convergence distance of the virtual cameras. These parameters have a major role on perceived depth and comfort of the VE, however they have fixed values in commercially available HMDs. For the purpose of achieving maximum depth perception and comfort, an optimization is needed on scene parameters. Since there is not a single setting for the optimized scene depth, these parameters have to be adjusted based on scene's depth range and content.

Increasing the perceived depth and general experience without causing discomfort is a challenge in stereoscopic VEs. Detecting the region which has viewer's attention and setting stereoscopic parameters according to that is a method used for enhancing scene depth without invoking discomfort [14]. Modifying stereoscopic parameters can be done by estimating the viewer's gaze positions [15] [16]. Another method is using an eye tracker which provides more accurate gaze points compared to estimation [17] [18].

While these issues are present, there are not many work done in VR field regarding investigating physiological and affective responses caused by VEs and evaluating a stereoscopic parameter optimization algorithm based on gaze data from eye tracker which works in real time in terms of depth perception and general improvements. Within this study, first we aimed to investigate physiological and affective responses caused by VEs which are specifically designed to induce certain emotions and cybersickness symptoms. We conducted a user study and gathered data by measuring physiological data of participants while they are experiencing VEs and collecting

their self assessments. We analyzed the data we collected and investigated their relationship with scene context in a comprehensive study and we reported our findings. We also developed a gaze-directed stereoscopic camera control method which works in real time and uses viewer's gaze points to modify VE's stereoscopic parameters. Furthermore we conducted another user study to evaluate the quality of our method by comparing it with default HMD settings in two VEs. We reported our findings after analyzing participants' self assessments.

1.1 Scope of the Thesis

Within this thesis, first we aimed to investigate and analyze the physiological and affective responses caused by experiencing VR scenes in a comprehensive study. For this purpose, we created (i) an emotionally neutral VE, (ii) a VE which induces unpleasant emotions and (iii) a cybersickness symptoms causing VE where all have different properties. We concentrated on finding which physiological measurement can be used for detecting viewer's emotional state and whether they are feeling cybersickness symptoms or not. We conducted an extensive analysis and reported our findings in detail.

We also intended to propose a method which optimizes stereoscopic scene parameters in real time based on viewer's gaze data without requiring any pre-processing, in order to improve the scene depth and comfort. Our goal was to find strengths and improvements of our method and for this purpose, we compared it with the default fixed parameters setup of HTC Vive by conducting a user study where viewers experience an indoor and an outdoor interactive VE. We reported our method's performance after a comprehensive analysis.

The main contributions of the thesis are given below:

- We analyze the physiological and affective responses induced by immersive dynamic VEs through an extensive analysis and investigate the interaction between cybersickness and emotions.
- We conduct a formal user study using VEs with varying context and collect

broad set of measures of physiological and affective responses.

- We propose a method to optimize stereoscopic camera parameters in real time using viewer's gaze input.
- We present a formal study to compare our dynamically modified stereoscopic parameters with default VR parameters in terms of various subjective assessments (image quality, perceived depth, visual comfort) using two unique VEs with different context.

1.2 Outline of the Thesis

The rest of the chapters are outlined below.

In Chapter 2, first we briefly introduce our study and present an overview of previous works. We explain our experiment and data collection in detail. The results of our user study are presented and discussed.

In Chapter 3, we explain our second study and give background information regarding components of our work and address previous works. Our proposed method is explained in detail. We evaluate our method after a formal user study and discuss our findings.

Finally, in Chapter 4, we conclude our work by our evaluated results. We suggest further improvements based on our findings.

CHAPTER 2

PHYSIOLOGICAL AND AFFECTIVE RESPONSES CAUSED BY DYNAMIC VIRTUAL ENVIRONMENTS

Virtual reality (VR) is improving and becoming widely used in many different professional fields such as education, health, archaeology and entertainment [1]. Furthermore, VR technology has become more popular among consumers after the releases of commercially available VR headsets. Unfortunately, experiencing virtual environments (VEs) can cause variety of unwanted physiological and affective responses on humans. Physiological responses can range from change in heart rate and pupil diameter, affective responses are related to emotions and mood. Some of these responses can effect the quality of VR experience and reduce it greatly. Monitoring and reducing these responses is crucial for improving user experience quality in VR.

Many researchers have conducted studies to investigate the users' responses caused by VEs with different context. For example, it is revealed that people felt the corresponding responses when they experience anxious and relaxing VEs [4]. Additionally, various studies showed that VEs are effective enough to change a participant's behaviour towards a condition they fear, which makes using VEs to overcome phobias possible [2]. Apart from phobias, some studies revealed that VR usage on post-traumatic stress disorder patients have positively affected their treatment [3].

As previous studies showed, VR environments can cause a great change in user's emotional state, which is why it is really important to measure people's feelings correctly while they are experiencing a VE. With correct measurements, VE developers could use the user's emotional state to improve the user experience and VE quality, thus they will have more control over VEs they create. Many methods have been proposed and used for monitoring user's emotions in desktop environments. Ques-

tionnaires like Visual Analogue Scale (VAS) [6] and Positive and Negative Affect Schedule (PANAS) [5] are widely used in many studies. On the other hand, recording physiological measurements such as pupil dilation, heart rate, facial expressions, skin conductance, electroencephalogram (EEG) signals and body temperature [7, 8, 9]. is another widely used method.

VEs have another effect aside from changing user's emotional state which is a phenomenon called VR sickness, also called cybersickness. People who are experiencing VR sickness show symptoms such as disorientation, nausea, headaches, sweating and dizziness [10] which greatly reduces the quality of the VE and results in a poor overall user experience. Despite being a widely known phenomenon, conducted studies still cannot fully explain causes of cybersickness. Many theories try to explain it, most popular ones are sensory conflict and postural instability [19].

Apart from its causes, accurately detecting the cybersickness levels induced while experiencing a VE is crucial for improving the overall quality of VR experience. Simulator sickness questionnaire (SSQ) [11] is widely used for this purpose by the researchers in order to measure users' cybersickness level. Participants can rate the severity of 16 symptoms they feel during the VR experience with Nausea, Oculomotor and Disorientation subscales.

Using physiological measurements is another common alternative to the questionnaires. Heart rate, breathing rate, electrogastrogram (EGG), electrooculogram (EOG), electrocardiogram (ECG), photoplethysmogram (PPG), galvanic skin response (GSR) and skin temperature have been used to approximate cybersickness symptoms [12, 13].

Our purpose is to detect and analyze the physiological responses such as change in pupil diameter, blink count, heart rate, and affective responses for instance happiness and anxiety changes of participants while they are experiencing our VEs which are specifically designed to make them feel varying of emotions and cause cybersickness. To achieve our goal, we created 3 VEs with distinct contexts. Each virtual environment has a different aim; the first one is emotionally neutral campfire environment, the second one is a hospital environment that causes unpleasant symptoms and the third one is a roller coaster environment that causes cybersickness symptoms. Gaze

point, pupil diameter, heart rate and blinks were recorded while participants were experiencing these VEs. Eye fixations and saccadic movements were computed using the collected gaze data. Aside from physiological measurements, participants' self assessments of level of motion sickness sensitivity, emotional state and cybersickness symptoms were collected.

We investigated the relationship between objective physiological measurements and subjective self assessments and impact of the affective VEs we have created on these outcomes with a detailed approach, and examined how cybersickness interacted with emotions through a comprehensive analysis. In addition to that, the link between susceptibility to motion sickness and observed cybersickness was explored. The results of this work were published as a journal article [20] and presented in CASA (Computer Animation and Social Agents) 2019 conference.

2.1 Related Works

Many works have explored the physiological effects of VR. Kim et al. [21] gave a task to the participants in a virtual environment and monitored their electrophysiological signals while they are completing the task. Aside from objective measurements, in order to report the level of symptoms they were feeling, participants filled a simulator sickness questionnaire. Results have shown that increased cybersickness severity also increased physiological responses such as gastric tachyarrhythmia, blink rate and heart rate. Dennison, Wisti & D'Zmura [12] conducted a similar study and monitored participants' physiological changes during VR experience. They discovered that physiological responses such as changes in blinking, stomach activity and breathing are related to cybersickness severity and can be used to detect severity levels. Nalivaiko et al. [13] divided the participants into two groups based on their initial skin temperature. They found that skin temperature of the group with lower initial temperature increased due to cybersickness.

Emotional changes caused by VR have also been studied widely besides cybersickness. Riva et al. [4] explored the effects of VEs with distinct emotional context on users. They measured users' affective responses with questionnaires and discovered

that relaxing and anxious VEs caused corresponding feelings on participants. Meehan et al. [22] investigated how stressful and nonthreatening VEs affect users physiologically. They found out that heart rate is a reliable metric which can be used to measure presence in a VE.

While studies cited above explored how to detect the viewer's emotional state or if he/she is having cybersickness symptoms, they mostly required attaching sensors to users' body to monitor their physiological activity during VR sessions. Gaze data such as eye fixations, pupil dilation and saccadic eye movements are used as an alternative method for predicting user's current emotional state without attaching extra sensors to users. Bradley et al. [7] investigated effects of emotional arousal on pupil dilation while users are viewing desktop images. In their study, they showed neutral, pleasant and unpleasant images to the participants and found that pupil dilation changes were larger while viewing pleasant and unpleasant images compared to neutral images. After measuring skin conductance and heart rate, they also found a relation between skin conductance and pupil dilation. Similarly, Snowden et al. [23] showed neutral and fearful unpleasant images and observed that pupil diameter was larger after showing unpleasant images. Partala & Surakka [24] investigated how pupil diameter is affected by listening to arousing sounds. Participants listened to neutral, pleasant and unpleasant sounds and results showed that while participants were listening to pleasant and unpleasant sounds, their pupil sizes were larger. Gianakakis et al. [25] discovered that participants' blink rate increased while they were in stress and felt anxiety. McCall et al. [26] explored the gaze and body movements behavior in different VEs. Their results showed that while experiencing a virtual art museum, participants avoided looking at unpleasant images. They also discovered that in another VE containing realistic human models showing pleasant and unpleasant emotions, participants looked at the faces of human models showing unpleasant emotions.

As seen on works cited above, even though studied many times in desktop environments, very few research has been conducted in VR for investigating the effects of VEs with distinct contexts on pupil dilation. Chen et al. [27] explored the effects of different emotions on pupil diameter in VR by creating VEs with happiness, sadness, disgust and fear contexts and measuring pupil dilation while participants explore

those VEs. They also showed additional feedback to participants by measuring the heart rate and showing heart beat cues using combinations of visual, audio and haptic channels. They discovered that VEs with negative emotions, especially disgust VE, increased pupil dilation more than VEs with positive emotions. They also found out that showing the heart beat cues with haptic-visual feedback combination affected the pupil dilation greater than the other feedback combinations.

As mentioned earlier, an important advantage of monitoring pupil dilation to predict user's emotional state is that no extra equipment such as sensors attached to user in order to measure ECG or EGG signals is needed. A VR headset integrated with an eye tracker is sufficient for pupil measurements.

2.2 Experimental Evaluation

2.2.1 Equipment

We used a desktop computer with an Intel(R) Core(TM) i7-6700 CPU, 16 GB RAM, and a Nvidia GeForce GTX 1050 Ti graphics card. Our virtual scenes displayed in an HTC Vive head-mounted display (HMD) containing the Tobii Pro eye trackers. Integrated eye trackers have output frequency of 120 Hz with estimated accuracy of 0.5 degree. For the measuring heart rate of participants', we used Polar OH1 heart rate sensor.

2.2.2 Virtual Environments

We created three VEs with explicit themes for our study (Figure 2.1). Campfire environment (CF) has emotionally neutral elements and setting. Its purpose is to observe participants' behaviour in a VR environment which doesn't have any emotionally charged or cybersickness causing elements. Hospital environment (HH) is the second environment which is designed to induce unpleasant emotions on participants without causing cybersickness using horror themed elements such as terrifying characters, narrow and dark corridors, sound effects. Third environment, Roller Coaster (RC), is specifically created to cause cybersickness symptoms while remaining emotionally

neutral thus participants' physiological states can be measured while cybersickness is present but affective factors are not.

Virtual cameras in VEs moved through predefined paths, head and position tracking of the HMD was disabled for the purpose of ensuring that each participant precisely experiences the same content as they traverse our VEs thus minimizing uncontrolled factors in our experiment. All 3 VEs had similar lighting settings in order to keep pupil diameter changes based on light to minimum. Each VE took 125 seconds to traverse.

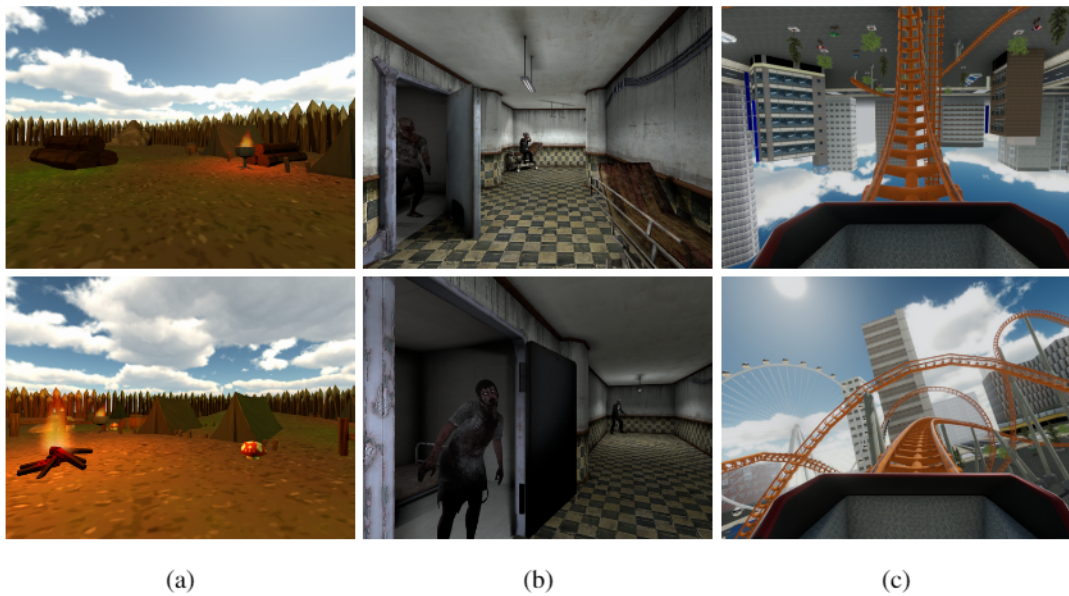


Figure 2.1: Sample frames from (a) Campfire VE (b) Hospital VE (c) Roller Coaster VE

2.2.3 Subjects

27 people volunteered for our study however, 7 of them with mean age 30.29 ± 6.32 (five males, two females) failed to complete the experiment. 1, 2 and 4 volunteers ended the experiment during CF, HH and RC VEs respectively. Their data are not included in our analysis. Remaining 20 volunteers, who will be referred as participants from this point, with mean age 27 ± 8.46 (15 males, five females) completed the study with success.

Using a five-point Likert scale questionnaire (0: “no previous experience,” 4: “very experienced”), we gathered all volunteers’ previous VR experience levels. Participants’ mean experience level was 0.75 ± 0.89 and 11 of the participants experienced VR for the first time.

As an additional precaution, participants were told not to consume caffeine [28] or smoke [29] on the day they will attend to our experiment since studies showed that these products significantly affect pupil responses such as dilation and speed.

2.2.4 Procedure

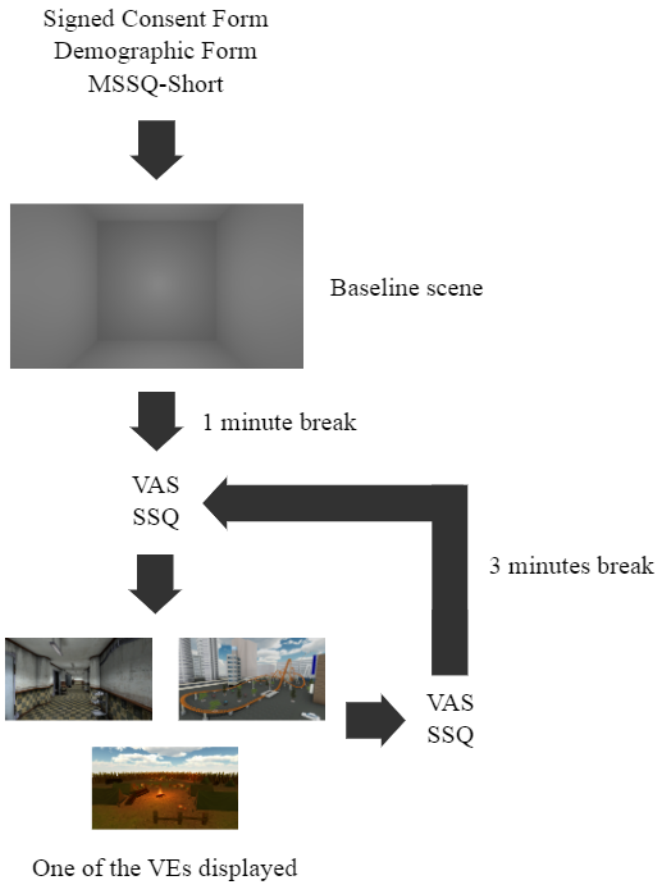


Figure 2.2: User study flow.

When participants first arrived, they were asked to sign a consent form. After their approval, they completed a demographic form containing questions regarding their names, ages, VR experiences and university departments they graduated. Following that, they were seated and filled the revised version of motion sickness susceptibility

questionnaire (MSSQ) by Golding [30]. Before the experiment started, participants were informed about the general instructions to consider while experiencing VR environments; then HMD was placed on their head and heart rate sensor was attached to their forearm.

After putting on the HMD, a baseline test scene is shown to participants before our 3 main VEs in order to eliminate the effects of brightness in pupil diameter measurement (Section 2.3). This scene is an entirely empty room without any elements in it. In the baseline test, virtual room is completely dark in the beginning and virtual camera in the room has a fixed 3D position (Figure 2.2). During the test, brightness of the room increased in 11 steps for 60 seconds and average pupil diameter of the participants recorded for each brightness level. After the baseline test, participants removed the HMD and took a one minute break.

After the break, participants put on the HMD again and experienced 3 VEs in a random order. Participants traversed CF and HH VEs only once, however In RC, they rode the roller coaster for 2 laps where coaster's speed increased in the second lap.

At the beginning of the each VE, general instructions about experiment displayed on HMD screen for 30 seconds. Following that, the eye tracker calibration phase started where 5 dots appeared on the screen one by one and participants had to follow them with their eyes. The calibration process repeated until it was successful in order to measure gaze data as accurately as possible. After the calibration succeeded, eye tracker position guide which shows the status of right and left eye trackers was displayed on the HMD screen and participants were asked to adjust the HMD until it was positioned correctly and accepted by the system. After the HMD positioned correctly, VE movement started. When participants reached the end point in VE and finished the movement, they were informed to remove the HMD. After completing each VE, participants took a three minutes break before starting experiencing the next VE.

Before and after experiencing each VE, participants filled a VAS questionnaire [6] where they rated how happy, sad, comfortable and anxious they felt at that moment in a seven-point Likert scale (1: "not at all," 7: "very") and SSQ [11] where they rated how severely they feel 16 different cybersickness symptoms with nausea, Oculomotor and Disorientation subscales (See Appendices A and B for details).

2.3 Information Gathered

While participants were experiencing each VE, their pupil diameter and gaze data for both eyes were recorded with 15.15 Hz sampling rate. Participants' blink counts were measured real time. On the other hand, using PyGaze library [31], saccade rate, saccade mean speed, fixation count and fixation duration were calculated offline. In order to have consistent measurements, blinks and fixations were counted within a window of 5 seconds.

To eliminate the effect of brightness on pupil diameter, measurements of the baseline test were used as in Chen et al.'s work [27]. First of all, for all 11 luminance levels from baseline scene, average pupil diameter of each participant was calculated. Following that, we captured screenshots in baseline scene whenever brightness changed and each second in other 3 VEs. All screenshots converted into grayscale images and their average luminance values were computed. Grayscale images from VEs were matched with images from baseline scene according to their luminance values.

For each participant, pupil diameter measurements related to a screenshot (measurements within 0.5 second of a screenshot) were subtracted from corresponding average pupil diameter value of that participant from matched baseline scene luminance level. We calculated pupil dilation caused by emotional arousal with this procedure. This

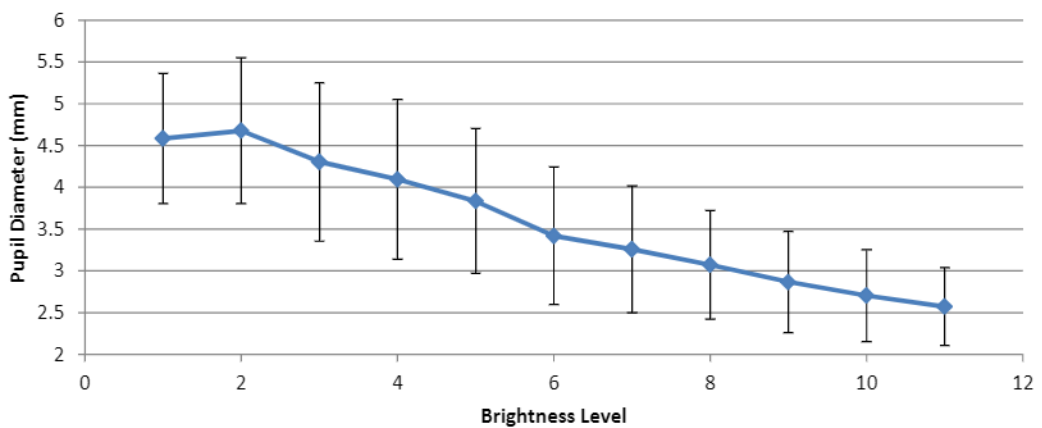


Figure 2.3: Average pupil diameter length for each brightness step. The error bars represent the corresponding ± 1 standard deviation.

difference gives the pupil dilation response, which we will be referred as pupil dilation, caused by emotions and cybersickness at given instant. This procedure carried on offline. For each luminance level in baseline scene, average pupil diameter of participants is shown in Figure 2.3.

Polar OH1 optical heart rate sensor was used to record a participant's heart rate once per second. The validity of the heart rate measurement by Polar OH1 was established with respect to the criterion measure in Schubert et al. [32]. MSSQ and SSQ scores were calculated as explained in Golding [30] and Kennedy et al. [11], respectively. VAS ratings were accumulated as is.

2.4 Results

Mean curves of momentary pupil dilation, blink count per 5 seconds, fixation count per 5 seconds and momentary heart rate as averaged over all participants' data throughout the three VEs is shown in Figure 2.4. With a running average of 10 samples, pupil dilation curves were smoothed in the figure.

Reflection of CF environment's neutral and calm setting compared to the other two VEs is noticeable on heart rate and pupil dilation curves. For HH, the pupil dilation curve reaches its the highest value (1.27 mm) and the largest average (0.77 mm) while the fixation curve has the lowest average (0.63). The heart rate and blink curves also have their lowest averages (0.61 and 76.28 bpm) and display higher variance (0.06 and 0.87 bpm^2) than for CF (0.03 and 0.22 bpm^2). During RC, the heart rate curve peaks (84.6 bpm) and exhibits the largest variance (7.28 bpm^2). Blink curve also displays the largest variance (0.13) in RC. The effect of the faster second half of RC can be clearly observed on the heart rate curve.

As seen on Table 2.1, physiological measurements and questionnaire outcomes as averaged over study sample showed that average pupil dilation was the highest in HH due to unpleasant emotional setting of the VE and the lowest in CF (Fig. 2.5.a) where environment does not have any emotionally charged elements. Cybersickness causing elements in RC resulted in the highest averages of heart rate (Fig. 2.5.b) and saccade mean speed (Fig. 2.5.d). Average blink count was the highest in CF and the lowest in

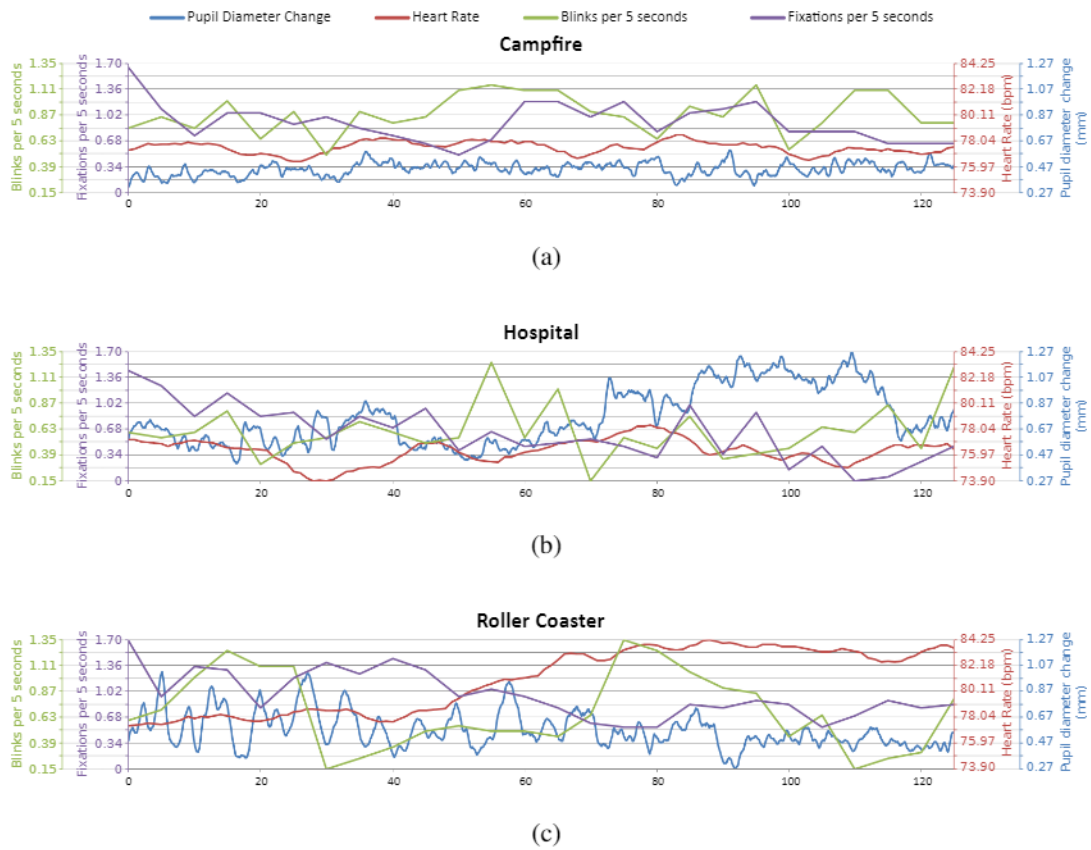


Figure 2.4: Mean curves of blinks per 5 seconds, fixations per 5 seconds, momentary heart rate and momentary pupil dilation for the three VEs. (a) Campfire VE (b) Hospital VE (c) Roller Coaster VE. The horizontal axis represents the time passed since the start of a VE in seconds.

HH (Fig. 2.5.c). Fixation duration was the highest in RC and the lowest in HH (Fig. 2.5.f). Saccade rate was the highest in HH (Fig. 2.5.e). These outcomes revealed that unpleasant emotions caused participants to blink and fixate less and make more saccadic eye movements.

Previous research showed that viewing pleasant or unpleasant images causes higher pupil dilation. In our work, unpleasant HH caused the largest pupil dilation thus supporting the previous statement. On the other hand, neutral themed CF and RC which causes cybersickness symptoms did not affect pupil dilation significantly. Unpleasant emotions also caused participants to fixate less and their saccade rates were significantly higher. They can be used as indicator for unpleasant emotions.

Table 2.1: Differences in physiological measurements and questionnaire outcomes between the VEs. Both groups are in descending significance order.

	Campfire ($M \pm SD$)	Hospital ($M \pm SD$)	Roller Coaster ($M \pm SD$)	<i>Significance</i>
Saccade Mean Speed (pixel/millisecond)	0.63 \pm 0.31	0.83 \pm 0.46	2.26 \pm 0.95	$F_{1,24,23.91} = 37.21$; $p < 0.001$
Pupil Diameter Change (mm)	0.45 \pm 0.40	0.77 \pm 0.56	0.57 \pm 0.48	$F_{2,38} = 13.32$; $p < 0.001$
Saccade Rate (saccade/minute)	0.07 \pm 0.05	0.10 \pm 0.04	0.06 \pm 0.04	$F_{2,38} = 7.35$; $p < 0.01$
Heart Rate (bpm)	77.52 \pm 14.55	76.28 \pm 15.58	81.01 \pm 15.78	$F_{2,38} = 6.541$; $p < 0.01$
Fixation Count	24.05 \pm 11.30	16.40 \pm 7.61	25.35 \pm 11.11	$F_{2,38} = 5.23$; $p < 0.05$
Fixation Duration (millisecond)	1779.69 \pm 629.20	1599.92 \pm 237.3	1924.79 \pm 682.50	$F_{2,38} = 5.04$; $p < 0.05$
Blink Count	22.90 \pm 29.98	15.90 \pm 17.90	17.65 \pm 17.17	$F_{2,38} = 1.92$; $p = 0.16$
Oculomotor Difference	10.99 \pm 14.07	12.13 \pm 18.63	28.43 \pm 32.15	$F_{2,38} = 15.16$; $p < 0.001$
Total SSQ Difference	11.97 \pm 18.42	17.58 \pm 30.87	41.70 \pm 38.47	$F_{1,62,30.79} = 11.51$; $p < 0.001$
Nausea Difference	7.16 \pm 14.11	13.83 \pm 25.86	31.01 \pm 29.38	$F_{2,38} = 9.42$; $p < 0.001$
Disorientation Difference	13.92 \pm 31.44	22.97 \pm 47.06	57.77 \pm 54.14	$F_{1,26,23.96} = 5.57$; $p < 0.05$

Aside from emotions, we found out saccade mean speed and heart rate were significantly higher in RC where participants felt cybersickness symptoms. These metrics can be used as indicator of cybersickness.

Effects of Scene Context on Cybersickness and Emotions. We observed highest average change in total SSQ (Fig. 2.5.g), nausea, oculomotor and disorientation scores (Table. 2.2) after RC. All VEs decreased comfort feeling (Fig. 2.5.j, Table. 2.2) where possibly due to unpleasant emotional setting and cybersickness causing elements, this increase was greater after HH and RC. Another possible effect of unpleasant emotions and cybersickness can be seen on average anxiety where HH and

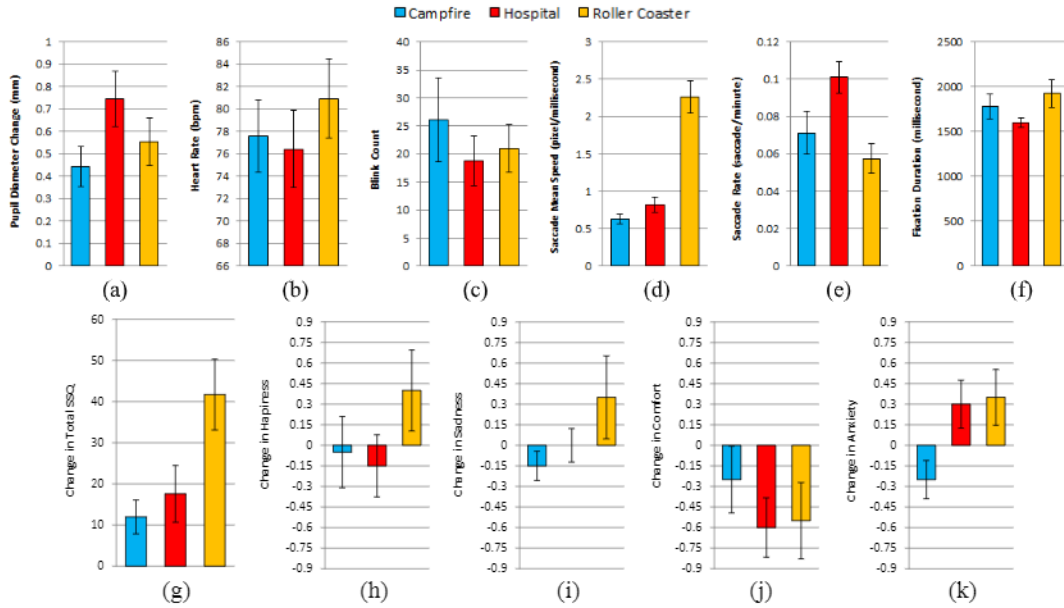


Figure 2.5: Color bars indicate means for physiological response measures and questionnaire outcomes in the three VEs: (a) Pupil dilation (b) Heart rate (c) Blink count (d) Saccade mean speed (e) Saccade rate (f) Fixation duration (g) Change in total SSQ score (h) Change in happiness feeling (i) Change in sadness feeling (j) Change in comfort feeling (k) Change in anxiety feeling. The error bars represent the corresponding ± 1 standard error of the mean.

RC increased it while neutral CF reduced it (Fig. 2.5.k, Table. 2.2). Average sadness decreased after CF while cybersickness symptoms increased it (Fig. 2.5.i). Unexpectedly, unpleasant HH did not affect sadness levels of participants. Another interesting finding was that CF and HH decreased average happiness, on the other hand with a large standard deviation from the sample mean, RC increased it (Fig. 2.5.h, Table. 2.2). Despite increasing happiness, RC also increased sadness the most with a large standard deviation (Fig. 2.5.i, Table. 2.2).

Significant Effects of Scene Context. To investigate the effects of VEs with distinct scene context on physiological responses in detail, we ran one-way repeated measure ANOVAs on the physiological data we collected from the participants while they were experiencing all 3 VEs. We found statistically significant effect of scene context on pupil dilation ($p < 0.001$), saccade mean speed ($p < 0.001$), saccade rate ($p < 0.01$), heart rate ($p < 0.01$), fixation count ($p < 0.05$) and fixation duration ($p < 0.05$)

Table 2.2: SSQ scores and VAS ratings before and after the VEs.

SSQ					
		Nausea	Oculomotor	Disorientation	Total SSQ
Campfire	Before	8.10 ± 11.79	10.99 ± 17.53	12.53 ± 25.63	11.97 ± 17.16
	After	15.26 ± 15.79	21.98 ± 18.24	26.45 ± 38.85	23.94 ± 21.98
Hospital	Before	7.16 ± 11.24	12.89 ± 14.80	12.53 ± 19.64	12.53 ± 14.11
	After	20.99 ± 29.80	25.01 ± 23.62	35.50 ± 54.18	30.11 ± 35.17
Roller	Before	7.63 ± 10.71	12.51 ± 12.75	11.83 ± 15.42	12.34 ± 10.85
Coaster	After	38.64 ± 37.01	40.93 ± 37.50	69.60 ± 63.64	54.04 ± 44.48

VAS					
		Happiness	Sadness	Comfort	Anxiety
Campfire	Before	4.50 ± 1.43	1.65 ± 1.11	5.00 ± 1.52	1.85 ± 1.35
	After	4.45 ± 1.32	1.50 ± 0.92	4.75 ± 1.41	1.60 ± 1.07
Hospital	Before	4.65 ± 1.35	1.65 ± 1.15	5.10 ± 1.48	1.65 ± 1.53
	After	4.50 ± 1.63	1.65 ± 1.15	4.50 ± 1.63	1.95 ± 1.50
Roller	Before	4.40 ± 1.07	1.55 ± 0.92	4.95 ± 1.32	1.65 ± 1.19
Coaster	After	4.80 ± 1.63	1.90 ± 1.61	4.40 ± 1.83	2.00 ± 1.73

(Table 2.1). To further analyze the data, we applied post hoc tests with Bonferroni correction and they revealed the following:

- Pupil dilation was significantly the highest in unpleasant HH scene ($p < 0.001$ between HH and CF and $p < 0.01$ between HH and RC).
- Saccade mean speed was found to be significantly higher in RC than other VEs after the Greenhouse-Geisser correction was applied to eliminate sphericity violation ($p < 0.001$).
- HH caused significantly higher saccade rate compared to the other VEs ($p = 0.05$ between HH and CF and $p = 0.001$ between HH and RC).
- Heart rate was significantly different between HH and RC ($p < 0.05$) while it was almost significantly different between CF and RC. However it wasn't significantly different between CF and HH.
- Fixation count was significantly the highest in HH ($p < 0.05$ between HH and both CF and RC).

- There was a significant difference between fixation durations measured during HH and RC ($p < 0.05$) while the difference between CF versus HH or RC were not significant.

Furthermore, in order to analyse VEs effect on emotions and cybersickness, we ran one-way repeated measure ANOVAs on the score changes of questionnaires filled before and after each VE session, SSQ and VAS. Our analyses revealed that the scene context had significant effects on changes in nausea ($p < 0.001$), oculomotor ($p < 0.001$) and disorientation ($p < 0.05$) scores of SSQ as well as the total SSQ score ($p < 0.001$) (Table 2.1). Post hoc tests with Bonferroni correction revealed:

- RC caused the significantly highest nausea score change ($p < 0.01$ between RC and CF and $p < 0.05$ between RC and HH).
- After Greenhouse-Geisser correction was applied to eliminate sphericity violation, SSQ oculomotor score changes were almost significantly different between HH and RC ($p = 0.055$). There were not any significant differences between CF-HH and CF-RC in terms of SSQ oculomotor score change.
- RC caused the significantly highest disorientation score change ($p < 0.001$ between RC and CF and $p < 0.01$ between RC and HH).
- After Huynh-Feldt correction was applied to eliminate sphericity violation, RC was found to have caused the significantly highest total SSQ score change ($p < 0.01$).

Through our analysis, we found that scene context does not have a significant effect on changes between pre- and post-session in VAS questionnaire responses, that is, in feelings of happiness, sadness, anxiety or comfort.

Correlations Between Objective Measurements and Subjective Outcomes. With the purpose of exploring the relationship between physiological responses measurements and the questionnaire outcomes, we used Pearson's correlation with linear regression. Correlation results for all VEs are given in Table 2.3. We did not find a significant correlation which is present in all three VEs. The correlation between

heart rate and change in total SSQ score was positively correlated across all VEs however it was only significant for HH ($p < 0.05$). Similarly, there was a positive correlation between blink count and change in anxiety feeling in all VEs but it was significantly so for HH ($p < 0.05$). Another overall positive correlation was found between fixation count and change in comfort feeling which was significantly so only for RC ($p < 0.05$). Saccade rate was found to have positive correlations with total SSQ scores and changes in disorientation feeling in general, where this relationship is only significant in RC ($p < 0.05$). Another overall positive correlation emerged between blink count and changes in both nausea and oculomotor discomfort scores, particularly in CF and HH VEs where these correlations were significant ($p < 0.05$ and $p < 0.01$, respectively). We also found a significant negative correlation between fixation duration and change in sadness in CF ($p < 0.001$) and in HH ($p < 0.01$).

Correlation Between Cybersickness and Emotions. We found an interesting relationship between SSQ and VAS outcomes after using Pearson’s correlation with linear regression. For all 3 VEs, there was a significantly negative correlation between change in total SSQ score and change in happiness ($r = -0.55$, $p < 0.05$ for CF; $r = -0.45$, $p < 0.05$ for HH; $r = -0.49$, $p < 0.05$ for RC).

Relationship Between Motion Sickness Sensitivity and Cybersickness. In order to analyse data collected by questionnaires based on motion sickness sensitivity, we divided the participants into two groups depending on their MSSQ total scores. The mean MSSQ score of healthy people without any type of vestibulopathy ($n = 167$) was revealed to be at 36.4 percentile in a previous work by Paillard et al. [33], which is converted to 7.9 MSSQ total score. Using this score as a threshold, 14 of the 20 participants in our study established the group of people with more than average motion sickness sensitivity. For comparison of these two groups, we ran independent samples t test on total SSQ score changes for each VE. Our analysis did not reveal a significant difference between two groups on changes in total SSQ scores, however there was an almost significant difference ($p = 0.071$) on RC outcomes.

For exploring the relationship between VR sickness and motion sickness sensitivity in detail, we investigated the correlation between SSQ and MSSQ score changes for all VEs. In our analysis, we found a positive correlation between change in total SSQ

Table 2.3: Correlations between physiological measures and questionnaire scores for each VE.

Campfire								
	Change in Nausea	Change in Oculomotor	Change in Disorientation	Change in Total SSQ	Change in Happiness	Change in Sadness	Change in Comfort	Change in Anxiety
Pupil Change	0.07	0.25	0.14	0.18	-0.31	0.04	-0.52*	0.02
Heart Rate	0.35	0.42	0.34	0.42	0.12	-0.20	-0.13	-0.17
Fixation Duration	-0.50*	-0.23	-0.21	-0.33	0.40	-0.77***	0.15	-0.60**
Fixation Count	0.22	-0.05	-0.05	0.03	0.06	0.34	0.25	0.05
Blink Count	0.47*	0.53*	0.11	0.40	-0.10	0.23	-0.21	0.34
Saccade Rate	-0.01	0.09	0.15	0.10	-0.34	0.00	-0.25	0.05
Saccade Mean Speed	0.12	-0.01	0.22	0.14	-0.14	0.03	-0.13	-0.16

Hospital								
	Change in Nausea	Change in Oculomotor	Change in Disorientation	Change in Total SSQ	Change in Happiness	Change in Sadness	Change in Comfort	Change in Anxiety
Pupil Change	0.05	0.12	0.13	0.11	0.36	-0.20	0.20	0.08
Heart Rate	0.49*	0.53*	0.42	0.49*	-0.04	0.01	-0.06	0.16
Fixation Duration	0.37	0.29	0.54*	0.43	-0.27	-0.64**	-0.08	-0.36
Fixation Count	-0.02	-0.08	-0.16	-0.01	0.31	0.04	0.28	-0.29
Blink Count	0.61**	0.66**	0.38	0.55**	0.06	0.40	-0.30	0.46*
Saccade Rate	0.23	0.26	0.27	0.27	-0.30	-0.28	-0.32	0.09
Saccade Mean Speed	0.01	0.06	0.04	0.04	-0.07	-0.29	0.12	-0.29

Roller Coaster								
	Change in Nausea	Change in Oculomotor	Change in Disorientation	Change in Total SSQ	Change in Happiness	Change in Sadness	Change in Comfort	Change in Anxiety
Pupil Change	-0.03	0.27	0.00	0.10	0.10	-0.10	0.09	-0.23
Heart Rate	0.18	0.27	0.22	0.25	-0.10	0.22	0.01	0.13
Fixation Duration	0.11	0.11	0.36	0.21	-0.43	0.00	-0.54*	0.21
Fixation Count	-0.07	-0.11	0.04	-0.05	0.24	0.02	0.53*	-0.38
Blink Count	0.04	0.22	-0.18	0.04	0.20	0.17	0.29	0.27
Saccade Rate	0.63**	0.44	0.48*	0.55*	-0.31	0.55*	0.10	-0.15
Saccade Mean Speed	0.03	-0.05	0.12	0.04	0.23	0.11	0.14	0.12

*Correlation is significant at 0.05 level. **Correlation is significant at 0.01 level. ***Correlation is significant at 0.001 level.

and MSSQ scores for RC ($p < 0.01$) which revealed that participants who are more susceptible to motion sickness showed more cybersickness symptoms.

2.5 Threats to Validity

In our user study, we showed the VEs in a completely random order to the participants in order to minimize the threats to internal validity. For the purpose of minimizing the threats to external validity, we avoided using abstract VEs in order to achieve an

experience which is close to the real world VR scenarios. We also used a widely available commercial HMD in our study. However, disabled HMD head tracking and predefined paths in VEs for achieving a controlled experiment decreased the external validity due to limiting HMD's features.

2.6 Discussion

In our analysis, we found that context of the scene has statistically significant effect on pupil dilation, heart rate, fixation duration, fixation count, saccade mean speed and saccade rate. Among these measurements, pupil dilation and saccade mean speed were effected by scene context the most significantly. We did not observe a significant effect of scene context on measured happiness, sadness, comfort and anxiety feelings while we found that it was significantly effective on nausea, disorientation and oculomotor discomfort in addition to total feeling of cybersickness.

The uneventful and relaxing environment of CF decreased the anxiety feeling of the participants, on the other hand unpleasant elements in HH and cybersickness causing roller-coaster ride in RC scene increased it. Significantly the highest pupil dilation and saccade rate was seen on HH, caused by unpleasant elements and sounds in the VE. A notable result was compared to RC, HH decreased the comfort feeling more which indicates the important effect of unpleasant sensory feedback and how it can be as effective as cybersickness symptoms in context of affecting and reducing the general comfort levels of a person.

Our results showed that cybersickness symptoms were clearly felt by participants the most while experiencing RC. Interestingly, participants also felt cybersickness while traversing CF and HH as seen from change in SSQ scores. These result revealed that even experiencing a VE which is specifically designed not to cause cybersickness and be uneventful can cause cybersickness, especially in people who are more susceptible to it while using a highly immersive setup.

Another interesting finding revealed in our analysis was that change in happiness and cybersickness had significant negative correlation between them. On the other hand, despite being the VE that caused the greatest increase in cybersickness, RC

increased happiness while other VEs decreased it. We examined this behaviour and found that since participants who are less susceptible to motion sickness had not felt cybersickness as much as the ones who are susceptible to it, had enjoyed experiencing RC and felt happier.

Finally, even though our findings were obtained after extensive analysis, in order to better understand the interaction of such complex network of inputs and outputs, further studies with larger sample sizes are necessary.

2.6.1 Related Publications

Cebeci, B., Celikcan, U., & Capin, T. K. (2019). A comprehensive study of the affective and physiological responses induced by dynamic virtual reality environments. *Computer Animation and Virtual Worlds*, 30(3-4), e1893.

CHAPTER 3

OPTIMIZING STEREOSCOPIC PARAMETERS IN REAL TIME USING A GAZE-DIRECTED METHOD

Providing visually realistic and comfortable experience is still a challenge in VR applications which prevents them to reach their full potential. Poorly optimized stereo parameter settings will result in eye fatigue and discomfort while viewing 3D stereoscopic images, thus further causes accommodation-convergence conflict and large disparities. Additionally, experiencing interactive stereoscopic virtual environments (VEs) for extensive periods can further cause eye strain, visual fatigue, headaches and inability to fuse two images (diplopia) [14].

Based on our eyes' positions, we see the world in two slightly different views which creates binocular vision. Stereoscopic parameters such as convergence distance and interaxial separation produce these different views, also called screen disparities, which results in perceived depth. Finding a comfortable range of this perceived depth is a challenge while creating stereoscopic content. Knowing the region which has viewer's attention and modifying stereoscopic parameters according to that is known for increasing the general viewing comfort [14]. For this reason, a method is needed to detect the region of the viewer's attention and modify the stereoscopic parameters based on the objects in that region in order to achieve highest viewing comfort possible.

A number of studies modified stereoscopic parameters, thus perceived depth of the scene, in post production of 3D videos [34, 35], on the other hand, other studies performed this parameter optimization in real time since it is more applicable to interactive stereoscopic scenes. However, keeping the perceived depth in a comfortable range in VEs where viewer's are free to move around is challenging. To overcome

this, Oskam et al. [15] and Celikcan et al. [16] optimized stereoscopic parameters in real time based on scene elements. However, these methods estimate the region of viewer's attention while optimizing stereoscopic parameters. In order to avoid estimation, eye tracking can be used for more accurate optimization. Earlier works showed that tracking and using viewer's gaze movements in optimization can reduce fusion time thus increase the scene comfort [17]. Also, it has been demonstrated that viewer's gaze can improve comfort of binocular viewing while preserving original depth of the scene [18].

Although these stereoscopic parameter optimization methods showed promising results on increasing viewing experience on traditional monitor-based 3D displays, there are not many works conducted for immersive VR displays regarding this topic. Widely available commercial HMDs generally have fixed stereoscopic camera parameters regardless of the scene content, which reduces the perceived depth and general visual comfort.

In this part of the thesis, our goal is to investigate the effects of dynamically modifying stereoscopic camera parameters on the visual comfort and depth perception in immersive stereoscopic VEs using an HMD. To achieve our goal, we developed a stereoscopic camera control method based on eye tracking. We conducted a user study to evaluate the quality of our method based on participants' subjective self assessments. For the user study, we created an indoor and an outdoor interactive VEs and compared our method's results with the a state-of-art commercial HMD, HTC Vive's default settings.

3.1 Background

3.1.1 Stereo Geometry

Depth cues are very important for human visual system to use objects' spatial relationships to perceive them. *Motion cues* such as motion parallax and kinetic depth, *oculomotor cues* based on eye movements, *pictorial cues* such as relative height and size, *binocular cues* are the visual cues used for depth perception. Especially binoc-

ular depth cues have a critical importance in stereo geometry. Our left and right eyes receive different views due to the position differences between them, which causes different retinal images, thus binocular vision is created as a result. Human visual system uses two parameters in real world for creating binocular vision; *binocular disparity*, the difference between two retinal images, *convergence* which is an eye movement based on the rotation of the eyes that enables our eyes to fixate at a certain point.

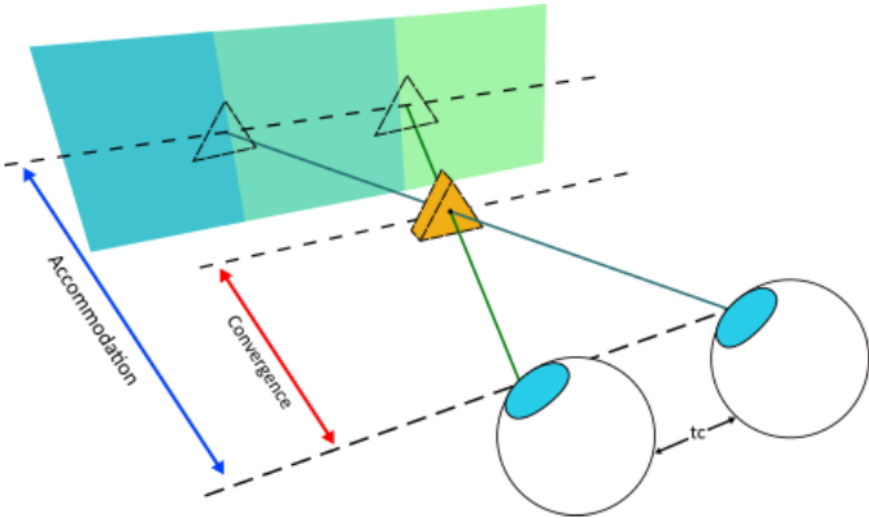


Figure 3.1: Accommodation and convergence conflict

Similar to real world parameters, there are two major parameters while creating a stereoscopic image: *interaxial separation* (t_c) and *convergence distance* (Z_c). As seen in Figure 3.2a, interaxial separation corresponds to the distance between two cameras. Convergence distance is the distance between the cameras and the focus plane (zero parallax plane) where both cameras converge. Interaxial separation (t_c) affects the image disparity directly thus it will change the perceived depth amount of the final created image. It has a positive correlation with perceived depth and disparity. Compared to t_c , convergence distance does not affect the perceived depth, whereas increasing it shifts the perceived depth forward and decreasing it shifts the perceived depth backward.

The disparity of an object which is positioned at scene distance (Z) is given as:

$$d = f_c t_c \left(\frac{1}{Z_c} - \frac{1}{Z} \right) \quad (31)$$

which shows that image disparity depends on interaxial separation (t_c), convergence distance (Z_c) and focal length of the cameras (f_c).

There are three conditions based on an object's position (Z) in the scene. The first one is called zero parallax setting, the object is located on the focus plane where two cameras converge ($Z = Z_c$). In this condition, retinal position of the object also appear at that distance which will cause it to appear at the surface of the physical screen. In $Z > Z_c$ condition, object has a positive disparity and appears behind the screen where in $Z < Z_c$ condition, object has a negative disparity and appears in front of the screen.

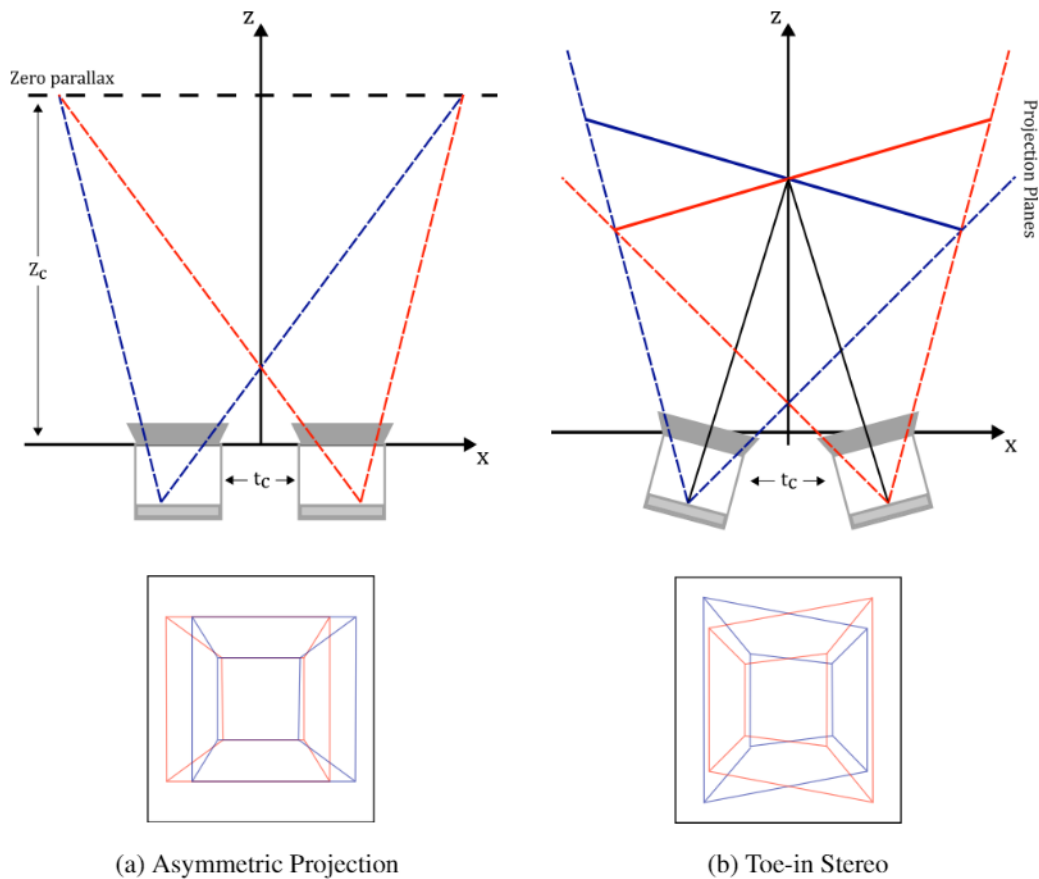


Figure 3.2: Different stereoscopic scene rendering methods

3.1.2 Stereoscopic Scene Rendering Method

Asymmetric projection [36] is a widely used stereoscopic scene rendering method for HMDs. In this method, cameras have an asymmetric frustum after the projection matrix is modified as seen in Figure 3.2a. This projection matrix modification process extends one side of the view frustum while condensing the other side.

This method is preferred over another method called toe-in stereo where virtual cameras rotate based on convergence distance (Figure 3.2b). This rotation causes distortion in rendered images [36], therefore commercial HMDs such as HTC Vive and Oculus Rift uses asymmetric projection.

3.1.3 Accommodation and Convergence Conflict

In the real world while looking at an object, our eyes accommodate and converge at the same point where the object is positioned. On the other hand, while viewing a stereoscopic image displayed on the screen, the viewer's eyes accommodate on the display screen, however their eyes converge on depth of the object they are looking at as seen in Figure 3.1. Since this condition does not occur in real world, it creates a conflict between accommodation and convergence of the eyes. Due to this unnatural experience, general comfort while viewing stereoscopic images on a 3D display or HMD greatly decreases.

3.2 Related Work On Improving Stereo Viewing Comfort

In order to increase the general comfort of a stereoscopic scene, many studies modified stereoscopic parameters in different methods. One of the earliest conducted studies about this topic is made by Ware et al. [37], where they implemented a stereo parameter modification algorithm which works in real time and reduces discomfort. Similarly, Kulshreshth et al. [38] proposed a dynamic stereo parameter adjustment algorithm but they used the depth map of the VE to find optimum stereo parameters. They found that their dynamic stereo parameter adjustment algorithm was preferred over the fixed stereo parameters by the subjects. Celikcan et al. [16] optimized stereo

parameters based on importance of the scene objects and the distance between them and virtual cameras, and improved scene depth considering viewer's disparity range. Their results showed that their method improved the general quality of the stereoscopic viewing experience. Oskam et al. [15] modified virtual cameras' convergence distance and interaxial distance dynamically based on virtual objects and viewer's position in the VE and they increased scene comfort without decreasing perceived realism. Aside from these works which used a desktop 3D display in their studies, Sherstyuk et al. [39] used an HMD and optimized the convergence distance based on the virtual hand model's position. They stated that their dynamic eye convergence model increased the VR quality. Although these studies improved the quality of stereoscopic scenes, they based their methods on gaze estimation which is not as accurate as using the gaze points directly.

Therefore, instead of using viewer's gaze based on estimation, detecting and tracking their gaze points with an eye tracker will provide more accurate and reliable results while developing stereoscopic scene parameters optimization methods. Thus, in order to enhance stereoscopic viewing experience, many studies used eye tracker in their optimization methods.

Foveated rendering, a rendering technique which uses an eye tracker to acquire the gaze center and reduces the quality of the image in the outside area of the gaze point, is widely used in both screen-based and VR stereoscopic displays. Arabadzhiyska et al. [40] improved the quality of foveated rendering by estimating ending position of a saccadic eye movement in screen based stereoscopic displays. Pai et al. [41] and Romero et al. [42] used foveated rendering in VR however they did not validate their studies with a user study. Aside from foveated rendering, Conti et al. [43] used the viewer's gaze to implement a dynamic depth of field (DOF) blur effect in a CAVE Automatic Virtual Environment, however participants did not preferred DOF blur. Kellnhofer et al. [44] improved the depth of stereoscopic images using eye fixation points in stereoscopic 3D displays.

The effect of dynamic adjustment of stereo parameters based on gaze feedback is investigated in a number of studies using stereoscopic 3D displays. Gurrieri [18] used viewer's gaze points to apply disparity correction in order to increase comfort and

decrease accommodation and convergence conflict. However, they manually generated disparity maps in their method which is not applicable for real time applications such as interactive 3D games. Bernhard et al. [17] avoided manual adjustment steps and used a gaze-directed stereoscopic parameter optimization method in an abstract scene which works in real time and found that their method increased comfort and scene depth while lowering fusion times. Aside from their stereo parameter adjustment algorithm based on depth maps, Kulshreshth et al. [38] also developed a stereo parameter modification model using viewer's gaze points. They noted that their model performed better while virtual objects are close to the camera.

Unlike traditional stereoscopic 3D displays, not many studies are conducted regarding this topic in VR. Kouliris et al. [45] implemented a gaze predictor method using machine learning. For collecting data to train their classifier, they used an eye tracker integrated in an HMD. Their method predicted viewers' gaze points during a video game viewing session and modified stereoscopic parameters dynamically. They stated that their method can make accurate gaze point predictions and improves general scene quality. Despite using eye tracker only for classifier training purpose and not in real time optimization, their method showed improvement over other methods they have compared with in their study.

Konrad et al. [46] dynamically changed ocular parallax stereo parameter using viewers' gaze points captured with an eye tracker integrated HMD. They found that modifying ocular parallax also improved the realistic depth feeling in VR environments. Jacobs et al. [47] dynamically modified vergence angle and interaxial distance of virtual cameras based on the depth of viewers' gazes in VR. They also implemented a probability model to improve tracked eye movements' accuracy. Yet their subjective assessments and objective measurements revealed that their method unexpectedly decreased visual comfort of the VE which shows the challenges of stereoscopic camera parameter modification without causing discomfort.

Although there are studies in the literature which modify stereoscopic parameters in order to improve general quality of VR using eye tracker systems, our gaze-directed stereoscopic camera control method is an improvement over for those mentioned methods, mainly for not requiring any pre-processing such as training a machine

learning model and can work with any 3D scene in real time. Unlike our study, similar works [45] [47] used limited to non interactable VEs where freely moving around the VE is not available.

3.3 Approach

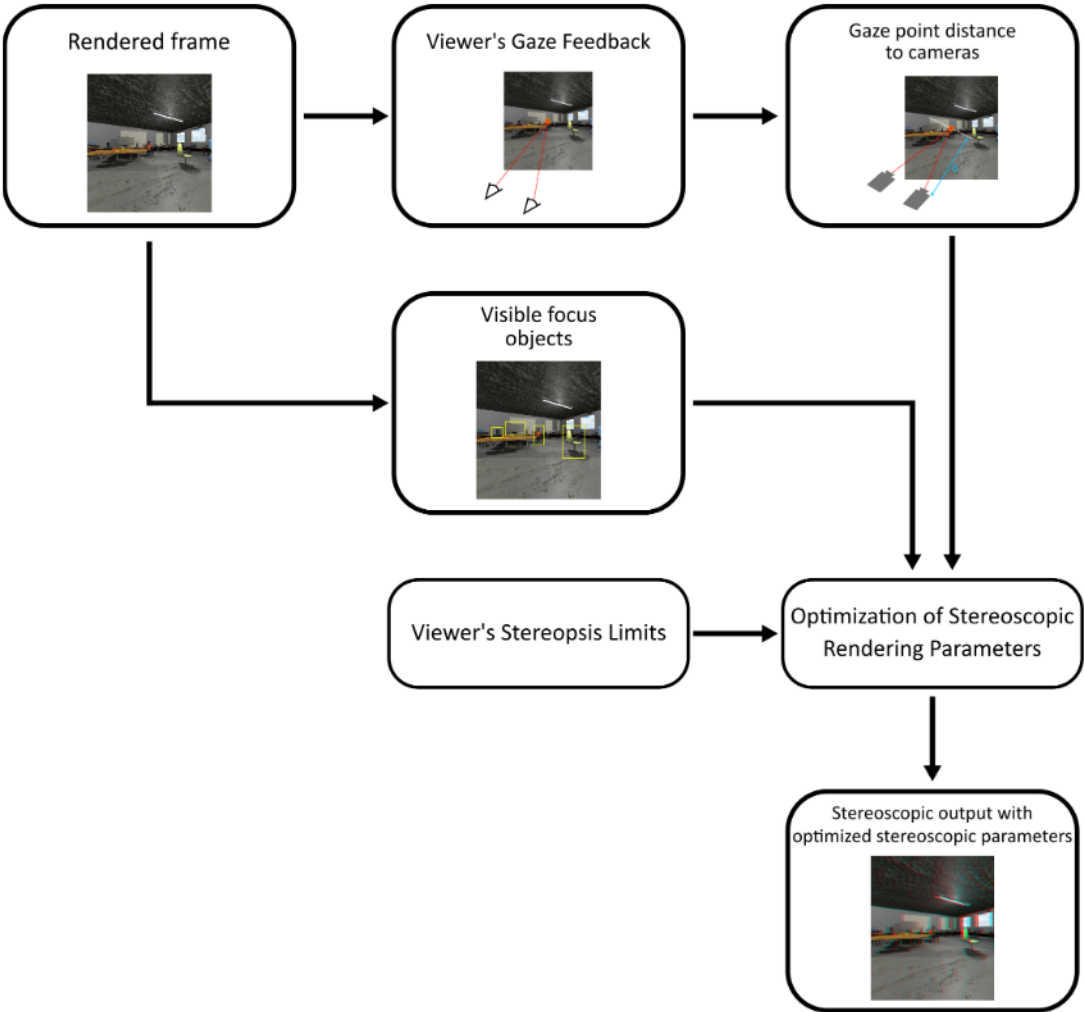


Figure 3.3: Flow diagram of gaze-directed method

3.3.1 Gaze-Directed Stereoscopic Camera Control

As mentioned in the previous chapters, commercial HMD devices such as HTC Vive have fixed stereoscopic camera parameters which results in poor depth perception and general comfort. For the improvement of these concepts, we propose a gaze-directed

method which uses a built-in eye tracker in HTC Vive to modify the stereoscopic parameters in real time. Our method's flow can be seen in Figure 3.3.

The main idea of our method is based on changing the scene objects' depth by positioning them in a region called viewer's *zone of comfort*, in order to minimize the visual discomfort in stereoscopic displays [48]. This process is also called *stereo-grading* [45]. Without this stereo-grading process, viewer's eyes accommodate and converge on separate distances which causes blurred or double vision, further results in eyestrain, headache and fatigue. Therefore, to avoid any unwanted effects, we performed stereo grading by modifying the projections of the objects in VEs.

3.3.2 Stereopsis Limits Calculation

Calculating viewer's stereopsis limits, i.e., comfortable depth perception range is crucial to avoid visual artifacts causing eye strains on too high disparity values or low depth feeling on too low disparity values.

For this purpose, a scene for stereopsis limits calculation which includes a realistic human model, is created. At the beginning, the human model positioned close to the virtual cameras as seen in Figure 3.4a. The viewers adjusted the disparity to find a point which makes the stereoscopic scene not comfortable to view. We saved this personalized disparity value as minimum disparity (d_{min}). Following that, the human model positioned further to the virtual cameras as seen in Figure 3.4b and the viewer again adjusted the disparity similar to last step. This disparity value saved as maximum disparity (d_{max}). This calculation phase is performed at the beginning of the experiment for each viewer and their unique (d_{min}) and (d_{max}) values used in the stereoscopic parameters optimization.

3.3.3 Camera Convergence Distance Modification

Camera convergence distance of the virtual cameras positioned as close as possible to the area in 3D environment which has the attention of the viewer in order to minimize the comfort reducing effects. As mentioned in Section 3.2, several works used gaze



Figure 3.4: Stereopsis limits calculation steps

approximation of the viewer while they are experiencing VE in order to achieve this goal. We used an eye tracker to obtain precise and accurate gaze points of the viewers to eliminate the need for approximating gaze points.

To find out which virtual object has the attention of the viewer, we cast rays from the virtual cameras to the gaze point at that moment. After calculating the distance between the object which has viewer’s attention and virtual cameras, we set the camera convergence distance based on the newly calculated distance each frame. A summary of the camera convergence distance modification algorithm can be seen in Algorithm 1.

Algorithm 1 Camera convergence distance setting algorithm

- 1: $obj \leftarrow getTrackedObject()$
 - 2: \triangleright Finding the object viewer is looking at
 - 3: $dist \leftarrow calculateDistance(obj)$
 - 4: \triangleright Calculate the distance between virtual cameras and the object viewer is looking at
 - 5: $SetConvergenceDistance(dist)$
 - 6: \triangleright Set convergence distance according to the new distance calculated.
-

However, this approach may result in unwanted effects considering frame-to-frame calculations. Viewers can quickly shift their attention from a far object to closer

object and vice versa. Changing camera convergence distance immediately based on these movements will decrease the viewing comfort. As a solution, we set the camera convergence distance gradually based on the value from previous frame:

$$f(x(t)) = \begin{cases} x(t-1) + x_1, & \text{if } x(t) - x(t-1) \leq x_1; \\ x(t-1) + x_2, & \text{if } x(t) - x(t-1) \geq x_2; \\ x(t-1) + k(x(t) - x(t-1)), & \text{otherwise.} \end{cases}$$

where $x_1 \in \mathbf{R}^-$, $x_2 \in \mathbf{R}^+$ and k is chosen to be $0 < k < 1$.

3.3.4 Optimizing the Interaxial Camera Distance

For the interaxial camera distance optimization, we determined and labeled focus objects, which are potential objects that will have viewer's attention in the VE. We labelled focus objects offline for both VEs. Each frame, visible focus objects to the virtual cameras are detected and their distances to the cameras are calculated. Objects far from the virtual cameras more than the maximum allowed distance are not included in calculations.

In order to use focus objects for optimizing interaxial distance value each frame, we formulated an energy function $E(t_c)$ for the purpose of maximizing total scene disparity, thus increasing general depth feeling of the stereoscopic scene. $E(t_c)$ is formulated as

$$E(t_c) = \sum_{i=1}^n f_c t_c \left(\frac{1}{Z_c} - \frac{1}{Z_i} \right) \quad (32)$$

where Z_i is forward distance of i^{th} focus object from the virtual cameras and n is the number of focus objects found for current frame. Our algorithm for finding focus objects and their forward distances from cameras is outlined in Algorithm 2.

Algorithm 2 Focus object distance calculation algorithm

```
1:  $e[] \leftarrow getFocusObjects()$ 
2:  $\triangleright$  Get all focus objects in the current scene
3:  $j \leftarrow 0$ 
4: for  $\forall e[i]$  do
5:   if  $e[i]$  is visible by camera then
6:      $e[i].Z \leftarrow ForwardDistanceFromCamera()$ 
7:     if  $e[i].Z \leq D_{max}$  then
8:        $\triangleright D_{max}$  : maximum allowed forward distance
9:        $o[j] \leftarrow e[i]$ 
10:       $\triangleright$  implies  $o[j].Z \leftarrow e[i].Z$ 
11:       $j \leftarrow j + 1$ 
12:     end if
13:   end if
14: end for
15: return  $o[]$ 
```

Our model maximizes $E(t_c)$ based on following constraints:

$$d_{max} \geq fct_c \left(\frac{1}{Z_c} - \frac{1}{Z_i} \right) \geq d_{min}, \forall i | 1 \leq i \leq n, \quad (33)$$

where d_{max} and d_{min} obtained from disparity calibration. Using these constraints, we kept the scene depth in a comfortable perceivable depth range while viewers experienced our VEs.

We used improved stochastic ranking based evolutionary strategy (ISRES) algorithm [49] for the optimization of our energy function. In constrained optimization, the ISRES algorithm is a well known solution which is also available in NLOpt library [50]. In order to optimize our energy function in real time, we used multi-threading to run the algorithm while viewer is traversing our VEs.

3.3.5 Projection Matrix Modification

For adjusting stereoscopic view of the HMD based on our newly calculated convergence distance and interaxial separation values, we used asymmetric projection matrix modification method as described in Avan et al.'s [51] work. They swapped the left and right eye projection matrices in order to make the modification method suitable for HMDs.

The projection matrix we used in our work is:

$$P = \begin{bmatrix} \frac{2n}{r-l} & 0 & \frac{r+l}{r-l} & 0 \\ 0 & \frac{2n}{t-b} & \frac{t+b}{t-b} & 0 \\ 0 & 0 & \frac{-(f+n)}{f-b} & \frac{-2fn}{f-n} \\ 0 & 0 & -1 & 0 \end{bmatrix} \quad (34)$$

where l, r, t, b, n and f corresponds to left, right, top, bottom, near and far plane positions of the view frustum. $\frac{t+b}{t-b}$ and $\frac{r+l}{r-l}$ create vertical and horizontal off-axis asymmetry [52]. Because of the built-in vertical asymmetry of the HMD lenses, only the horizontal off-axis asymmetry is modified and vertical off-axis asymmetry (t and b values) remain unaltered.

New l and r values for left frustum are calculated as

$$l = -\left(Z_c q \tan\left(\frac{\theta}{2}\right) + \frac{t_c}{2}\right) \frac{n}{Z_c} \quad r = \left(Z_c q \tan\left(\frac{\theta}{2}\right) - \frac{t_c}{2}\right) \frac{n}{Z_c} \quad (35)$$

and for the right frustum, they are calculated as

$$l = -\left(Z_c q \tan\left(\frac{\theta}{2}\right) - \frac{t_c}{2}\right) \frac{n}{Z_c} \quad r = \left(Z_c q \tan\left(\frac{\theta}{2}\right) + \frac{t_c}{2}\right) \frac{n}{Z_c} \quad (36)$$

where q is the aspect ratio and θ is the vertical FOV. Newly calculated values of l and r used in matrix P in Eq. 34.

3.4 Experimental Evaluation

3.4.1 Equipment

We developed and conducted our study on a PC which has Intel(R) Core(TM) i7-6700 CPU, 16 GB RAM and an Nvidia GeForce GTX 1050 Ti graphics card. As VR set, we used an HTC Vive head-mounted display (HMD) containing the Tobii Pro eye trackers that have 120 Hz output frequency with estimated accuracy of 0.5 degree.

3.4.2 Virtual Environments

For our study, we created two distinct VEs; the first one is a classroom environment (CR) which is an indoor environment that contains animated realistic human models (Figure 3.5a), the second one is suburban street (SS) which is an outdoor environment with animated realistic drone and robot models (Figure 3.5b). Using HTC Vive controllers, participants moved freely around in the VEs. In SS, participants further interacted with the VE using a virtual gun which tracks the controllers movements.

3.4.3 Subjects

21 people volunteered for our study however, three volunteers with mean age 27 ± 1.41 failed to participate in the study. Despite passing the Monoyer chart test, one of them had blurred vision which prevented a high quality VR experience and others failed to pass eye tracker calibration. Remaining 18 volunteers, who will be referred as participants from this point, with mean age 24.88 ± 5.32 (13 males, 5 females) completed the study with success. Participants' ages varied between 20 and 42.

Using a five-point Likert scale questionnaire (0: "no previous experience," 4: "very experienced"), participants rated their previous VR experience levels. Participants' mean experience level was 1.27 ± 1.28 and 6 of the participants experienced VR for the first time. All participants volunteered for our study and did not received payment or anything else in return.

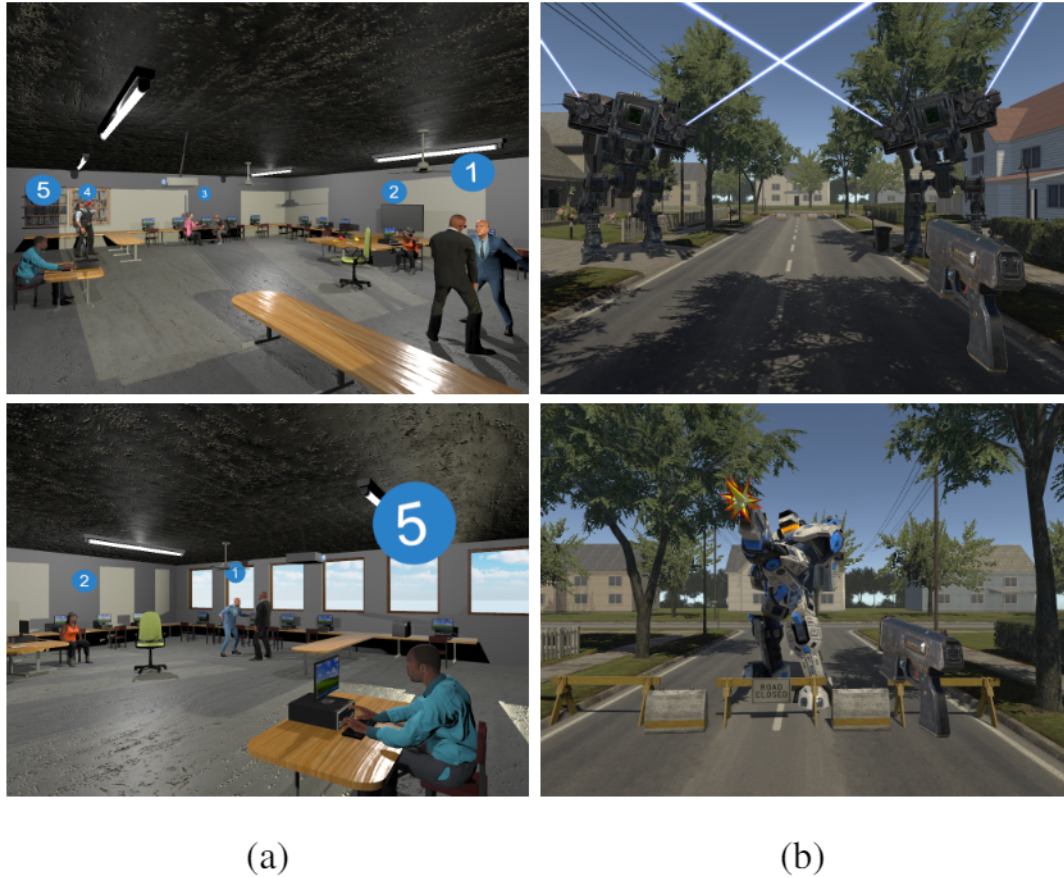


Figure 3.5: Sample frames from (a) CR and (b) SS

3.4.4 Procedure

Before participants started the experiment, first their general vision acuity and stereoscopic viewing ability were checked using Monoyer chart and random dot stereogram test respectively. All participants scored at least 9 out of 10 in Monoyer test and passed random stereogram test thus qualified for the experiment.

Following that, participants were asked to sign a consent form. After their approval, they completed a demographic form containing questions regarding their names, ages, VR experiences and university departments they graduated. After that, experiment procedure and usage of the HTC Vive controllers explained to the participants.

As a last step, each participant's interpupillary distance was measured and the HTC Vive HMD was adjusted according to measured distance value. In order to achieve

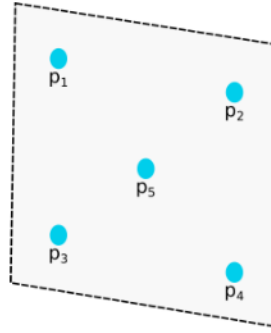


Figure 3.6: Eye tracker calibration points

accurate tracking of the eye movements, the eye tracker has to be placed on the head of the user correctly and calibrated first. When participants put on the HMD, headset is adjusted until it is in an acceptable position by the eye tracking system. Following that, eye tracker calibration phase started with displaying 5 dots on the screen one by one as seen in Figure 3.6. Participants followed these dots only with their eyes. If the calibration is successful, depth perception calibration phase started. This phase is carried on as explained in Section 3.3.2.

For comparing our method, participants experienced each VE twice in random order; with the default HTC Vive parameters where no stereoscopic camera parameter optimization is present, and with our Gaze-Directed Stereoscopic Camera Control where stereoscopic camera parameters optimized in real time. Participants were told to examine their surroundings while traversing the VEs for a better methods comparison.

Firstly, CR sessions showed to the participants. They were asked to go near to the 5 checkpoints placed across the VE in an unrestricted order using HTC Vive controllers. After 5th checkpoint is reached, session concluded. After experiencing CR twice with different methods, SS sessions started. Participants traversed a suburban street where robot and drone enemies attacked as they progressed. Participants were asked to shoot them using HTC Vive controllers. When they reached at the end of the VE, session ended. After experiencing SS with the other method, experiment concluded.

After each session, the participants removed the HMD and were asked to evaluate the visual quality, visual comfort, perceived depth and overall experience of the VE they viewed in that session on a 7-point Likert scale questionnaire where 1 indicates

the worst score and 7 indicates the best (See Appendix C for details). While some studies investigated realistic depth [46] [53], we aimed to measure depth perceived by participants and explore its possible improvements.

3.5 Results

3.5.1 Frame Rate

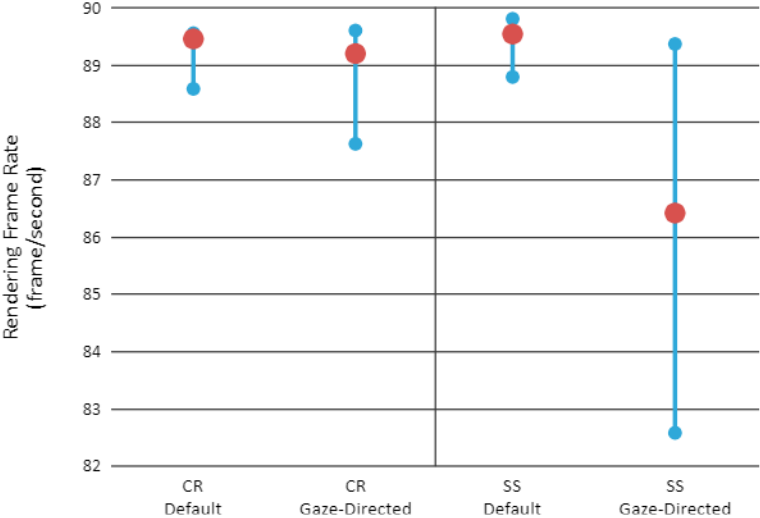


Figure 3.7: Ranges of frame rates of sessions are shown with blue bars. Red dots mark the averages.

Virtual Reality applications are required to meet averagely 90 frames per second while running in order to avoid unwanted effects on the users. Considering this, we had to optimize our VEs and camera optimization method to match this requirement while real time stereoscopic camera optimization is present. We recorded average frame rate of each session and results can be seen on Figure 3.7.

Our frame rate recordings showed that sessions with default HTC Vive configuration had slightly higher frame rates compared to sessions with our optimization method. This result was expected since there are no extra elements which require additional computational power in default HTC Vive settings unlike our method which works with additional hardware and software such as an eye tracker and libraries for the stereoscopic parameters optimization. While this difference is unnoticeable in CR

which has approximately 598 thousand polygons per frame, it is more significant in SS which has approximately 903 thousand polygons per frame. Focus objects are another factor to cause frame drops in SS for our method. On average 16 focus objects per frame were used in optimization although this number increased to 20 focus objects per frame in SS where it also peaked at 25 at the beginning part of the VE.

These results showed that our method causes frame rate drops in VEs with higher polygon count per frame and increased number of focus object. Despite the fact that frame rate drops in our method, in the end it does not effect the VR experience and our method can still catch up with the frame rate requirements of the VR applications.

3.5.2 Evaluation Scores

Average scores of participants' self assessments for four different measures (Image Quality, Perceived Depth, Visual Comfort, Overall Experience) where 1 is the lowest score and 7 is the highest, for both Default HTC Vive setup and our Gaze-Directed Stereoscopic Camera Control Method can be seen on Table 3.1a for CR and Table 3.1b for SS.

Results on Table 3.1 showed that our method overperformed HTC Vive camera setup and had higher average scores for all four measurement metrics in both VEs, where the differences between average scores of default setup and our method are greater in SS. Figure 3.8 shows the average scores in detail.

3.5.3 Statistical Analysis

For analysis we conducted two-tailed paired sample t-tests to participants' self assessment scores. Analysis showed that there was a statistically significant difference between our method and default HTC Vive setup in terms of perceived depth in CR, image quality, perceived depth, visual comfort and overall experience in SS. These results revealed the following:

- Our method's perceived depth ($M = 6.06, SD = 0.73$) provided significantly

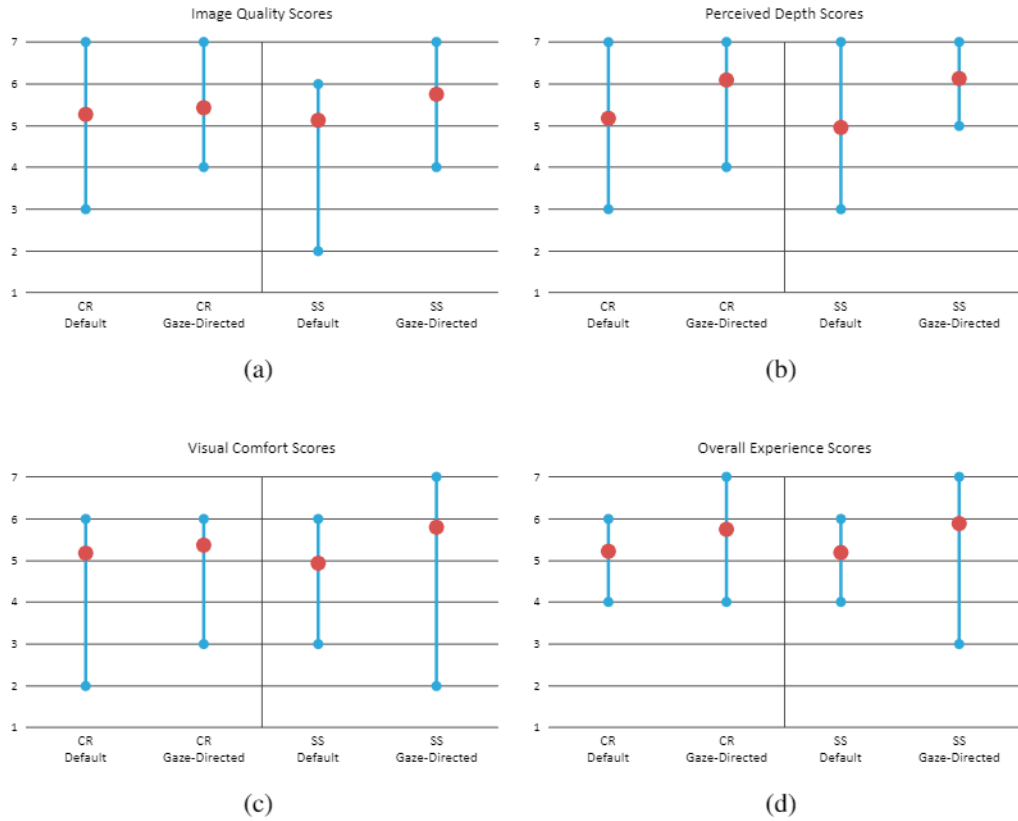


Figure 3.8: Ranges of scores given by the user study participants are shown with blue bars for (a) image quality, (b) perceived depth, (c) visual comfort, and (d) overall experience. Red dots mark the respective averages.

better depth perception than provided by default parameters ($M = 5.17, SD = 0.99$) in CR ($t = 3.332, df = 17, p = 0.004$).

- Our method's overall experience score ($M = 5.67, SD = 0.68$) was almost significantly better compared to default parameters ($M = 5.22, SD = 0.80$) sessions in CR ($t = 2.046, df = 17, p = 0.057$).
- Image quality in our method ($M = 5.72, SD = 0.67$) was significantly better than default parameters' image quality ($M = 5.11, SD = 0.90$) in SS ($t = 2.829, df = 17, p = 0.012$).
- Perceived depth of our method ($M = 6.11, SD = 0.58$) was significantly better than depth perception in default parameters sessions ($M = 4.94, SD = 0.87$) in SS ($t = 4.302, df = 17, p < 0.001$).

- Our method’s visual comfort ($M = 5.78, SD = 1.35$) was significantly higher than default parameters’ visual comfort ($M = 4.94, SD = 0.80$) in SS ($t = 2.185, df = 17, p < 0.043$).
- Our method provided significantly better overall experience ($M = 5.83, SD = 0.99$) than default parameters ($M = 5.17, SD = 0.51$) sessions in SS ($t = 2.817, df = 17, p = 0.012$).

As seen from the analysis results, our Gaze-Directed Stereoscopic Camera Control Method did not yield significant effect on image quality, visual comfort and overall experience aside from perceived depth in CR. Thus, we believe that our method can be used to improve overall depth perception for indoor scenes. On the other hand, our method clearly performed better in SS, an outside environment with wide depth range and more engaging VE interaction, providing significant improvement over default HTC Vive parameters on all four measures.

Table 3.1: Averages of participants’ ratings on 7-point Likert scale (a) for CR (b) for SS.

(a)

	Image Quality ($M \pm SD$)	Perceived Depth ($M \pm SD$)	Visual Comfort ($M \pm SD$)	Overall Experience ($M \pm SD$)
Experience with Default Settings	5.33 \pm 1.00	5.17 \pm 0.96	5.17 \pm 1.12	5.22 \pm 0.79
Experience with Gaze-Directed Camera Control	5.44 \pm 0.76	6.06 \pm 0.70*	5.39 \pm 0.83	5.67 \pm 0.67

(b)

	Image Quality ($M \pm SD$)	Perceived Depth ($M \pm SD$)	Visual Comfort ($M \pm SD$)	Overall Experience ($M \pm SD$)
Experience with Default Settings	5.11 \pm 0.87	4.94 \pm 0.85	4.94 \pm 0.78	5.17 \pm 0.50
Experience with Gaze-Directed Camera Control	5.72 \pm 0.65*	6.11 \pm 0.57*	5.78 \pm 1.31*	5.83 \pm 0.96*

3.5.4 Comparing With Saliency Guided Camera Control Method

In addition to our analysis, we also compared our Gaze-Directed Stereoscopic Camera Control method with a Saliency Guided Camera Control Method which uses Zhang et al.'s [54] visual saliency methods, in an unpublished manuscript [55] with the same participants and VEs as our study and without changing the experiment procedure. Along with the sessions where our method and default VR parameters are present, participants experienced the same VEs with saliency guided method modifying stereoscopic parameters in real time. Analysis showed that, saliency guided method showed a significant improvement over default VR parameters only in terms of perceived depth in SS ($p = 0.044$). However, our method provided significantly better perceived depth compared to saliency guided method in SS ($p = 0.021$). The results showed that aside from improvements over default VR parameters, our method also provided better VR experience compared to the saliency guided method, especially in the outdoor scene.

3.6 Threats to Validity

In our user study, while participants experienced the sessions with our method and default VR parameters in a random order, the order of the VEs did not change. To further eliminate possible threats to internal validity, instead of showing SS after CR, order of the VEs can be randomized as well. Similar to the user study conducted in Chapter 2, we used a widely available commercial HMD and created VEs which are applicable in real world VR scenarios. HMD's features were not restricted and participants moved freely in VEs similar to the real VR applications. For the purpose of minimizing threats to external validity and acquiring more generalized results, our gaze-directed stereoscopic camera control method can be tested with additional indoor and outdoor VEs with different contexts.

3.7 Discussion

In our study, we proposed a novel automated gaze-directed stereoscopic camera control method which works in real time for the purpose of increasing visual comfort and depth feeling of VEs, thus improving general quality of VR experience. For testing our method, we created two interactive VEs with different settings and we conducted a user study with 21 participants where each participant experienced both VEs twice, once with our method is present and once with default HTC Vive fixed stereoscopic parameters.

We compared our method and default HTC Vive parameters in terms of VE's image quality, perceived depth, visual comfort and overall experience based on participants' self assessments. Our results showed that our method showed an improvement in all four metrics of the VE we have measured in SS, while in CR only improvement was in perceived depth. These results suggest that our method works significantly better in VEs with more engaging interaction and wide depth range.

In conclusion, our study showed that dynamically modifying stereoscopic parameters based on gaze data in VEs is an applicable method to improve the VR experience of commercial HMDs. Besides the success of our gaze-directed method, further studies with larger sample sizes and different VEs are necessary to see the full potential of our method.

CHAPTER 4

CONCLUSION

In this thesis, we investigated the physiological and affective responses induced by experiencing VEs with different contexts in our first study. Through our comprehensive analysis of participants' objective physiological measurements and subjective self assessments, we distinguished whether a response is caused by cybersickness symptoms or the unpleasant setting of the VE, explored the interaction between emotions and cybersickness, examined the link between motion sickness susceptibility and cybersickness. The results of our user study revealed that context of the VE can cause significant physiological responses; pupil dilation, saccade rate and fixation count were significantly higher when participants were experiencing VE with unpleasant setting, saccade mean speed was the highest while cybersickness symptoms are present. These findings demonstrate how VR technology can affect our emotional state and cause physiological responses in human body.

In our second study, we introduced a real time working stereoscopic camera parameters optimization method based on viewer's gaze and scene contents in VR in order to improve VR experience by increasing perceived depth and general comfort. Results of our user study showed that dynamically modifying stereoscopic parameters based on gaze information in outdoor VE provided improved image quality, perceived depth, visual comfort and better overall experience compared to default fixed VR parameters. On the other hand, changing stereoscopic parameters dynamically in indoor VE still showed a promising result by providing a higher perceived depth quality. Our study demonstrated that our gaze-directed stereoscopic camera control method can be used as a suitable solution in VR in order to increase the quality of the experience, especially in VEs with wide depth range and more interaction.

4.1 Future Work

Even though we gathered valuable information in the first study and found multiple physiological measurements has a relationship with unpleasant scene setting, only saccade mean speed was found to be highest when cybersickness symptoms are present. In a further study with increased range of physiological response measurements, additional metrics can be found and by using them, our stereoscopic camera control method's effects on cybersickness can be explored. Besides the success of our stereoscopic camera control method, it can be integrated with foveated rendering technique and the effects of the foveated rendering can be investigated to find further possible improvements. Sense of presence in VEs is also another important concept which can show the effectiveness of the VEs and can be measured using questionnaires [56]. As an additional metric, sense of presence can be measured in order to explore the effects of our stereoscopic camera control method on presence in VR.

REFERENCES

- [1] K. Berntsen, R. C. Palacios, and E. Herranz, "Virtual reality and its uses: a systematic literature review," in *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality*, pp. 435–439, ACM, 2016.
- [2] N. Morina, H. Ijntema, K. Meyerbröker, and P. M. Emmelkamp, "Can virtual reality exposure therapy gains be generalized to real-life? a meta-analysis of studies applying behavioral assessments," *Behaviour research and therapy*, vol. 74, pp. 18–24, 2015.
- [3] A. Rizzo, J. Cukor, M. Gerardi, S. Alley, C. Reist, M. Roy, B. O. Rothbaum, and J. Difede, "Virtual reality exposure for ptsd due to military combat and terrorist attacks," *Journal of Contemporary Psychotherapy*, vol. 45, no. 4, pp. 255–264, 2015.
- [4] G. Riva, F. Mantovani, C. S. Capideville, A. Preziosa, F. Morganti, D. Villani, A. Gaggioli, C. Botella, and M. Alcañiz, "Affective interactions using virtual reality: the link between presence and emotions," *CyberPsychology & Behavior*, vol. 10, no. 1, pp. 45–56, 2007.
- [5] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: the panas scales.," *Journal of personality and social psychology*, vol. 54, no. 6, p. 1063, 1988.
- [6] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cognition & emotion*, vol. 9, no. 1, pp. 87–108, 1995.
- [7] M. M. Bradley, L. Miccoli, M. A. Escrig, and P. J. Lang, "The pupil as a measure of emotional arousal and autonomic activation," *Psychophysiology*, vol. 45, no. 4, pp. 602–607, 2008.
- [8] P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system

- activity distinguishes among emotions,” *Science*, vol. 221, no. 4616, pp. 1208–1210, 1983.
- [9] M. Soleymani, S. Asghari-Esfeden, Y. Fu, and M. Pantic, “Analysis of eeg signals and facial expressions for continuous emotion detection,” *IEEE Transactions on Affective Computing*, vol. 7, no. 1, pp. 17–28, 2016.
- [10] J. J. LaViola Jr, “A discussion of cybersickness in virtual environments,” *ACM SIGCHI Bulletin*, vol. 32, no. 1, pp. 47–56, 2000.
- [11] 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*, vol. 3, no. 3, pp. 203–220, 1993.
- [12] M. S. Dennison, A. Z. Wisti, and M. D’Zmura, “Use of physiological signals to predict cybersickness,” *Displays*, vol. 44, pp. 42–52, 2016.
- [13] E. Nalivaiko, S. L. Davis, K. L. Blackmore, A. Vakulin, and K. V. Nesbitt, “Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction time,” *Physiology & behavior*, vol. 151, pp. 583–590, 2015.
- [14] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, “Vergence–accommodation conflicts hinder visual performance and cause visual fatigue,” *Journal of vision*, vol. 8, no. 3, pp. 33–33, 2008.
- [15] T. Oskam, A. Hornung, H. Bowles, K. Mitchell, and M. H. Gross, “Oscam-optimized stereoscopic camera control for interactive 3d.,” *ACM Trans. Graph.*, vol. 30, no. 6, p. 189, 2011.
- [16] U. Celikcan, G. Cimen, E. B. Kevinc, and T. Capin, “Attention-aware disparity control in interactive environments,” *The Visual Computer*, vol. 29, no. 6-8, pp. 685–694, 2013.
- [17] M. Bernhard, C. Dell’mour, M. Hecher, E. Stavrakis, and M. Wimmer, “The effects of fast disparity adjustment in gaze-controlled stereoscopic applications,” in *Proceedings of the Symposium on Eye Tracking Research and Applications*, pp. 111–118, ACM, 2014.

- [18] L. E. Gurrieri, "Improvements in the visualization of stereoscopic 3d imagery," *3D Research*, vol. 6, no. 3, p. 25, 2015.
- [19] S. Davis, K. Nesbitt, and E. Nalivaiko, "A systematic review of cybersickness," in *Proceedings of the 2014 Conference on Interactive Entertainment*, pp. 1–9, ACM, 2014.
- [20] B. Cebeci, U. Celikcan, and T. K. Capin, "A comprehensive study of the affective and physiological responses induced by dynamic virtual reality environments," *Computer Animation and Virtual Worlds*, vol. 30, no. 3-4, p. e1893, 2019.
- [21] Y. Y. Kim, H. J. Kim, E. N. Kim, H. D. Ko, and H. T. Kim, "Characteristic changes in the physiological components of cybersickness," *Psychophysiology*, vol. 42, no. 5, pp. 616–625, 2005.
- [22] M. Meehan, S. Razzaque, B. Insko, M. Whitton, and F. P. Brooks, "Review of four studies on the use of physiological reaction as a measure of presence in stressful virtual environments," *Applied psychophysiology and biofeedback*, vol. 30, no. 3, pp. 239–258, 2005.
- [23] R. J. Snowden, K. R. O'Farrell, D. Burley, J. T. Erichsen, N. V. Newton, and N. S. Gray, "The pupil's response to affective pictures: Role of image duration, habituation, and viewing mode," *Psychophysiology*, vol. 53, no. 8, pp. 1217–1223, 2016.
- [24] T. Partala and V. Surakka, "Pupil size variation as an indication of affective processing," *International journal of human-computer studies*, vol. 59, no. 1-2, pp. 185–198, 2003.
- [25] G. Giannakakis, M. Pediaditis, D. Manousos, E. Kazantzaki, F. Chiarugi, P. G. Simos, K. Marias, and M. Tsiknakis, "Stress and anxiety detection using facial cues from videos," *Biomedical Signal Processing and Control*, vol. 31, pp. 89–101, 2017.
- [26] C. McCall, L. K. Hildebrandt, R. Hartmann, B. M. Baczkowski, and T. Singer, "Introducing the wunderkammer as a tool for emotion research: Unconstrained

- gaze and movement patterns in three emotionally evocative virtual worlds,” *Computers in Human Behavior*, vol. 59, pp. 93–107, 2016.
- [27] H. Chen, A. Dey, M. Billingham, and R. W. Lindeman, “Exploring pupil dilation in emotional virtual reality environments,” 2017.
- [28] S. Abokyi, J. Owusu-Mensah, and K. Osei, “Caffeine intake is associated with pupil dilation and enhanced accommodation,” *Eye*, vol. 31, no. 4, p. 615, 2017.
- [29] U. Erdem, F. C. Gundogan, U. A. Dinc, U. Yolcu, A. Ilhan, and S. Altun, “Acute effect of cigarette smoking on pupil size and ocular aberrations: A pre-and postsmoking study,” *Journal of ophthalmology*, vol. 2015, 2015.
- [30] J. F. Golding, “Motion sickness susceptibility questionnaire revised and its relationship to other forms of sickness,” *Brain research bulletin*, vol. 47, no. 5, pp. 507–516, 1998.
- [31] PyGaze, “Open-source toolbox for eye tracking in python.”
- [32] M. M. Schubert, A. Clark, and B. Annie, “The polar® oh1 optical heart rate sensor is valid during moderate-vigorous exercise,” *Sports medicine international open*, vol. 2, no. 03, pp. E67–E70, 2018.
- [33] A. Paillard, G. Quarck, F. Paolino, P. Denise, M. Paolino, J. F. Golding, and V. Ghulyan-Bedikian, “Motion sickness susceptibility in healthy subjects and vestibular patients: effects of gender, age and trait-anxiety,” *Journal of Vestibular Research*, vol. 23, no. 4, 5, pp. 203–209, 2013.
- [34] M. Lang, A. Hornung, O. Wang, S. Poulakos, A. Smolic, and M. Gross, “Non-linear disparity mapping for stereoscopic 3d,” in *ACM Transactions on Graphics (TOG)*, vol. 29, p. 75, ACM, 2010.
- [35] P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, H.-P. Seidel, and W. Matusik, “A luminance-contrast-aware disparity model and applications,” *ACM Transactions on Graphics (TOG)*, vol. 31, no. 6, p. 184, 2012.
- [36] A. J. Woods, T. Docherty, and R. Koch, “Image distortions in stereoscopic video systems,” in *Stereoscopic displays and applications IV*, vol. 1915, pp. 36–48, International Society for Optics and Photonics, 1993.

- [37] C. Ware, C. Gobrecht, and M. A. Paton, “Dynamic adjustment of stereo display parameters,” *IEEE transactions on systems, man, and cybernetics-part A: systems and humans*, vol. 28, no. 1, pp. 56–65, 1998.
- [38] A. Kulshreshtha and J. J. LaViola Jr, “Dynamic stereoscopic 3d parameter adjustment for enhanced depth discrimination,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 177–187, ACM, 2016.
- [39] A. Sherstyuk, A. Dey, C. Sandor, *et al.*, “Dynamic eye convergence for head-mounted displays improves user performance in virtual environments,” in *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pp. 23–30, ACM, 2012.
- [40] E. Arabadzhiyska, O. T. Tursun, K. Myszkowski, H.-P. Seidel, and P. Didyk, “Saccade landing position prediction for gaze-contingent rendering,” *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, p. 50, 2017.
- [41] Y. S. Pai, B. Tag, B. Outram, N. Vontin, K. Sugiura, and K. Kunze, “Gazesim: simulating foveated rendering using depth in eye gaze for vr,” in *ACM SIGGRAPH 2016 Posters*, p. 75, ACM, 2016.
- [42] M. F. Romero-Rondón, L. Sassatelli, F. Precioso, and R. Aparicio-Pardo, “Foveated streaming of virtual reality videos,” in *Proceedings of the 9th ACM Multimedia Systems Conference*, pp. 494–497, ACM, 2018.
- [43] J. Conti, B. Ozell, E. Paquette, and P. Renaud, “Adjusting stereoscopic parameters by evaluating the point of regard in a virtual environment,” *Computers & Graphics*, vol. 69, pp. 24–35, 2017.
- [44] P. Kellnhofer, P. Didyk, K. Myszkowski, M. M. Hefeeda, H.-P. Seidel, and W. Matusik, “Gazestereo3d: Seamless disparity manipulations,” *ACM Transactions on Graphics (TOG)*, vol. 35, no. 4, p. 68, 2016.
- [45] G. A. Koulteris, G. Drettakis, D. Cunningham, and K. Mania, “Gaze prediction using machine learning for dynamic stereo manipulation in games,” in *2016 IEEE Virtual Reality (VR)*, pp. 113–120, IEEE, 2016.
- [46] R. Konrad, A. Angelopoulos, and G. Wetzstein, “Gaze-contingent ocular parallax rendering for virtual reality,” *arXiv preprint arXiv:1906.09740*, 2019.

- [47] J. Jacobs, X. Wang, and M. Alexa, “Keep it simple: Depth-based dynamic adjustment of rendering for head-mounted displays decreases visual comfort,” *ACM Transactions on Applied Perception (TAP)*, vol. 16, no. 3, pp. 1–16, 2019.
- [48] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks, “The zone of comfort: Predicting visual discomfort with stereo displays,” *Journal of vision*, vol. 11, no. 8, pp. 11–11, 2011.
- [49] T. P. Runarsson and X. Yao, “Stochastic ranking for constrained evolutionary optimization,” *IEEE Transactions on evolutionary computation*, vol. 4, no. 3, pp. 284–294, 2000.
- [50] S. Johnson, “The nlopt nonlinear-optimization package,” URL <http://ab-initio.mit.edu/nlopt>, 2011.
- [51] E. Avan, U. Celikcan, T. K. Capin, and H. Gurcay, “Enhancing user experience in virtual reality with radial basis function interpolation based stereoscopic camera control,” *arXiv preprint arXiv:1911.04446*, 2019.
- [52] “Opendgl projection matrix.” http://www.songho.ca/opengl/gl_projectionmatrix.html. Accessed: 25-October-2019.
- [53] X. Liu and G. Wang, “Depth perception optimization based on multi-viewing spaces,” in *Proceedings of the 4th International Conference on Computer Science and Application Engineering*, pp. 1–6, 2020.
- [54] J. Zhang, S. Sclaroff, Z. Lin, X. Shen, B. Price, and R. Mech, “Minimum barrier salient object detection at 80 fps,” *IEEE International Conference on Computer Vision (ICCV)*, pp. 1404–1412, 2015.
- [55] B. Cebeci, , M. B. Askin, U. Celikcan, and T. K. Capin, “Two novel approaches towards optimal stereoscopic rendering: Gaze-directed and saliency-based methods for real-time enhancement of vr experience with hmds.” Unpublished manuscript, 2020.
- [56] B. G. Witmer and M. J. Singer, “Measuring presence in virtual environments: A presence questionnaire,” *Presence*, vol. 7, no. 3, pp. 225–240, 1998.

APPENDIX A

VAS QUESTIONNAIRE

Aşağıda verilen duyguları şu anda ne derece hissettiğinizi 1 ("hiç") ile 7 ("çok") arasında puanlayınız.

Şu anda kendinizi ne kadar Mutlu hissediyorsunuz?

Hiç Mutlu Hissetmiyorum				Mutlu Hissediyorum				Çok Mutlu Hissediyorum
1	2	3	4	5	6	7		

Şu anda kendinizi ne kadar Üzgün hissediyorsunuz?

Hiç Üzgün Hissetmiyorum				Üzgün Hissediyorum				Çok Üzgün Hissediyorum
1	2	3	4	5	6	7		

Şu anda kendinizi ne kadar Rahat hissediyorsunuz?

Hiç Rahat Hissetmiyorum				Rahat Hissediyorum				Çok Rahat Hissediyorum
1	2	3	4	5	6	7		

Şu anda kendinizi ne kadar Endişeli hissediyorsunuz?

Hiç Endişeli Hissetmiyorum				Endişeli Hissediyorum				Çok Endişeli Hissediyorum
1	2	3	4	5	6	7		

APPENDIX B

SIMULATOR SICKNESS QUESTIONNAIRE

No _____

Date _____

Sanal Gerçeklik Tutması Anketi





Aşağıdaki belirtilerin şu anda sizi ne ölçüde etkilediğini işaretleyin.

1. Genel rahatsızlık	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
2. Yorgunluk	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
3. Baş ağrısı	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
4. Göz yorgunluğu	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
5. Odaklanmada zorluk	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
6. Tükürükte artış	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
7. Terleme	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
8. Mide bulantısı	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
9. Konsantrasyonda güçlük	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
10. Baştta doluluk hissi	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
11. Bulanık görüş	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
12. Göz açıkken baş dönmesi	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
13. Göz kapalıyken baş dönmesi	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
14. Vertigo	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
15. Az miktarda mide bulantısı	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>
16. Geçirmek	<u>Hiç</u>	<u>Hafif</u>	<u>Orta</u>	<u>Şiddetli</u>

APPENDIX C

SCENE EVALUATION QUESTIONNAIRE

User ID Number:

Test Scene # :	
İzlediğiniz sekansın ortalama Resim (Görüntü) Kalitesini Değerlendiriniz	
İzlediğiniz sekansın verdiği ortalama Derinlik Hissini Değerlendiriniz	
İzleme sırasında hissettiğiniz ortalama Görsel Konforu Değerlendiriniz	
İzlediğiniz sekansa bir Genel Değerlendirme Puanı Veriniz	

Faydalı Açıklamalar

1- Resim (Görüntü) Kalitesi için:

Kötü 3B görüntü üretimi, çift görüntüye (ghosting efektine) ve görüntü keskinliği kaybına yol açmaktadır.

2-Derinlik Hissi için:

İyi üretilmiş 3B görüntüler, 2B görüntülerden farkedilir düzeyde ayrılacak denli iyi derinlik hissi sunmaktadır.

3- Görsel Konfor için:

Kötü üretilmiş 3B görüntüler size görsel rahatsızlık verir, gözlerinizi daha fazla yorar.