東海大学大学院 平成 26 年度博士論文

Study on 3D Gaze Measurement and Its Application to Analyze Visually Induced Motion Sickness in Stereoscopic Environment

指導 濱本 和彦 教授

東海大学大学院 総合理工学研究科

総合理工学専攻

SUNU WIBIRAMA

Abstract

This doctoral thesis presents a study on 3D gaze measurement and its implementation to analyze visually induced motion sickness (VIMS) in stereoscopic environment. One important safety issue in stereoscopic environment is VIMS. VIMS is a condition in which users of dynamic 3D contents feel symptoms of nausea, dizziness, or visual fatigue during or after exposure while they are being physically still. In this study, I present a novel 3D gaze tracking method with simpler 3D calibration technique. The proposed 3D gaze tracking system, along with electrocardiography (ECG) and simulator sickness questionnaire (SSQ), are used to investigate VIMS during exposure of dynamic 3D contents.

Chapter 1 contains background, purposes, and scientific contributions of this study. Decoupling of accommodation and vergence in viewing mechanism during dynamic 3D contents is one contributing factor of VIMS. Since vergence provides veridical depth information, investigation of human gaze in not only horizontal (X) and vertical (Y), but also in depth direction (Z) is important. Previous studies of VIMS used SSQ as subjective measurement. Objective measurement was performed using ECG and 2D gaze tracking. However, there was no research work that investigated relationship of 3D gaze, ECG data, and SSQ during VIMS occurrence. Furthermore, gaze tracking systems in previous studies were not compatible with active shutter glasses. Thus, the purpose of this study is to develop a novel 3D gaze tracking system that is compatible with active shutter glasses. The developed 3D gaze tracking system, ECG system, and SSQ information are used to investigate VIMS in stereoscopic environment.

Chapter 2 explains theory of stereoscopic environment, depth perception, and contributing factors of VIMS. To achieve 3D experience, human brain perceives left and right image as a single image. This mental ability of the brain to perceive two slight different images and extract depth information from them is called depth perception. Motion sickness is generally induced by vestibular stimuli while vision can also be a contributing factor. On the other side, VIMS occurs strictly caused by visual factor without vestibular stimulation.

Chapter 3 elaborates the development of 3D gaze tracking system, including design consideration, algorithm, and experimental validation. Optimized geometric method with only three calibration points in 3D calibration session is used to compute 3D point of gaze accurately. Experimental validation was performed to confirm accuracy of the proposed

3D gaze tracking system. The results show that the proposed 3D gaze tracking algorithm achieves better accuracy than conventional geometric method by average errors 0.83, 0.87, and 1.06 cm in X, Y, and Z direction, respectively. Comparison of the proposed 3D gaze tracking system with subjective judgment in depth measurement is also explained.

Chapter 4 presents a novel investigation of VIMS using the proposed 3D gaze tracking system, ECG, and SSQ in two different dynamic 3D contents containing low and high dynamic motions stimuli, respectively. Two-way statistical Analysis of Variance (ANOVA) on SSQ data shows that nausea and disorientation symptoms increase as amount of dynamic motions increase (nausea: $p<0.005$; disorientation: $p<0.05$). To reduce VIMS, ECG data suggests that user should perform voluntary gaze fixation at one point when experiencing vertical and horizontal motions in dynamic 3D contents. Observation of 3D gaze tracking data reveals that depth gaze is compressed by sustained forward motion. This finding may strengthen visual fatigue and VIMS caused by decoupling of accommodation and vergence in dynamic 3D contents. Furthermore, user who experiences VIMS tends to have unstable depth gaze than ones who does not experience VIMS.

Chapter 5 contains conclusion, implications of this study, and future works. In this study, I present a novel 3D gaze tracking system with simpler 3D calibration that is compatible with active shutter glasses. I demonstrate that detection of VIMS in stereoscopic environment can be done by observing 3D point of gaze using low frame rate consumer-grade cameras ($\pm$ 25 Hz). Experimental results show that horizontal and vertical motions are effective contributing factors of VIMS. Controlled gaze fixation during provoking scenes should be performed to reduce VIMS. Intense oscillation of depth gaze during provoking scene can be used as indicator of VIMS occurrence. The most important contribution of this study for general society is promoting methods for development of user-friendly 3D contents that considers safety issues and human factors. Based on this study, further development of mobile 3D gaze tracking system that allows free head movement is possible to expand functionality of the 3D gaze tracking in large screen immersive virtual environment.

Contents

List of Figures

List of Tables

Glossary

ANOVA Analysis of variance. 56, 82

CAVE Cave automatic virtual environment. 10

CCD Charge-coupled device. 40

CG Computer graphics. 69

CRT Cathode ray tube. 9

DLT Direct linear transformation. 44

DOF Depth of focus. 32

DVI Digital visual interface. 31

ECG Electrocardiography. 2, 10, 67

EOG Electrooculography. 4, 10

Geom. Geometric method. 9

HDDP Horizontally double-density pixel. 9

HMD Head mounted display. 10

HRV Heart rate variability. 67

IPD Interpupillary distance, distance between left and right eyeball. 27

IR LED Infrared light-emitting diode. 41

IR Infrared. 30, 38

IVE Immersive virtual environment. 10

LCD Liquid crystal display. 9, 10

LCP Left-camera projection. 50

MLP Multi-layered perceptron. 8, 9

MSQ Motion sickness questionnaire. 10, 64

NN Neural network method. 9

OKN Optokinetic nystagmus. 93

OpenCV Open Source Computer Vision library. 42

OpenGL Open Graphics Library. 42

PSOM Parameterized self-organizing map. 8, 9

RCP Right-camera projection. 50

RSA Respiratory sinus arrhythmia. 68

SA Sinoatrial node inside cardiac of human. 67

SE Stereoscopic environment, defined as active stereoscopic system with time-sharing mechanism for left and right eye. 1

SSC Scleral search coil. 4, 93

SSQ Simulator sickness questionnaire. 2, 10, 64

SVD Singular value decomposition, a method in linear algebra to solve $Ax = B$. 49

TV Television. 10

VIMS Visually induced motion sickness. 1, 31, 64

VOG Videooculography. 4

Acknowledgments

I would like to express my deepest gratitude to my supervisor Professor Dr. Kazuhiko Hamamoto for all advices, personal attention, and continuous encouragement during my doctoral program in Japan. It was a great experience to learn many things in Hamamoto Laboratory, ranging from engineering and science to culture and interaction. Thus, I thank all Hamamoto Laboratory members for all supports and cheers.

I am also deeply grateful to Professor Dr. Akira Tomono, Professor Dr. Hiroshi Ishii, Professor Dr. Kiyoshi Nosu, and Professor Dr. Kagayaki Kuroda for useful comments and suggestions for this thesis.

Special thanks to Japan International Cooperation Agency (JICA) and Tokai University for their financial support and chances to publish, present, and contribute in international conferences. I am really grateful to Mrs. Maria Sutherland and Ms. Tomomi Nishigaki of JICA Yokohama for their suggestion and assistance during my doctoral study. My appreciation is extended to Head of Graduate School of Science and Technology Tokai University, all lecturers, professors, and staffs for their supports.

Great appreciate for TexPad editor, LaTeX typesetting engine, and MacBookPro 7,1 that performed an estimated 20 quadrillion (10^6) CPU cycle to prepare thesis manuscript, thereby consuming about 19.6 kWh of electric energy and producing 6.8 kg of carbon dioxide.

Finally, I would like to thank all family members, including my wife, my son, both of our parents, brothers, and sisters for continuous support during my stay in Chiba, Japan.

Research is a transformation.
A transformation of knowledge consumer to knowledge producer,
a transformation of dependence to independence,
a transformation of student to colleague.
After all, the greatest encouragement of your achievements is
the feeling that you are doing something worthwhile with your life.
This research is dedicated to my wife and my beloved son
Thank you for your support during three years in Japan
Miseries, tears, and cheers were paid off by this work

三年間お世話になりました
楽しい時も、悲しい時も
色々な事を一緒に頑張りました
心から、本当にありがとうございました

Chapter 1

Introduction

1.1 Background

Stereoscopic environment (**SE**) becomes a ubiquitous technology in various areas, such as medical applications [1,2], augmented reality [3], rehabilitation [4], and industrial training [5]. In this study, I define stereoscopic environment by the following conditions: *a*) based on active 3D technology, i.e. 3D image is presented using time-sharing mechanism for left and right eye based on stereopsis concept; *b*) 3D image is presented on a single stereoscopic monitor; *c*) only one user views 3D image, multiple subjects at the same time are not supported; and *d*) user wears an active shutter glasses to see 3D image properly.

As the usage of SE increases, concerns on safety and human factors in 3D image technology also increase [6, 7]. One important safety issue in SE is *visually induced motion sickness* (**VIMS**). VIMS is a condition in which users of dynamic 3D contents feel symptoms of nausea, dizziness, or visual fatigue during or after exposure while they are being physically still [8, 9].

Unnatural viewing mechanism in SE is one cause of visual fatigue and visually induced motion sickness [10–13]. Users generally focus their view to the screen during stereoscopic contents exposure while their vergence eye movement changes according to depth cues in 3D contents. The decoupling of vergence and accommodation affects how user perceives depth information in SE [14]. Furthermore, Akai [14] explained several major problems

caused by inaccurate depth perception in SE, including incorrect size perception of virtual 3D object, virtual 3D object is seen at wrong distance, and eye strain. Since vergence provides veridical depth information [15], measuring 3D point of gaze in SE is important.

In their previous works [16–18], Jinjakam and Hamamoto investigated VIMS in large screen immersive virtual environment. The observation of VIMS was performed based on simulator sickness questionnaire (**SSQ**). Jinjakam found that users experienced VIMS when viewing a 3D computer graphics movie containing various dynamic motions. One well-known theory that explains this phenomenon is sensory-conflict theory [19, 20]. This theory states that during exposure of dynamic motions in visual contents, there is a mismatch between what the vestibular system is transducing and what vision system is signaling about body orientation. This sensory-conflict is detected when user experiences vection during exposure of 3D movie.

However, previous observation by Jinjakam and Hamamoto [16–18] were mainly based on subjective judgment, i.e. no objective measurement was undertaken during the experiment. Furthermore, users filled in the SSQ form after exposure of 3D contents. Thus, physiological effect of VIMS during exposure of 3D contents, duration of VIMS, and part of scene that cause VIMS were still unknown.

Other methods to objectively investigate physiological effect of VIMS are electrocardiography (**ECG**) or 3D gaze tracking data. ECG measures heartbeats and their variability through electrodes attached on chest. 3D gaze tracking estimates point of gaze using computer vision techniques in not only X (horizontal) and Y (vertical) direction, but also Z (depth) direction.

Previous research works on 3D gaze tracking in SE revealed that there was a trade-off between accuracy and practicality. High accuracy was achieved by implementing neural network based calibration methods [21–24]. The experimental results of those neural network based techniques were accurate. However, user had to fulfill intricate and tiring 3D calibration procedure to train the neural network algorithm.

On the other side, there were some researchers that proposed geometric method for 3D gaze tracking [25–28]. This method estimated 3D point of gaze by crossing extended virtual rays from both eyes to its fixation on 2D plane. Since there was no 3D calibration session, this method was quite practical. However, the experimental results were not ac-

curate due to general assumption that eyeball radius and distance between two eyes were same for all users. The inaccuracy of 3D point of gaze estimation was enlarged by unlikely intersection of virtual rays in 3D space. Furthermore, all previous research works were performed in anaglyph stereoscopic 3D technology (i.e. there was no information on how 3D gaze tracking was used simultaneously with active shutter glasses).

Therefore, development of precise 3D gaze tracking system to investigate VIMS in SE takes into account several constraints:

- user has to wear active shutter glasses to see 3D image;
- the active shutter glasses should be properly installed on the gaze tracking headgear;
- the eye camera should be properly aligned to reduce inaccuracy in 3D gaze measurement while allowing user to gaze without any obstruction;
- stereo camera cannot be used to measure depth of virtual 3D objects since those objects are generated using stereoscopic technology;
- the gaze tracking system should provide simpler 3D gaze calibration technique to avoid fatigue;
- the gaze tracking system estimates eye radius and distance between two eyes to improve its accuracy; and
- the gaze tracking system can be implemented to investigate physiological effect of VIMS during exposure of stereoscopic 3D movie.

1.2 Measurement of eye movements

Eye is a unique part of human body. It serves as a visual sensor, an ocular motor system, as well as an indicator for healthiness of vestibular and visual system. Rayner [29] proposed three eras of research in eye movements: (1879–1920) indicated by discovery of many basic eye movements facts, including saccadic eye movements, latency in eye movements, and the size of perceptual span; (1930–1958) characterized by focus in applied

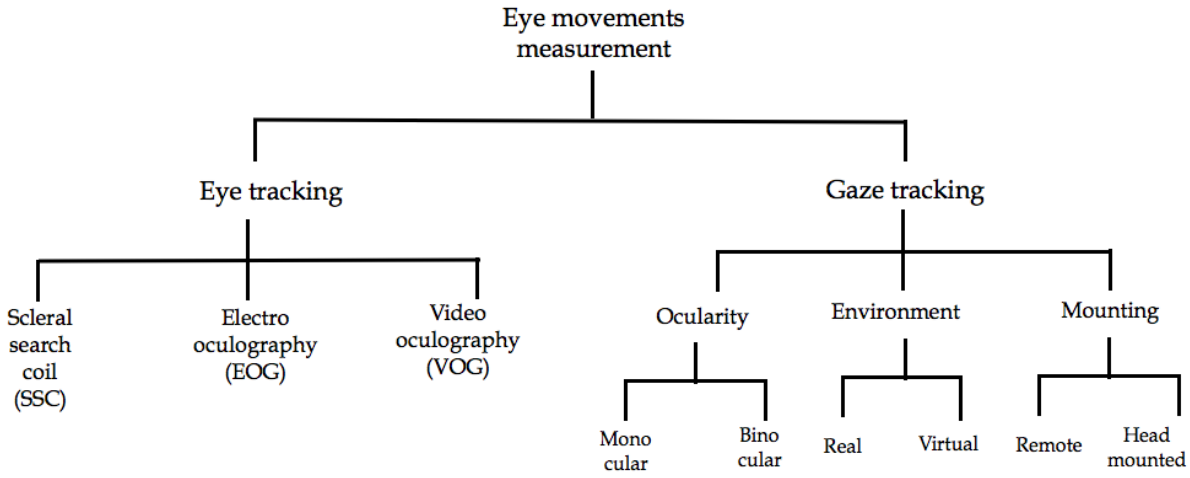


Figure 1.1: Classification of eye movements measurement

applications that were used in behavior and psychology research area; (1970–1998) marked by improvements in recording technology and methods of measuring eye movements with higher accuracy.

Methods of measuring eye movements can be classified into eye tracking and gaze tracking system (Fig. 1.1). Eye tracking is an electronic device that measures the position of eyeball either by directly attaching a measurement device on the eye and its surrounding, or by using a video-recording system. Eye tracking merely measures the Euclidean position of eyeball without information of gaze and attention of the user.

On the other side, gaze tracking measures both the position of eyeball and gaze of user. Gaze tracking is generally used as computer interface, assistive technology, medical, and ergonomic purpose. Gaze tracking systems are classified based on their type of ocularity (monocular or binocular), environment of measurement (real or virtual), and type of mounting (head-mounted or remote).

1.2.1 Classification of eye tracking

Eye tracking methods can be classified into *scleral search coil* (**SSC**), *electrooculography* (**EOG**), and *videooculography* (**VOG**). SSC is generally described as the most precise eye tracking technique since the measurement is performed by directly attaching a contact

lens containing coil magnetometer to the surface of eyeball [30]. To adhere the coil firmly on the eyeball, a vacuum is made by suction through polyvinyl tube. The user is then surrounded by alternating-current magnetic field generated from external field coils. The eyeball position is detected through change of electrical voltage generated by the magnetic field [31].

Nevertheless, there are some disadvantages of this method. For example, SSC suffers from coil slipping and possibility of electrical leads breakage [31,32]. The coil slipping and electrical leads breakage occur due to eye blink and force exerted by the eyelids. Moreover, this method can only be used not more than 40 minutes due to various effects such as lid irritation, conjunctival hyperemia (red eye), corneal staining, and reduction of visual acuity [33, 34].

EOG is one kind of eye tracking technique based on change in electrical potential between the cornea and ocular fundus [35,36]. Normally, EOG requires attachment of several electrodes surrounding the area of the eye. For this reason, EOG can be used as a diagnostic device for everyday activity [37]. The EOG is also preferred for recording eye movements in sleep research and in infants behavior research [38].

Advances in digital image processing algorithm and video recording technology over the past two decades have encouraged researchers to develop video-based eye tracking system known as VOG. This technique uses single or stereo camera mounted in front of the eyeball to measure horizontal, vertical, and torsional eye movements. These three directional movements are often known as 3D eye movements [39]. The small number of existing image processing systems for eye movements measurement fall into two broad categories: those based iris striations tracking using cross correlation or template matching method [39–43] and those employing natural or artificial landmark tracking [44–49].

1.2.2 Classification of gaze tracking

I reviewed 50 research papers of gaze tracking development from 2000 to 2014 [21–28, 50–91]. I classified gaze tracking systems based on type of ocularity (monocular or binocular), environment of measurement (real or virtual), or type of mounting (remote or head-mounted) as shown in Fig.1.1 and Table 1.1. Note that a gaze tracking system

Table 1.1: Literature review of research in gaze tracking during 2000–2014

Ocularity	No. of Papers	Environment	No. of Papers	Mounting	No.of papers	No.of subjects	No. of papers	Year	No.of papers
Monocular	28	Real 2D	10	Remote	12	1-10	41	2000-2002	5
Binocular	22	Real 3D	11	Head-	38	11-20	5	2003-2005	5
		Virtual 2D	21	mounted		>20	4	2006-2008	14
		Virtual 3D	8					2009-2011	15
								2012-2014	11

classified based on its ocularity is also available to be classified based on their measurement environment or type of mounting. For example: gaze tracking system **A** is classified into monocular type. At the same time, gaze tracking **A** is also a remote gaze tracking used in virtual 2D environment.

Ocularity means amount of camera needed to extract user's gaze. Monocular type uses only single eye camera while the binocular one tracks both left and right eye. As shown in Table 1.1, there are 28 and 22 papers for monocular [24, 50, 52, 57, 58, 60, 61, 64, 65, 67–69, 71–78, 80, 82, 84, 85, 88–91] and binocular type [21–23, 25–28, 51, 53–56, 59, 62, 63, 66, 70, 79, 81, 83, 86, 87], respectively.

Based on their measurement environment, gaze tracking systems can be classified into real and virtual environment. If the gaze tracking system is used in real environment, the system is normally equipped with eye and scene camera. Eye camera captures user's eye while scene camera captures real scene seen by user.

The eye and scene camera are installed on special glasses or helmet for user's flexibility and mobility. The personal computer used for data processing is commonly in compact size or portable model. Researcher estimates correspondence between eye coordinate and point of gaze in real environment by calibrating eye and scene camera simultaneously.

In virtual environment, scene camera is not used since users see digital stimuli generated using computer graphics technology. The stimuli are presented on large 2D screen or computer monitor. Gaze tracker can be either head-mounted or remote mounting type since portability and mobility are not the main issue. Depends on the method, users are normally asked to fix their heads or to locate their heads on a chin rest to avoid measurement error.

Based on direction of gaze, both gaze tracking systems in real and virtual environment

are classified into 2D (horizontal and vertical) or 3D (horizontal, vertical, and depth) type. There are two types of calibration: 2D and 3D calibration. In real environment, 2D and 3D calibration are normally performed by providing 2D calibration points on a wall or 3D points on a known 3D object, such as 3D cube [52] or 2D planes located at various distance [54]. There are 10 and 11 research papers of gaze tracking research in real 2D environment [60, 64, 65, 67, 76–79, 82, 90] and real 3D environment [50–57, 66, 83, 87], respectively.

There are 21 and 8 research papers of gaze tracking in virtual 2D environment [58, 59, 61–63, 68–75, 80, 81, 84–86, 88, 89, 91] and virtual 3D environment [21–28]. In virtual environment, 2D and 3D stimuli are presented in computer monitor. 3D stimuli are generated by stereoscopic technology. All previous research works of 3D gaze tracking in virtual 3D environment used passive technology that required the usage of anaglyph red-blue or polarized 3D glasses. Although active 3D technology is now common in consumer-level personal computer, capturing eye images through active shutter glasses is rather difficult since the shutters of left and right eye open and close alternately. Thus, all researchers [21–28] preferred to capture eye images through anaglyph red-blue or polarized glasses. Further explanation of 3D gaze tracking in virtual 3D environment is presented in section 1.3.

I also classified gaze tracking research based on subjects involved in experiment and year of research. Most of previous works involved 1 until 10 subjects for their experimental validation (41 papers). Several research works exhaustively validated their system with 11–20 subjects (5 papers) or more than 20 subjects (4 papers). I also found that amount of research papers increased after 2005. During 2000–2005, I found 10 research papers while other 40 papers were published in 2006–2014.

1.3 3D gaze tracking in virtual 3D environment

In this study, I focused literature review on 3D gaze tracking in virtual 3D environment since the aim of research is accurate 3D gaze tracking to analyze VIMS in stereoscopic contents. Based on method of measurement, previous research works of 3D gaze tracking in virtual 3D environment fall into two broad categories: those based on neural

network [21–24] and geometric method [25–28]. Essig *et al*. [21, 22] used parameterized self-organizing map (**PSOM**) [92,93], which is a rapidly learning variant of self-organizing maps proposed by Kohonen [94]. Soon after the 2D calibration finished, the user was asked to gaze at 3D calibration points arranged in 3 x 3 x 3 array. The 3D calibration points were shown in anaglyph stereoscopic technology. In summary, PSOM heuristically translated 2D gaze on screen to 3D position in virtual space using 27 calibration points.

Lee *et al*. [24] developed a monocular 3D gaze tracker. Lee *et al.* used first and forth purkinje images, inter-distance between these purkinje images, and the size of pupil as inputs of multi-layered perceptron (**MLP**) [95,96], a variance of neural network algorithm. First, the user calibrated their gaze on 2D plane by seeing at 4 calibration points. Next, the user was asked to calibrate their gaze in 3D calibration session. In the 3D calibration session, five virtual planes were positioned at various distance ranging from 10 to 50 cm. Each of virtual plane consisted of 9 calibration points. Those 45 calibration points were needed to train the MLP algorithm with six iterations. Lee *et al.* reported an impressive result with average errors of the 3D gaze estimation about 0.48 cm on the X-axis, 0.77 cm on the Y-axis, and 4.59 cm along the Z-axis in 3D space.

Although those neural network based 3D gaze trackers yield rigorous results by large iterations in training session and computation of parameters, the user normally has to gaze at more than 20 calibration points in 3D gaze calibration session, which is time-consuming and exhausting.

The second approach is geometric approach. This method approximates 3D point of gaze using virtual rays extending from both eyes to its fixation on a 2D plane [25–28]. The 3D point of gaze was estimated by intersecting virtual rays in 3D space. Normally, the position of the eyeball was assumed unchanged during measurement. This position was estimated based on average interpupillary distance, which was ranging from 5 to 7 cm [97, 98]. Duchowski [26] and Daugherty *et al*. [27] used geometric method to measure 3D point of gaze over anaglyph stereogram video. Duchowski used head-mounted gaze tracker embedded in head-mounted display while Daugherty preferred remote gaze tracker to estimate 3D point of gaze. Boev *et al*. [28] also used geometric method to measure binocular volume of interest on a display equipped with lenticular lens.

Although those conventional geometric methods did not carry out complicated 3D gaze

Table 1.2: Comparison of several 3D gaze tracking methods in virtual 3D environment

	NN [22]	**NN [24]**	**Geom. [27]**	**Geom. [28]**	**This study**
Year	2006	2012	2010	2012	2014
3D technology	passive (anaglyph)	passive (anaglyph)	passive (anaglyph)	passive (anaglyph)	active (shutter)
Display	CRT monitor	LCD display	LCD display	HDDP display	LCD display
Type of device	commercial EyeLink II	lab-made prototype	commercial Tobii ET-1750	commercial EyeLink 1000	lab-made prototype
Mounting	head-mounted	head-mounted	remote	remote	head-mounted
2D calibration	○	○	○	○	○
3D calibration	○	○	×	×	○
Amount of 3D stimuli	27	45	×	×	3
Amount of eye camera	2	1	1	1	4
Core algorithm	PSOM	MLP	Conventional geometric	Conventional geometric	Optimized geometric
Average error in cm	$\pm 0.52(X)$, $\pm 0.82(Y)$, $\pm 2.53(Z)$	$\pm 0.48(X)$, $\pm 0.77(Y)$, $\pm 4.59(Z)$	n/a	$\pm 0.4(X)$, $\pm 0.4(Y)$, $\pm 5(Z)$	$\pm 0.83(X)$, $\pm 0.87(Y)$, $\pm 1.06(Z)$

Abbreviation: neural network method (**NN**), geometric method (**Geom.**), cathode ray tube (**CRT**), liquid crystal display (**LCD**), parameterized self-organizing map (**PSOM**), multi-layered perceptron (**MLP**), horizontally double-density pixel (**HDDP**), not available (n/a).

calibration session, the 3D gaze measurement result was not accurate since intersection of virtual rays in 3D space was unlikely. Furthermore, they assumed that distance between two eyeballs and radius of eyeball were same for all users. In this study, I encompass simplicity of geometric method and accuracy of neural network. I developed an optimized geometric method by providing a tailored measurement of user dependent parameters and simple 3D calibration using only three calibration points. Table 1.2 shows comparison of 3D gaze tracking system in this study and previous research works.

Table 1.3: Research on visually induced motion sickness (VIMS) using subjective judgment and objective measurement method during 2002 - 2014.

Name of research	**Display**	**Participant (Age)**	**Measurement methods**		
			Questionnaire	**Bio-signal**	**Eye response**
VIMS 2002 [99]	Single large screen	10 (50–71)	7-ratings	ECG, blood pressure	pupillometer
VIMS 2003 [100]	HMD	16 (20–25)	vection scale	—	EOG
VIMS 2004 [101]	Single large screen	9 (18–32)	MSQ	—	2D gaze tracking
VIMS 2007 [102]	Single large screen	7 (avg.23)	7-ratings	ECG, blood pressure	pupillometer
VIMS 2007 [103]	Rear-projection TV	12 (avg.22)	Bagshaw and Stott's Scale	—	2D gaze tracking
VIMS 2008 [104]	Single large screen	56 (18–25)	SSQ	ECG, blood pressure	—
VIMS 2008 [6]	Four screens CAVE	18 (avg.20)	SSQ	ECG	—
VIMS 2013 [105]	LCD monitor	19 (avg.20)	SSQ	—	2D gaze tracking
This research 2014	LCD monitor	20 (21–27)	SSQ	ECG	3D gaze tracking

Note:
a) Abbreviation: electrocardiography (**ECG**), simulator sickness questionnaire (**SSQ**), motion sickness questionnaire (**MSQ**), television (**TV**), immersive virtual environment (**IVE**), cave automatic virtual environment (**CAVE**), head mounted display (**HMD**), electrooculography (**EOG**), liquid crystal display (**LCD**).
b) Measurement of eye responses can be performed either by measuring the size of pupil (pupillometer) or direction of gaze (gaze tracking).

1.4 Measurement of visually induced motion sickness

Previous research works found that visually induced motion sickness (VIMS) affected heart rate variability [99,102,104,106] as well as gaze behavior of user [103,105,107,108]. To investigate the occurrence of VIMS, researchers normally combined subjective opinion

with one of objective measurement, either electrocardiography (ECG) data or gaze tracking. One cause of VIMS is decoupling of accommodation and convergence during stereoscopic exposure [11], which also affects user depth perception on the virtual contents [15]. Other cause of VIMS is sensory-conflict between vestibular system (balance system located inside the inner part of ear) and vision system [19, 20]. These contributing factors of VIMS can be observed using 3D gaze measurement device during exposure of stereoscopic 3D contents.

Nevertheless, most of gaze measurements in those previous research works were performed only in 2D direction, i.e. only X and Y gaze position were obtained during exposure of stimuli [103, 107, 108]. Furthermore, instead of using real scenery presented in stereoscopic 3D, optical flow contents or alternate black-and-white vertical stripes were generally presented to the user [103, 107, 108]. Therefore, there was no information of how stereoscopic 3D movie of real scenery affected 3D gaze behavior and heart rate variability at the same time. To the best knowledge of the author, there was no research work that combined simulator sickness questionnaire (SSQ), electrocardiography (ECG), and 3D gaze tracking in one experiment to investigate VIMS.

In this research, I present a novel investigation method to get better understanding about VIMS during stereoscopic 3D contents. SSQ, ECG, and 3D gaze position measurement are incorporated in stereoscopic environment. SSQ score is used to check whether a participant experiences VIMS or not. However, SSQ score merely informs subjective condition after exposure. Furthermore, SSQ score does not reflect objective information about physiological effect of VIMS. Hence, additional investigation of occurrence and duration of VIMS are needed.

ECG data can be used to detect the occurrence of VIMS during exposure of stereoscopic contents. By measuring heart rate variability, activity of both *sympathetic* and *parasympathetic* nerves are known. These nerves serve as a clear and objective indicator whether a participant experiences VIMS in particular time.

Observing 3D gaze position during stereoscopic exposure is important because human gaze provides physiological cues of VIMS invoked by various motions and scenes. In summary, gaze behavior in 3D space is valuable information to understand symptoms of VIMS during provoking scenes.

1.5 Purpose and contribution of research

The purpose of this study is to develop an accurate 3D gaze tracking system that can be implemented in stereoscopic environment to analyze visually induced motion sickness (VIMS). Based on several constraints mentioned in section 1.1, I developed a head-mounted 3D gaze tracking system that can be used properly with consumer-level active shutter glasses. I considered proper installation of eye camera to obtain clear eye image without obstructing field of view of the user. To achieve this goal, infrared reflectors were installed in the gaze tracking system to capture brief eye images while allowing user to view the 3D movie. I also took into account appropriate installation of Nvidia 3D Vision® glasses by providing "Y"-shaped shutter glasses holder on the gaze tracking headgear.

I developed a new algorithm by improving conventional geometric method and refining the capability of the system to measure gaze in depth direction. First, I developed a measurement method to obtain customized user dependent parameters using computer vision techniques. Next, I developed 3D calibration technique using only three calibration points to avoid exhausting 3D calibration session. I performed experimental validation to confirm the accuracy of the system. I also compared the 3D gaze tracking system with subjective depth judgment to inspect usefulness of the system for development of more user-friendly stereoscopic contents. I implemented the proposed 3D gaze tracking system, ECG, and SSQ to observe visually induced motion sickness (VIMS) during exposure of stereoscopic 3D movie. I also discuss the relationship between VIMS, ECG, and 3D gaze behavior.

There are two main scientific contributions in this study:

1. I developed a novel gaze tracking system to measure 3D point of gaze accurately using only three calibration points in 3D calibration session.

2. I conducted a novel investigation of VIMS in stereoscopic environment using the 3D gaze tracking, electrocardiography (ECG), and simulator sickness questionnaire (SSQ).

Two implications of my study include:

1. Future development of gaze-to-perception conversion for better human and computer interaction in stereoscopic environment.

2. Promotion and encouragement of more user-friendly stereoscopic contents that consider safety issues and human factors.

This doctoral thesis is structured as follows:

- Chapter 1 explains background, purposes, and scientific contributions of this research. This chapter also provides several basic understanding of gaze tracking and measurement of VIMS to help reader appreciating the rest of this doctoral thesis.

- Chapter 2 contains literature review of stereoscopic environment, including depth perception, and the Nvidia 3D Vision® system. This chapter also presents a brief overview of VIMS and contributing factors of VIMS.

- Chapter 3 elaborates the proposed 3D gaze tracking system, including design consideration, algorithm, and experimental validation. Comparison of the proposed 3D gaze tracking system with subjective judgment in depth measurement is also explained.

- Chapter 4 explains a novel investigation of VIMS using the proposed 3D gaze tracking system, simulator sickness questionnaire (SSQ), and electrocardiography (ECG).

- Chapter 5 contains conclusion, implications of this study, and several possible future works.

Bibliography

[1] D. Meziat, M. H. Van Horn, S. Weeks, and E. Bullitt, "3D stereo interactive medical visualization," *IEEE Computer Graphics and Applications*, vol. 25, no. 5, pp. 67–71, 2005.

[2] H. Liao, *3D Medical Imaging and Augmented Reality for Image-Guided Surgery*, ch. 27, pp. 589–602. Springer New York, 2011.

[3] H. Benko, R. Jota, and A. Wilson, "MirageTable: freehand interaction on a projected augmented reality tabletop," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 199–208, ACM, 2012.

[4] A. Gargantini, M. Bana, and F. Fabiani, "Using 3D for rebalancing the visual system of amblyopic children," in *The 2011 International Conference on Virtual Rehabilitation (ICVR)*, pp. 1–7.

[5] S. White, M. Prachyabrued, T. Chambers, C. Borst, and D. Reiners, "Low-cost simulated MIG welding for advancement in technical training," *Virtual reality*, vol. 15, no. 1, pp. 69–81, 2011.

[6] H. Watanabe and H. Ujike, "The Activity of ISO/Study Group on "Image Safety" and Three Biological Effect," in *Second International Symposium on Universal Communication 2008*, pp. 210–214, 2008.

[7] A. Solimini, A. Mannocci, D. Di Thiene, and G. La Torre, "A survey of visually induced symptoms and associated factors in spectators of three dimensional stereoscopic movies," *BMC Public Health*, vol. 12, no. 1, p. 779, 2012.

[8] R. S. Kennedy, J. Drexler, and R. C. Kennedy, "Research on Visually Induced Motion Sickness," *Applied Ergonomics*, vol. 41, no. 4, pp. 494–503, 2010.

[9] H. Ujike and H. Watanabe, "Effects of stereoscopic presentation on visually induced motion sickness," in *Proceedings of SPIE 7863, Stereoscopic Displays and Applications XXII,786314*, vol. 7863, pp. 786314–786314–6, 2011.

[10] H. Takada and M. Miyao, "Visual Fatigue and Motion Sickness Induced by 3D Video Clip," *Forma*, vol. 27, pp. S67–S76, 2012.

[11] 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. 33,1–30, 2008.

[12] K. Ukai and P. A. Howarth, "Visual fatigue caused by viewing stereoscopic motion images: Background, theories, and observations," *Displays*, vol. 29, no. 2, pp. 106–116, 2008.

[13] T. Takeda, K. Hashimoto, N. Hiruma, and Y. Fukui, "Characteristics of accommodation toward apparent depth," *Vision Research*, vol. 39, no. 12, pp. 2087–2097, 1999.

[14] C. Akai, *Depth Perception in Real and Virtual Environments: An Exploration to Individual Differences*. PhD thesis, 2007.

[15] M. Mon-Williams, J. R. Tresilian, and A. Roberts, "Vergence provides veridical depth perception from horizontal retinal image disparities," *Experimental Brain Research*, vol. 133, no. 3, pp. 407–413, 2000.

[16] C. Jinjakam and K. Hamamoto, "Study on Parallax Affect on Simulator Sickness in One-screen and Three-screen Immersive Virtual Environment," *Proceedings of The School of Information and Telecommunication Engineering Tokai University*, vol. 4, no. 1, pp. 34–39, 2011.

[17] C. Jinjakam and K. Hamamoto, "Analysis of Simulator Sickness in IVE by PCA," *Transaction of the Japan Society for Simulation Technology*, vol. 4, no. 4, pp. 194–199, 2012.

[18] C. Jinjakam and K. Hamamoto, "Parallax, position and height difference effects on simulator sickness in immersive virtual environment," in *Biomedical Engineering International Conference (BMEiCON), 2013 6th*, pp. 1–4, 2013.

[19] J. LaViola and J. Joseph, "A Discussion of Cybersickness in Virtual Environments," *SIGCHI Bulletin*, vol. 32, pp. 47–56, Jan. 2000.

[20] J. E. Bos, W. Bles, and E. L. Groen, "A Theory on Visually Induced Motion Sickness," *Displays*, vol. 29, no. 2, pp. 47–57, 2008.

[21] K. Essig, M. Pomplun, and H. Ritter, "Application of a novel neural approach to 3D gaze tracking: Vergence eye-movements in autostereograms," in *Proceedings of the Twenty-Sixth Annual Meeting of the Cognitive Science Society*, pp. 357–362, 2004.

[22] K. Essig, M. Pomplun, and H. Ritter, "A neural network for 3D gaze recording with binocular eye trackers," *The International Journal of Parallel, Emergent and Distributed Systems*, vol. 21, no. 2, pp. 79–95, 2006.

[23] M. Xiao, *Eye-movement Analysis for Advanced Human-computer Interaction*. PhD thesis, 2008. AAI3313759.

[24] J. W. Lee, C. W. Cho, K. Y. Shin, E. C. Lee, and K. R. Park, "3D gaze tracking method using Purkinje images on eye optical model and pupil," *Optics and Lasers in Engineering*, vol. 50, no. 5, pp. 736–751, 2012.

[25] R. Danforth, A. Duchowski, R. Geist, and E. McAliley, "A platform for gaze-contingent virtual environments," in *Smart Graphics (Papers from the 2000 AAAI Spring Symposium, Technical Report SS-00-04)*, pp. 66–70, 2000.

[26] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*. Springer-Verlag New York, Inc., 2nd ed., 2007.

[27] B. C. Daugherty, A. T. Duchowski, D. H. House, and C. Ramasamy, "Measuring vergence over stereoscopic video with a remote eye tracker," in *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, pp. 97–100, ACM, 2010.

[28] A. Boev, M. Hanhela, A. Gotchev, T. Utirainen, S. Jumisko-Pyykko, and M. Hannuksela, "Parameters of the human 3D gaze while observing portable autostereoscopic display: a model and measurement results," in *IS&T/SPIE Electronic Imaging*, pp. 830407–830407–11, International Society for Optics and Photonics, 2012.

[29] K. Rayner, "Eye Movements in Reading and Information Processing: 20 Years of Research," *Psychological Bulletin*, vol. 3, no. 124, pp. 372–422, 1998.

[30] D. Robinson, "A Method of Measuring Eye Movement using A Scleral Search Coil Technique in A Magnetic Field," *IEEE Transactions on Biomedical Engineering*, vol. 10, pp. 137–145, 1963.

[31] H. Collewijn, F. van der Mark, and T. C. Jansen, "Precise recording of human eye movements," *Vision Research*, vol. 15, no. 3, pp. 447–450, IN5, 1975.

[32] H. Collewijn, J. d. Steen, L. Ferman, and T. Jansen, "Human ocular counterroll: assessment of static and dynamic properties from electromagnetic scleral coil recordings," *Experimental Brain Research*, vol. 59, no. 1, pp. 1432–1106, 1985.

[33] A. L. Duwaer, G. Van Den Brink, G. Van Antwerpen, and C. J. Keemink, "Comparison of subjective and objective measurements of ocular alignment in the vertical direction," *Vision Research*, vol. 22, no. 8, pp. 983–987, 989, 1982.

[34] E. Irving, J. Zacher, R. Allison, and M. Callander, "Effects of Scleral Search Coil Wear on Visual Function," *Investigative Ophthalmology & Visual Science*, vol. 44, no. 5, pp. 1933–1938, 2003.

[35] O. H. Mowrer, T. C. Ruch, and N. E. Miller, "The Corneo-Retinal Potential Difference as the Basis of the Galvanometric Method of Recording Eye Movements," *American Journal of Physiology*, vol. 114, pp. 423–428, 1936.

[36] M. Brown, M. Marmor, Vaegan, E. Zrenner, M. Brigell, and M. Bach, "ISCEV Standard for Clinical Electro-oculography (EOG) 2006," *Documenta Ophtalmologica*, vol. 113, no. 3, pp. 205–212, 2006.

[37] A. Bulling, D. Roggen, and G. Troster, "Wearable EOG goggles: Seamless sensing and context-awareness in everyday environments," *Journal of Ambient Intelligence and Smart Environments*, vol. 1, no. 2, pp. 157–171, 2009.

[38] J. H. Manning, E. Courchesne, and P. T. Fox, "Intrinsic connectivity network mapping in young children during natural sleep," *NeuroImage*, vol. 83, no. 0, pp. 288 – 293, 2013.

[39] S. Wibirama, S. Tungjitkusolmun, and C. Pintavirooj, "Dual-Camera Acquisition for Accurate Measurement of Three-Dimensional Eye Movements," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 8, no. 3, pp. 238–246, 2013.

[40] M. Hatamian and D. J. Anderson, "Design Considerations for a Real-Time Ocular Counterroll Instrument," *IEEE Transactions on Biomedical Engineering*, vol. BME-30, no. 5, pp. 278–288, 1983.

[41] T. Vieville and D. Masse, "Ocular counter-rolling during active head tilting in humans," *Acta Otolaryngology*, vol. 103, no. 3-4, pp. 280–290, 1987.

[42] S. T. Moore, T. Haslwanter, I. S. Curthoys, and S. T. Smith, "A geometric basis for measurement of three-dimensional eye position using image processing," *Vision Research*, vol. 36, no. 3, pp. 445–459, 1996.

[43] D. Zhu, S. T. Moore, and T. Raphan, "Robust and real-time torsional eye position calculation using a template-matching technique," *Computer Methods and Programs in Biomedicine*, vol. 74, no. 3, pp. 201–209, 2004.

[44] J. A. Parker, R. V. Kenyon, and L. R. Young, "Measurement of Torsion from Multitemporal Images of the Eye Using Digital Signal Processing Techniques," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 1, pp. 28–36, 1985.

[45] S. Yamanobe, T. Taira, T. Morizono, T. Kamio, and T. Yagi, "Eye movement analysis system using computerized image recognition," *Archives of Otolaryngology Head & Neck Surgery*, vol. 116, no. 3, pp. 338–341, March 1990.

[46] D. Ott, F. Gehle, and R. Eckmiller, "Video-oculographic measurement of 3-dimensional eye rotations," *Journal of Neuroscience Methods*, vol. 35, no. 3, pp. 229–234, 1990.

[47] J. E. Bos and B. de Graaf, "Ocular torsion quantification with video images," *IEEE Transactions on Biomedical Engineering*, vol. 41, no. 4, pp. 351–357, 1994.

[48] J. Kim, "A simple pupil-independent method for recording eye movements in rodents using video," *Journal of Neuroscience Methods*, vol. 138, no. 1-2, pp. 165–171, 2004.

[49] A. A. Migliaccio, H. G. MacDougall, L. B. Minor, and C. C. Della Santina, "Inexpensive system for real-time 3-dimensional video-oculography using a fluorescent marker array," *Journal of Neuroscience Methods*, vol. 143, no. 2, pp. 141–150, 2005.

[50] A. Sugimoto, A. Nakayama, and T. Matsuyama, "Detecting a gazing region by visual direction and stereo cameras," in *Proceedings of the 16th International Conference on Pattern Recognition.*, vol. 3, pp. 278–282 vol.3, 2002.

[51] I. Mitsugami, N. Ukita, and M. Kidode, "Estimation of 3D gazed position using view lines," *Proceedings of The 12th International Conference on Image Analysis and Processing*, pp. 466–471, 2003.

[52] S. M. Munn and J. B. Pelz, "3D Point-of-regard, Position and Head Orientation from a Portable Monocular Video-based Eye Tracker," in *Proceedings of the 2008 Symposium on Eye Tracking Research and Applications*, ETRA '08, (New York, NY, USA), pp. 181–188, ACM, 2008.

[53] F. Pirri, M. Pizzoli, and A. Rudi, "A general method for the point of regard estimation in 3D space," in *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 921–928, 2011.

[54] S. Shimizu and H. Fujiyoshi, "Acquisition of 3D Gaze Information from Eyeball Movements Using Inside-out Camera," in *Proceedings of the 2nd Augmented Human International Conference*, AH '11, (New York, NY, USA), pp. 6:1–6:7, ACM, 2011.

[55] J. Stoll, S. Kohlbecher, S. Marx, E. Schneider, and W. Einhauser, "Mobile three dimensional gaze tracking," *Studies in health technology and informatics*, vol. 163, pp. 616–622, 2011.

[56] M. Y. Kim, S. Yang, and D. Kim, "Head-mounted binocular gaze detection for selective visual recognition systems," *Sensors and Actuators A: Physical*, vol. 187, no. 0, pp. 29–36, 2012.

[57] K. Essig, D. Dornbusch, D. Prinzhorn, H. Ritter, J. Maycock, and T. Schack, "Automatic Analysis of 3D Gaze Coordinates on Scene Objects Using Data from Eye-tracking and Motion-capture Systems," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, (New York, NY, USA), pp. 37–44, ACM, 2012.

[58] K. Asai, N. Osawa, H. Takahashi, Y. Sugimoto, S. Yamazaki, A. Samejima, and T. Tanimae, "Eye mark pointer in immersive projection display," in *IEEE Proceedings of Virtual Reality 2000.*, pp. 125–132, 2000.

[59] T. Ohno and N. Mukawa, "A Free-head, Simple Calibration, Gaze Tracking System That Enables Gaze-based Interaction," in *Proceedings of the 2004 Symposium on Eye Tracking Research and Applications*, ETRA '04, (New York, NY, USA), pp. 115–122, ACM, 2004.

[60] A. Tsukada, M. Shino, M. Devyver, and T. Kanade, "Illumination-free gaze estimation method for first-person vision wearable device," in *Proceedings of The 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pp. 2084–2091, 2011.

[61] A. Villanueva, G. Daunys, D. Hansen, M. Bohme, R. Cabeza, A. Meyer, and E. Barth, "A geometric approach to remote eye tracking," *Universal Access in the Information Society*, vol. 8, no. 4, pp. 241–257, 2009.

[62] Y. Ebisawa, K. Abo, and K. Fukumoto, *Head-Free, Remote Eye-Gaze Detection System with Easy Calibration Using Stereo-Calibrated Two Video Cameras*, vol. 174, pp. 151–155. Springer Berlin Heidelberg, 2011.

[63] Z. Zhu and Q. Ji, "Novel eye gaze tracking techniques under natural head movement," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 12, pp. 2246–2260, 2007.

[64] F. Li, S. Munn, and J. Pelz, "A model-based approach to video-based eye tracking," *Journal of Modern Optics*, vol. 55, no. 4-5, pp. 503–531, 2008.

[65] N. Schneider, P. Bex, E. Barth, and M. Dorr, "An Open-source Low-cost Eye-tracking System for Portable Real-time and Offline Tracking," in *Proceedings of the 1st Conference on Novel Gaze-Controlled Applications*, NGCA '11, (New York, NY, USA), pp. 8:1–8:4, ACM, 2011.

[66] K. Takemura, Y. Kohashi, T. Suenaga, J. Takamatsu, and T. Ogasawara, "Estimating 3D Point-of-regard and Visualizing Gaze Trajectories Under Natural Head Movements," in *Proceedings of the 2010 Symposium on Eye-Tracking Research and Applications*, ETRA '10, (New York, NY, USA), pp. 157–160, ACM, 2010.

[67] D. Li, J. Babcock, and D. J. Parkhurst, "openEyes: A Low-cost Head-mounted Eye-tracking Solution," in *Proceedings of the 2006 Symposium on Eye Tracking Research and Applications*, ETRA '06, (New York, NY, USA), pp. 95–100, ACM, 2006.

[68] T. Pfeiffer, "Towards Gaze Interaction in Immersive Virtual Reality : Evaluation of a Monocular Eye Tracking Set-Up," in *Virtuelle und Erweiterte RealitatFunfter Workshop der GIFachgruppe VRAR*, pp. 81–92, 2008.

[69] V. Tanriverdi and R. J. K. Jacob, "Interacting with Eye Movements in Virtual Environments," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '00, (New York, NY, USA), pp. 265–272, ACM, 2000.

[70] C. Zhang, J. Chi, Z. Zhang, X. Gao, T. Hu, and Z. Wang, "Gaze estimation in a gaze tracking system," *Science China Information Sciences*, vol. 54, no. 11, pp. 2295–2306, 2011.

[71] C. S. Lin, "An eye behavior measuring device for VR system," *Optics and Lasers in Engineering*, vol. 38, no. 6, pp. 333 – 359, 2002.

[72] C. W. Cho, J. W. Lee, K. Y. Shin, E. C. Lee, K. R. Park, H. Lee, and J. Cha, "Gaze Detection by Wearable Eye-Tracking and NIR LED-Based Head-Tracking Device Based on SVR," *ETRI Journal*, vol. 34, pp. 542–552, Aug. 2012.

[73] A. Haffegee and R. Barrow, "Eye Tracking and Gaze Based Interaction within Immersive Virtual Environments," in *Computational Science ICCS 2009* (G. Allen, J. Nabrzyski, E. Seidel, G. Albada, J. Dongarra, and P. Sloot, eds.), vol. 5545 of *Lecture Notes in Computer Science*, pp. 729–736, Springer Berlin Heidelberg, 2009.

[74] Y. J. Ko, E. C. Lee, and K. R. Park, "A robust gaze detection method by compensating for facial movements based on corneal specularities," *Pattern Recogn Letters*, vol. 29, no. 10, pp. 1474–1485, 2008.

[75] E. C. Lee and K. R. Park, "A robust eye gaze tracking method based on a virtual eyeball model," *Mach. Vision Appl.*, vol. 20, no. 5, pp. 319–337, 2009.

[76] W. J. Ryan, A. T. Duchowski, and S. T. Birchfield, "Limbus/Pupil Switching for Wearable Eye Tracking Under Variable Lighting Conditions," in *Proceedings of the 2008 Symposium on Eye Tracking Research and Applications*, ETRA '08, (New York, NY, USA), pp. 61–64, ACM, 2008.

[77] A. Tsukada and T. Kanade, "Automatic Acquisition of a 3D Eye Model for a Wearable First-person Vision Device," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, (New York, NY, USA), pp. 213–216, ACM, 2012.

[78] Y. Ishiguro, A. Mujibiya, T. Miyaki, and J. Rekimoto, "Aided Eyes: Eye Activity Sensing for Daily Life," in *Proceedings of the 1st Augmented Human International Conference*, AH '10, (New York, NY, USA), pp. 25:1–25:7, ACM, 2010.

[79] H. Yamazoe, A. Utsumi, T. Yonezawa, and S. Abe, "Remote Gaze Estimation with a Single Camera Based on Facial-feature Tracking Without Special Calibration Actions," in *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, (New York, NY, USA), pp. 245–250, ACM, 2008.

[80] J. Song, S. Cho, S.-Y. Baek, K. Lee, and H. Bang, "GaFinC: Gaze and Finger Control interface for 3D model manipulation in {CAD} application," *Computer-Aided Design*, vol. 46, no. 0, pp. 239 – 245, 2014. 2013 {SIAM} Conference on Geometric and Physical Modeling.

[81] C. Hennessey and J. Fiset, "Long Range Eye Tracking: Bringing Eye Tracking into the Living Room," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA '12, (New York, NY, USA), pp. 249–252, ACM, 2012.

[82] K. Lukander, S. Jagadeesan, H. Chi, and K. Muller, "OMG!: A New Robust, Wearable and Affordable Open Source Mobile Gaze Tracker," in *Proceedings of the 15th International Conference on Human-computer Interaction with Mobile Devices and Services*, MobileHCI '13, (New York, NY, USA), pp. 408–411, ACM, 2013.

[83] C. Hennessey and P. Lawrence, "Noncontact Binocular Eye-Gaze Tracking for Point-of-Gaze Estimation in Three Dimensions," *Biomedical Engineering, IEEE Transactions on*, vol. 56, no. 3, pp. 790–799, 2009.

[84] J. W. Lee, H. Heo, and K. R. Park, "A Novel Gaze Tracking Method Based on the Generation of Virtual Calibration Points," *Sensors*, vol. 13, no. 8, pp. 10802–10822, 2013.

[85] J. Barabas, R. Goldstein, H. Apfelbaum, R. Woods, R. Giorgi, and E. Peli, "Tracking the line of primary gaze in a walking simulator: Modeling and calibration," *Behavior Research Methods, Instruments, & Computers*, vol. 36, no. 4, pp. 757–770, 2004.

[86] T. Nagamatsu, J. Kamahara, and N. Tanaka, "Calibration-free Gaze Tracking Using a Binocular 3D Eye Model," in *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '09, (New York, NY, USA), pp. 3613–3618, ACM, 2009.

[87] L. Paletta, K. Santner, G. Fritz, and H. Mayer, "3D recovery of human gaze in natural environments," in *Proceedings of SPIE Vol. 8662 - Intelligent Robots and Computer Vision XXX: Algorithms and Techniques*, vol. 8662, pp. 86620K–86620K–10, 2013.

[88] C. H. Morimoto and M. R. Mimica, "Eye gaze tracking techniques for interactive applications," *Computer Vision and Image Understanding*, vol. 98, no. 1, pp. 4 – 24, 2005. Special Issue on Eye Detection and Tracking.

[89] J. W. Bang, E. C. Lee, and K. R. Park, "New computer interface combining gaze tracking and brainwave measurements," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 4, pp. 1646–1651, 2011.

[90] A. Villanueva and R. Cabeza, "A Novel Gaze Estimation System With One Calibration Point," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 38, no. 4, pp. 1123–1138, 2008.

[91] E. C. Lee, K. R. Park, M. C. Whang, and J. Park, "Robust Gaze Tracking Method for Stereoscopic Virtual Reality Systems," in *The 12th International Conference on Human-Computer Interaction: Intelligent Multimodal Interaction Environments* (J. Jacko, ed.), vol. 4552 of *Lecture Notes in Computer Science*, pp. 700–709, Springer Berlin Heidelberg, 2007.

[92] J. Walter and H. Ritter, "Rapid learning with parametrized self-organizing maps," *Neurocomputing*, vol. 12, no. 2-3, pp. 131–153, 1996. Current European Neurocomputing Research.

[93] J. Walter, C. Nolker, and H. Ritter, "The PSOM algorithm and applications," in *Proceedings of the symposium on neural computation*, pp. 758–764, 2000.

[94] T. Kohonen, "The self-organizing map," in *Proceedings of the IEEE*, vol. 78, pp. 1464–1480, 1990.

[95] D. W. Ruck, S. K. Rogers, M. Kabrisky, M. E. Oxley, and B. W. Suter, "The multilayer perceptron as an approximation to a Bayes optimal discriminant function," *IEEE Transactions on Neural Networks*, vol. 1, no. 4, pp. 296–298, 1990.

[96] R. Collobert and S. Bengio, "Links Between Perceptrons, MLPs and SVMs," in *Proceedings of the Twenty-first International Conference on Machine Learning*, ICML '04, (New York, NY, USA), pp. 23–, ACM, 2004.

[97] M. Poyade, A. Reyes-Lecuona, and R. Viciana-Abad, *Influence of Binocular Disparity in Depth Perception Mechanisms in Virtual Environments*, ch. 2, pp. 13–22. Springer London, 2009.

[98] M. Lambooij, M. Fortuin, I. Heynderickx, and W. IJsselsteijn, "Visual discomfort and visual fatigue of stereoscopic displays: a review," *Journal of Imaging Science*, vol. 53, no. 3, pp. 30201–(14), 2009.

[99] T. Ando, A. Tanaka, S. Fukasaku, R. Takada, M. Okada, K. Ukai, K. Shizuka, H. Oyamada, H. Toda, T. Taniyama, T. Usui, M. Yoshizawa, T. Kiryu, M. Takagi, S. Saida, and T. Bando, "Pupillary and cardiovascular responses to a video movie in senior human subjects," *Autonomic Neuroscience*, vol. 97, no. 2, pp. 129–135, 2002.

[100] N. A. Webb and M. J. Griffin, "Eye Movement, Vection, and Motion Sickness with Foveal and Peripheral Vision," *Aviation, Space, and Environmental Medicine*, vol. 74, no. 6, pp. 622–625, 2003.

[101] M. B. Flanagan, J. G. May, and T. G. Dobie, "The role of vection, eye movements and postural instability in the etiology of motion sickness," *Journal of Vestibular Research*, vol. 14, pp. 335–346, Jan. 2004.

[102] H. Oyamada, A. Iijima, A. Tanaka, K. Ukai, H. Toda, N. Sugita, M. Yoshizawa, and T. Bando, "A pilot study on pupillary and cardiovascular changes induced by stereoscopic video movies," *Journal of NeuroEngineering and Rehabilitation*, vol. 4, no. 1, pp. 1–7, 2007.

[103] C. Diels, K. Ukai, and P. A. Howarth, "Visually Induced Motion Sickness with Radial Displays: Effects of Gaze Angle and Fixation," *Aviation, Space, and Environmental Medicine*, vol. 78, no. 7, pp. 659–665, 2007.

[104] N. Sugita, M. Yoshizawa, A. Tanaka, K. Abe, S. Chiba, T. Yambe, and S. Nitta, "Quantitative evaluation of effects of visually-induced motion sickness based on causal coherence functions between blood pressure and heart rate," *Displays*, vol. 29, no. 2, pp. 167–175, 2008.

[105] C. Guo, C. W. Tsoi, Y. L. Wong, and K. C. Yu, "Visually Induced Motion Sickness during Computer Game Playing," in *Contemporary Ergonomics and Human Factors 2013: Proceedings of the International Conference on Ergonomics and Human Factors 2013, Cambridge, UK, 15-18 April 2013*, p. 51, Taylor & Francis.

[106] C. Jinjakam, *Study on Simulator Sickness in Immersive Virtual Environment and Proposal for Safety Guidelines*. PhD thesis, Graduate School of Science and Technology Tokai University, 2013.

[107] C. Diels and P. A. Howarth, "Visually induced motion sickness: Single-versus dual-axis motion," *Displays*, vol. 32, no. 4, pp. 175–180, 2011.

[108] J. Yang, C. Guo, R. So, and R. Cheung, "Effects of eye fixation on visually induced motion sickness: are they caused by changes in retinal slip velocity?," in *Proceedings of the human factors and ergonomics society 55th annual meething*, pp. 1220–1224, 2011.

Chapter 2

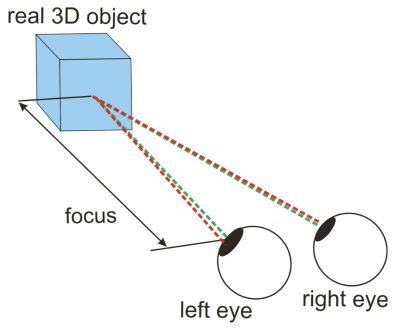
Stereoscopic Environment

2.1 Principle of stereoscopic vision

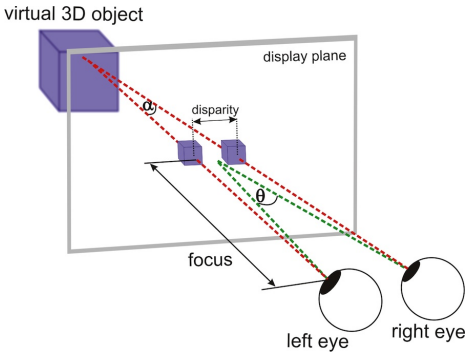
Stereoscopic display technology is developed based on *stereopsis* in human visual system. Human eyes are separated by interpupillary distance (**IPD**) ranging from 5 to 7 cm [1, 2]. As human eyes are separated horizontally, each eye has its own view of the world scene. To achieve 3D experience, human brain perceives left and right image as a single image. This mental ability of the brain to perceive two slight different images and extract depth information from them is called *depth perception*.

Depth perception through human eyes is achieved by using binocular vergence and lens accommodation that provide fixation and focus, respectively. In real environment, users fixate and focus their view on real 3D object. However, depth perception in virtual environment is achieved by unnatural viewing condition. During exposure of stereoscopic 3D contents, all images are projected onto a 2D screen that requires both eyes to always focus on the screen although virtual 3D object is gazed at different depths [3, 4]. These different viewing mechanisms are depicted in Fig.2.1.

Human eyes are sensitive in differentiating depth of two objects. The smallest depth interval between two objects that can be detected by human eye is known as *depth-discrimination threshold*. *Stereoacuity* is the depth-discrimination threshold when binocular disparity is the only cue to depth. Previously published work found that human could discriminate position of two objects with 14 arcsec of angular disparity, which is equivalent to 1.3 mm



(a) Stereoscopic viewing in real environment: vergence and accommodation are coupled.



(b) Stereoscopic viewing in virtual environment: user accommodates at screen while gazing to the position of virtual 3D object.

Figure 2.1: Concept of stereoscopic viewing: by fusing left and right images, the brain perceives depth of 3D object. Parallax angle is defined as the difference between angle formed by both eyes-virtual object and both eyes-display plane (θ-α).

depth interval at a distance of 1.12 meter [5]. However, the ability of perceiving depth decreases for distance beyond 2 meter. This is due to only slight change of vergence angle when the eyes see far object [3].

2.2 Depth perception cues

There are several cues that influence human depth perception in both real and virtual environment [4, 6, 7]:

1. *Accommodation*: the ability of crystal lens inside the eyeball to produce sharp focus of the viewed objects. Crystal lens becomes thicker when seeing near distance object. In contrast, crystal lens becomes thinner when seeing far distance object.

2. *Binocular disparity*: separation of left and right eye that produces slightly different images in left and right retina. Distance of two eyes is ranging from 5 to 7 cm. Binocular disparity is the strongest depth cue among the others.

3. *Vergence*: eye movement through equal angles in opposite directions to produce a disjunctive movement. Horizontal vergence occurs when a person changes fixation from an object in one depth plane to other depth planes.

4. *Aerial perspective*: effect of atmosphere on the appearance of scene and objects looked from distance location. If the atmosphere contains high degree of air, farther objects become bluer and less contrast.

5. *Linear perspective*: this cue is often known as "geometric" or "photographic" perspective. Perspective is an important monocular depth cue. This cue produces an illusion of scaling between foreground and background objects, e.g. foreground objects are seen bigger than background objects, parallel lines that recede into the distance appear to converge.

6. *Motion parallax*: depth cue created by moving-image displays or rotation of an object. Movement of farther objects is perceived slower. For example, when looking at outside scenery from a moving train, near electric poles seem to move faster than distant trees.

7. *Occlusion*: if a foreground object occludes a background object, the foreground object seems closer to us since human cannot see through the foreground object. This cue provides information on depth order instead of the amount of depth.

8. *Relative size*: size of 2D or 3D object increases as the distance between user and the object decreases, and vice versa.

9. *Relative height*: object located closer to us is perceived to be taller than object located farther.

10. *Shading*: variations of light source position and specularity of light produce different irradiance of surface. Brighter objects are commonly seen as nearer than the dim ones.

11. *Shadow*: obstruction of an opaque or semi-opaque object causes variations of irradiance on the surface of object.

12. *Texture gradient*: Textured surfaces of material provide depth cue since a nearer object shows more apparent texture than the farther ones.

Binocular disparity, accommodation, and vergence are classified into binocular cues since these cues are normally provided for both eyes [7]. These cues create depth sensation in most of 3D pictures. On the other side, shading, shadow, linear perspective, relative height, relative size, texture gradient, and motion perspective are classified as monocular cues since these cues can be perceived through only one eye. Even 2D picture can provide depth sensation by utilizing these cues [8, 9].

2.3 Nvidia 3D Vision® system

Most of current 3D technologies adapt conventional stereoscopic principle introduced by Wheatstone in 1838 [10]. Recently, active 3D technology is gaining wider attention. To produce 3D experience, an active stereoscopic display presents left and right image alternately with very high frequency to prevent user from recognizing flickering in the 3D contents. Generally, a frequency of 120 Hz is used in a stereoscopic 3D display.

To ensure that the right eye does not see image for the left eye and vice versa, the user has to wear an active shutter glasses that blind either the right or the left eye. The side that blinds user's eye has to alternate synchronously with the display. This 3D image visualization technique provides effective time-sharing of the display for the left and right eye. By providing a pair of stereo images with disparity on screen, the user perceives a stereoscopic 3D experience.

In this research, I used Nvidia 3D Vision® for viewing stereoscopic 3D contents [11]. Several prior research works utilized Nvidia 3D Vision®, including 3D haptic-based modeling system [12], augmented reality with freehand interaction [13], navigation for visually impaired people [14], digital exhibition of archaeological structure [15], welding training for industrial worker [16], and performance investigation for sport player [17]. Compared with commonly used anaglyph stereoscopic technology, Nvidia 3D vision® is better since it preserves original color of the 3D contents.

Nvidia 3D Vision® system consists of two main hardwares: infrared (**IR**) emitter and a lightweight active shutter glasses (50 gram of weight). The IR transmitter is used to synchronize the glasses with the frame rate of the running application. The 3D vision

system is coupled with a 23 inch I-O Data LCD-3D231XBR-S monitor. In order to do stereoscopic 3D rendering in OpenGL, the computer needs to be configured with the right graphics card. Since standard consumer-level graphics cards do not support stereoscopic 3D rendering, Nvidia Quadro® 5000 graphics card with driver version 275.36 was used in this research.

Nvidia Quadro® 5000 is a sophisticated graphics card that provides four times faster performance than other Nvidia graphics cards [18]. Nvidia Quadro® 5000 can process up to 950 million triangles per second, setting the standard for 3D performance benchmarks. It also drives an entire visual supercomputing platform, incorporating hardware and software that enables stereoscopic 3D rendering, scalable visualization, and 3D high-definition broadcasting. The graphics card also supports double digital visual interface (**DVI**) display ports, allowing easier extension for flexible experiment environment. Stereoscopic 3D rendering algorithm is explained in subsection 3.3.1.

2.4 Visually induced motion sickness

Visually induced motion sickness (**VIMS**) is a condition in which symptoms of nausea, dizziness, or visual fatigue are felt by viewers of dynamic 3D image during or after exposure while they are being physically still [19]. VIMS is often referred to as simulator sickness, cinerama sickness, or cybersickness depending on type of environment and motion seen by the viewers [20].VIMS and motion sickness produce similar symptoms, although stimuli that cause both sicknesses are not necessarily same. Motion sickness is generally induced by vestibular stimuli while vision can also be a contributing factor. On the other side, VIMS occurs strictly caused by visual factor without vestibular stimulation. This visual-vestibular relationship is a basic component that underlies the development of sensory-conflict theory [21].

According to Kennedy *et al.* [19], symptoms of VIMS can be clustered into three general types named: (1) nausea; (2) oculomotor; and (3) disorientation. Nausea (*N*) symptoms related to gastrointestinal distress such as nause, stomatch awareness, salivation, and burping. Oculomotor (*O*) symptoms related to visual disturbance such as eyestrain (also

known as *asthenopia*), which is discomfort resulting from dysfunctional visual states such as dryness of the eyes, orbital pain, heavy eyelids, difficulty in focusing, blurred vision, and headache. Disorientation (*D*) symptoms related to vestibular disturbances, such as dizziness and vertigo.

Lambooij *et al.* [2] described several factors from steroscopic 3D contents that may contribute to visual fatigue and VIMS:

1. **Excessive binocular disparity**

   In order to obtain 3D depth information, fusion limits of human visual system can be remarkably small. With longer stimuli and vergence eye movements, human brain can fuse disparity information about 4.93 degrees for crossed and 1.57 degrees for uncrossed without any *diplopia* (double vision). Normally, one-degree disparity is accepted as appropriate setting for most of 3D contents. If 3D contents are configured in excessive disparity setting, the contents may cause eyestrain.

2. **Mismatch of accommodation and convergence**

   During stereoscopic 3D exposure, accommodation distance is constant, but vergence distance varies depending on degree of screen disparity. This process of unnatural viewing is argued as main source of visual fatigue and VIMS [22].

3. **Seeing beyond zone of comfortable viewing**

   When human sees object in natural condition, range of depth of focus (**DOF**) goes along with range of fusion. Objects located farther from fixation point are perceived as more blurred. Limit of DOF under one degree of disparity and eyes focusing at infinity are about one-third of a diopter. Limit of DOF and disparity show high resemblance if both of them are calculated in distance, creating what is called as "comfortable viewing distances within one degree of disparity". As an example, range of comfortable viewing distances for 3D contents in a computer screen with 50 cm viewing distance is about 44 cm (near) to 58 cm (far). Outside this comfortable viewing range, possibility of VIMS occurrence is higher.

4. **Stereoscopic distortions**

   Imperfect production of 3D contents caused by choice of camera, camera configuration, 2D-to-3D conversion, coding, transmission, and rendering can induce VIMS.

5. **3D artifacts**

   Real-time 2D-to-3D conversion algorithms are promising method, particularly in real-time 3D contents transmission. However, such algorithms rarely consider accurate depth information by assuming that depth from screen disparity is enough. Such assumption may produce artifacts in 3D contents, such as incorrect depth values assigned to depth layers. This error further produces incorrect blurring, unnatural visualization, or flickering images.

Besides contents and displays factors, individual factors may affect the occurrence of VIMS. This situation is rather difficult to answer due to human complexity and individual proneness to motion sickness. According to La Viola [21], there are several individual factors that affect the occurrence of VIMS:

1. **Age**

   Susceptibility to VIMS is strongest between 2 to 12 years old. It decreases rapidly from 12 to 21 years and the more slowly thereafter. Around 50 years of age, VIMS is almost nonexistent.

2. **Illness**

   Illness is one important factor that contributes to the occurrence of VIMS. Someone suffering from cold, influenza, fatigue, sleep loss, hangover, upset stomach, ear infection, and upper respiratory illness tends to experience VIMS in virtual environment. Therefore, Frank *et al.* [23] suggested that user of virtual environment simulator should be healthy and free from illness.

3. **Position of user in virtual environment**

   Position of user in virtual environment or simulator may contribute to the occurrence of VIMS. Based on postural instability theory [24], sitting is better than standing. In large virtual environment, such as CAVE virtual reality [25] or HoloStage^{TM} [26],

user who controls the simulation (driver) is less susceptible to VIMS than those who are passive participants (passengers). To avoid VIMS, Jinjakam [27] suggested that the driver should be the tallest person among all users of simulator. The position of the passenger should be close to the driver while avoiding position near screen border connections.

Bibliography

[1] M. Poyade, A. Reyes-Lecuona, and R. Viciana-Abad, *Influence of Binocular Disparity in Depth Perception Mechanisms in Virtual Environments*, ch. 2, pp. 13–22. Springer London, 2009.

[2] M. Lambooij, M. Fortuin, I. Heynderickx, and W. IJsselsteijn, "Visual discomfort and visual fatigue of stereoscopic displays: a review," *Journal of Imaging Science*, vol. 53, no. 3, pp. 30201–(14), 2009.

[3] I. P. Howard and B. J. Rogers, *Geometry of Binocular Images*, vol. 2, pp. 1–9. Toronto: I Porteous, 2002.

[4] C. Akai, *Depth Perception in Real and Virtual Environments: An Exploration to Individual Differences*. PhD thesis, 2007.

[5] S. P. McKee and D. G. Taylor, "The precision of binocular and monocular depth judgments in natural settings," *Journal of Vision*, vol. 10, no. 10, pp. 1–13, 2010.

[6] L. Lipton, *The Stereographics Developer's Handbook - Background on Creating Images for CrystalEyes and SimulEyes*. Stereographics Corporation, 1997.

[7] 3DConsortium, "3DC Safety Guidelines for Dissemination of Human-friendly 3D," April 2010.

[8] I. P. Howard, *Seeing in Depth: Volume 1 Basic Mechanisms*. I Porteous, Toronto, 2002.

[9] I. P. Howard, *Seeing in Depth: Volume 2 Depth Perception*. I Porteous, Toronto, 2002.

[10] C. Wheatstone, "On some remarkable, and hitherto unobserved, phenomena of binocular vision," *Philosophical Transactions of the Royal Society of London*, vol. 128, pp. 371–394, 1838.

[11] Nvidia Corporation, "Nvidia 3D Vision Product Overview. Available:http://www.nvidia.com/object/3d-vision-main.html," May 2014.

[12] S. Kamuro, K. Minamizawa, and S. Tachi, "3D Haptic modeling system using ungrounded pen-shaped kinesthetic display," in *Virtual Reality Conference (VR), 2011 IEEE*, pp. 217–218.

[13] H. Benko, R. Jota, and A. Wilson, "MirageTable: freehand interaction on a projected augmented reality tabletop," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 199–208, ACM, 2012.

[14] S. Vitek, M. Klima, L. Husnik, and D. Spirk, "New possibilities for blind people navigation," in *International Conference on Applied Electronics*, pp. 1–4.

[15] H. T. Tanaka, K. Hachimura, K. Yano, S. Tanaka, K. Furukawa, T. Nishiura, M. Tsutida, W. Choi, and W. Wakita, "Multimodal digital archiving and reproduction of the world cultural heritage "Gion Festival in Kyoto"," 2010.

[16] S. White, M. Prachyabrued, T. Chambers, C. Borst, and D. Reiners, "Low-cost simulated MIG welding for advancement in technical training," *Virtual reality*, vol. 15, no. 1, pp. 69–81, 2011.

[17] T. Hoinville, A. Naceri, J. Ortiz, E. Bernier, and R. Chellali, "Performances of experienced and novice sportball players in heading virtual spinning soccer balls," in *Virtual Reality Conference (VR), 2011 IEEE*, pp. 83–86.

[18] Nvidia Corporation, "Nvidia Quadro 5000 Product Overview. Available: http://www.nvidia.com/object/product-quadro-5000-us.html," May 2014.

[19] R. S. Kennedy, J. Drexler, and R. C. Kennedy, "Research on Visually Induced Motion Sickness," *Applied Ergonomics*, vol. 41, no. 4, pp. 494–503, 2010.

[20] J. E. Bos, W. Bles, and E. L. Groen, "A Theory on Visually Induced Motion Sickness," *Displays*, vol. 29, no. 2, pp. 47–57, 2008.

[21] J. LaViola and J. Joseph, "A Discussion of Cybersickness in Virtual Environments," *SIGCHI Bulletin*, vol. 32, pp. 47–56, Jan. 2000.

[22] 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. 33,1–30, 2008.

[23] L. Frank, R. S. Kennedy, M. E. McCauley, K. R. S. Root, R. W., and A. Bittner, "Simulator Sickness: Sensory Motor Disturbances Induced in Flight Simulators," in *The Image II Conferences*, pp. 417–426, 1983.

[24] L. J. Hettinger and G. E. Riccio, "Visually Induced Motion Sickness in Virtual Environments," *Presence: Teleoper. Virtual Environ.*, vol. 1, pp. 306–310, July 1992.

[25] G. C. Burdea and P. Coiffet, *Virtual Reality Technology*. John Wiley & Sons, Inc., 2003.

[26] C. Jinjakam and K. Hamamoto, "Study on Parallax Affect on Simulator Sickness in One-screen and Three-screen Immersive Virtual Environment," *Proceedings of The School of Information and Telecommunication Engineering Tokai University*, vol. 4, no. 1, pp. 34–39, 2011.

[27] C. Jinjakam, *Study on Simulator Sickness in Immersive Virtual Environment and Proposal for Safety Guidelines*. PhD thesis, Graduate School of Science and Technology Tokai University, 2013.

Chapter 3

3D Gaze Measurement

3.1 Hardware configuration

3.1.1 Design consideration

There are three possible scenarios to use Nvidia 3D Vision® glasses with gaze tracking headgear, as shown in Fig.3.1. The first scenario is to directly put the cameras in front of the Nvidia 3D Vision® glasses. This type of installation is not complicated. However, the cameras cannot capture the left and right eye images simultaneously since the Nvidia 3D Vision® glasses alternately open and close the shutters with 60 Hz frequencies for each side. Moreover, direct cameras installation obstructs field of view of the user.

The second scenario is to install the gaze tracker cameras below the Nvidia 3D Vision® glasses. The advantage of this installation is similar to the first scenario. However, the vertical field of view of user will be limited and obstructed when gazing the bottom part of screen. The third scenario is to use additional hardware, such as infrared (**IR**) light and hot mirror (IR reflector). In this case, the vertical field of view of the user is preserved and the cameras are able to capture images of eyes clearly. However, since the glasses are slightly positioned farther from the eyes, the gaze tracking headgear has to take into account a mechanism to hold the Nvidia 3D Vision® glasses properly. Unfortunately, no commercial gaze tracking device is available for such purpose. Based on the third scenario, I developed a new gaze tracking headgear that is suitable to be used with Nvidia 3D Vision® glasses.

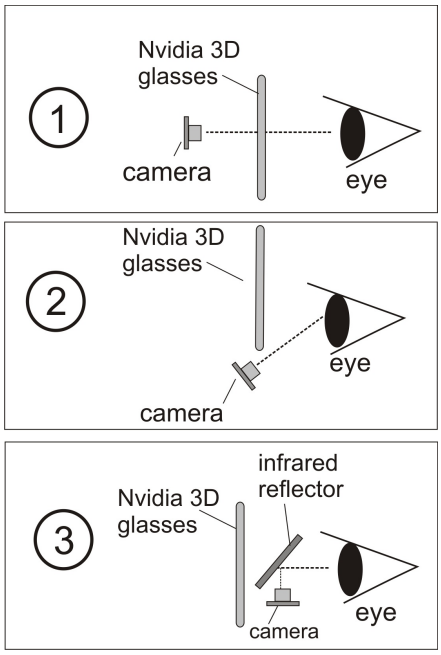


Figure 3.1: Three possible scenarios of Nvidia 3D Vision® glasses and camera installation. The proposed gaze tracking headgear is developed based on the third scenario.

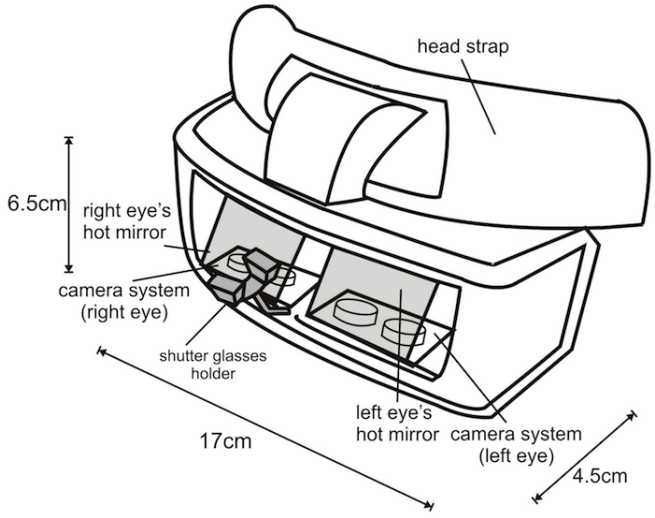


Figure 3.2: Schematic physical construction of gaze tracking headgear.

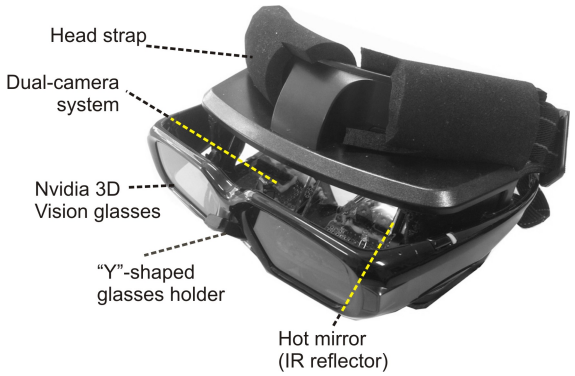


Figure 3.3: Front view of the proposed gaze tracking headgear.

3.1.2 Gaze tracking goggle

The Nvidia 3D Vision® glasses is designed to be used appropriately with prescription glasses. Thus, it has wider width range than normal prescription glasses. I considered the size of the gaze tracking headgear to not exceed the maximum width range of Nvidia 3D Vision® glasses and for such purpose, a modified Famicom gaming goggle (Nintendo Co. Ltd., Kyoto, Japan) was used in this research. Fig.3.2 and Fig.3.3 show schematic physical construction of the gaze tracking headgear and front view of gaze tracking headgear with Nvidia 3D Vision® glasses installed, respectively.

The physical size of Famicom goggle is 17 x 4.5 x 6.5 cm with 365 gram of weight. Front cover and two small LCD displays of the goggle were removed and replaced by a plastic frame to hold hot mirrors. Head strap foam was added to provide comfortable usage for user. Extra surface was added to mount dual-cameras system while maintaining stability, preventing large movements of the goggle during experiment. To hold the Nvidia 3D Vision® glasses properly, a thermoplastic "Y"-shaped shutter glasses holder was designed and added at the middle between left and right hot mirrors.

3.1.3 Dual-camera system

Two dual-camera systems were utilized as main capturing device for left and right eye. Each dual-camera system consists of two mini charge-coupled device (**CCD**) cameras (Analog Technologies, Inc., Santa Clara, United States) as depicted in Fig.3.4. The

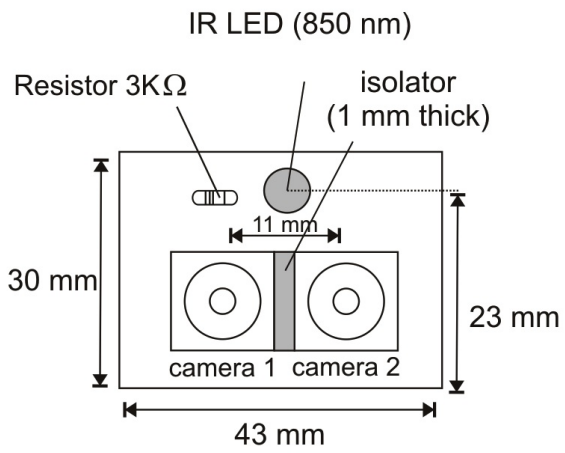


Figure 3.4: Schematic configuration of dual-camera system. Camera 1 and camera 2 are installed parallel to obtain rigorous matching result on stereo pair images.

size of the circuit board is 43 x 30 mm. The usage of two cameras for each eye is to comply with the proposed algorithm for measuring 3D point of gaze, which is mainly based on my previous research work [1]. One of the major problems in multiple vision system is the well-known "matching" problem, which is the determination of conjugate points between stereo pair images. To achieve rigorously matching result on stereo pair images, the two cameras were installed parallel in horizontal direction (Fig.3.4).

The camera size is 10.7 x 10.7 mm. Each camera is equipped by a 1/4" sensor with 400 lines of resolution. The focal length of the lens is 3.9 mm. Despite its size, the camera provides adequate image quality in light sensitivity up to 0.6 lux for image thresholding, edge detection, and ellipse fitting algorithms provided by the gaze tracking software. The camera is ideal for capturing the eye because of its small size and sensitivity to minimum light exposure.

Obtaining high contrast pupil image in visible light is unlikely. Thus, the camera was modified into IR-sensitive camera by removing the IR filters installed inside the camera. I also attached an IR passing filter (Kodak, Ltd. Tokyo, Japan) to the front of the camera lens to block visible light while passing IR light with wavelength longer than 750 nm. An infrared light-emitting diode (**IR LED**) with 850 nm of wavelength was used to create illumination on eye region. To minimize vertical gaze error while allowing the user to gaze at the designated scene, two 50 x 50 mm hot mirrors (Edmund Optics, New Jersey, United

States) were installed such that the hot mirror and the optical axis of the eye formed 45° angle. The thickness of the hot mirror is 3.3 mm, allowing it to be installed neatly in the front of the goggle.

3.1.4 Supporting system

To support real-time gaze tracking while rendering stereoscopic 3D contents, a personal computer with Intel® i7-2600 3.4GHz processor, 4GB memory, Nvidia Quadro® 5000 graphics card, and Windows XP® operating system was used. The gaze tracking headgear was used to capture real-time image sequence with 640 x 480 pixels resolution and 25 Hz sampling rate. An I-O Data LCD-3D231XBR-S monitor with 3D vision compatibility, 23 inches of size, and 1920 x 1080 pixel of screen resolution is used. The gaze tracking software was developed using Microsoft Visual C++ 2010, Open Frameworks C++ toolkit [2], Open Source Computer Vision (**OpenCV**) library [3], and Open Graphics Library (**OpenGL**) [4].

3.2 3D gaze tracking algorithm

3.2.1 Fundamental procedures

Left and right panel of Fig.3.5 show fundamental procedures of 3D gaze tracking and 3D calibration, respectively. Fig.3.6 shows a flowchart of the proposed 3D gaze tracking algorithm that is explained in the following subsections. The eyeball is assumed as a perfect sphere with known radius performing pure rotations around the center of the eyeball [5]. The radius of eyeball is assumed same for left and right eyeball. The proposed optimized geometric method uses user coordinate system. The origin of user coordinate system is located between left and right eyeball (left panel of Fig.3.5). The left and right eyeball are separated by interpupillary distance (IPD). The head of the user is stabilized using chinrest, so that there is no head movement during 3D gaze measurement process.

Before 3D gaze measurement is conducted, several initial parameters are measured (left panel of Fig.3.5): E_{height} (height of eye from tabletop), E_{ipd} (interpupillary distance), E_{rad}

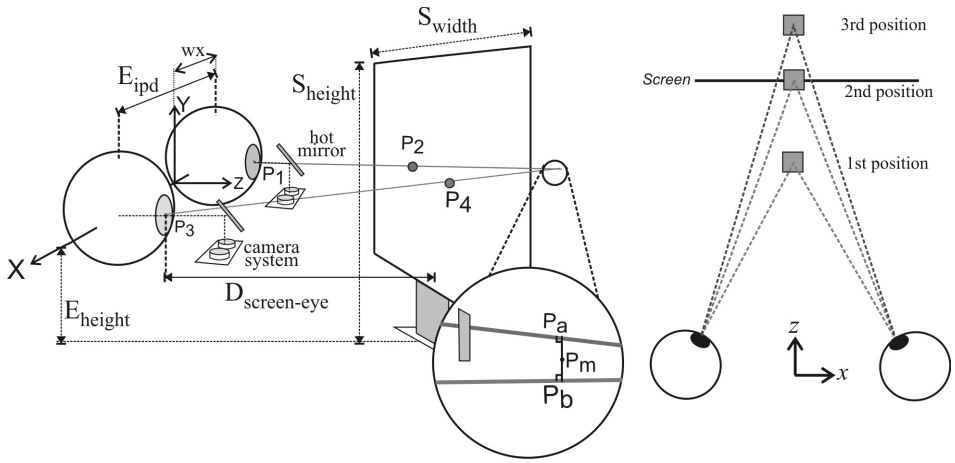


Figure 3.5: Schematic drawing of 3D gaze tracking method (left panel); 3D calibration to improve accuracy in *Z* dimension (right panel).

(radius of eyeball), $D_{screen-eye}$ (distance between screen and eye), S_{height} (height of screen from tabletop), and S_{width} (width of screen). wx is defined as half of E_{ipd}. E_{ipd} and E_{rad} are digitally measured using computer vision technique. Measurement of user dependent parameters is explained further in subsection 3.2.3.

3.2.2 Camera calibration

Camera calibration process is performed for each camera system as explained in previous research work [1]. A chessboard pattern is used to obtain intrinsic, extrinsic, and distortion parameters. Note that camera calibration is conducted only one time before the gaze tracking system is used. Once the intrinsic, extrinsic, and distortion parameters have been obtained, the gaze tracking system is ready to be used for experiment.

3.2.3 Measurement of user dependent parameters

Measurement of eyeball radius and IPD is necessary to estimate 3D pupil position accurately. Dong and Joo [6] found that eyeball and iris diameter were highly correlated. They suggested that eyeball diameter was about 2.11 times larger than iris diameter for normal adult people. Thus, I can estimate eyeball diameter from iris diameter using computer vision technique.

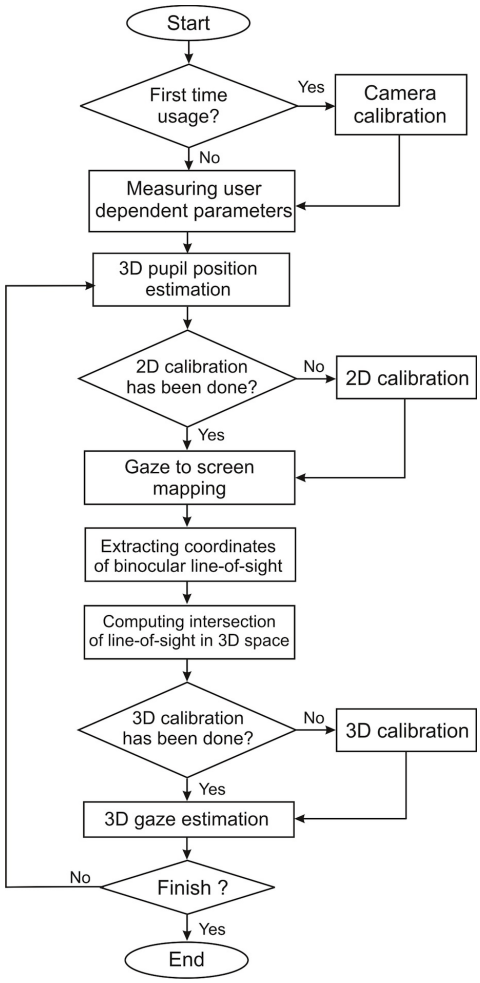


Figure 3.6: Flowchart of the proposed 3D gaze tracking.

Initially, I captures eye images using dual-camera system that related to the measured eye. I use left eye as reference. On each image captured by camera 1 and camera 2, precise pupil position is obtained automatically using thresholding, contour segmentation, and ellipse fitting algorithm. Based on 2D coordinate of the pupil center, I draw a circle manually to fit the edge of the iris. Given eyeball diameter is 2.11 times larger than iris diameter, the software estimates the eyeball diameter automatically.

Since the eyeball is assumed as a perfect sphere, I use a computer vision technique named Direct Linear Transformation (**DLT**) algorithm to measure eyeball diameter in 3D space [1]. On the image captured by camera 1, I mark two diametrically opposed points on eyeball diameter axis as $a(x_1, y_1)$ and $b(x_2, y_2)$ for left and right point, respectively. I apply

similar procedure for image captured by camera 2, resulting $c(x_3, y_3)$ and $d(x_4, y_4)$ for left and right point, respectively.

Next, I obtain 3D coordinate of the left point $A(X_l, Y_l, Z_l)$ from $a(x_1, y_1)$ and $c(x_3, y_3)$ using DLT algorithm. I also obtain the 3D coordinate of the right point $B(X_r, Y_r, Z_r)$ from $b(x_2, y_2)$ and $d(x_4, y_4)$ using DLT algorithm. I then compute the eyeball diameter ($E_{diameter}$) in pixel unit by measuring Euclidean distance of points A and B in 3D space:

$$E_{diameter} = \sqrt{(X_r - X_l)^2 + (Y_r - Y_l)^2 + (Z_r - Z_l)^2} \tag{3.1}$$

The eyeball radius (E_{rad}) is half of $E_{diameter}$.

To measure interpupillary distance (E_{ipd}) between left and right eyeball, I capture left and right eye image from camera 2 of left camera system and camera 1 of right camera system. The captured images are then rectified and merged into one image. On the merged image, left and right pupil position are detected automatically. I measure IPD in pixel unit by measuring Euclidean distance between left pupil (x_l, y_l) and right pupil (x_r, y_r):

$$E_{ipd} = \sqrt{(x_r - x_l)^2 + (y_r - y_l)^2} \tag{3.2}$$

The E_{rad} and E_{ipd} in pixel unit are converted to cm unit using *pixel-to-cm* conversion unit stored in gaze tracking system.

3.2.4 3D pupil position estimation

Given 2D position of left and right pupil, I use dual-camera system and DLT algorithm to estimate the angular position of the eyeball, as explained in [1]. In this research, I consider only horizontal (θ) and vertical (ϕ) angular position. Knowing θ, ϕ, E_{rad}, and E_{ipd}, real time 3D position of left pupil can be computed by the following steps:

First, the distance from center of left eyeball to origin of user coordinate system (wx_{left}) is defined as:

$$wx_{left} = -\left(\frac{E_{ipd}}{2}\right) \tag{3.3}$$

Second, rotational matrices of left eyeball related to user coordinate system are defined.

Horizontal (R_y) and vertical (R_x) rotational matrix in homogeneous format are defined as:

$$R_y = \begin{bmatrix} cos(\theta) & 0 & -sin(\theta) & 0 \\ 0 & 1 & 0 & 0 \\ sin(\theta) & 0 & cos(\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.4}$$

$$R_x = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & cos(\phi) & sin(\phi) & 0 \\ 0 & -sin(\phi) & cos(\phi) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.5}$$

Third, translation from center of left eyeball to origin of user coordinate system (T_1) and translation from origin of user coordinate system to center of left eyeball (T_2) in homogeneous format are defined as:

$$T_1 = \begin{bmatrix} 1 & 0 & 0 & -wx_{left} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.6}$$

$$T_2 = \begin{bmatrix} 1 & 0 & 0 & wx_{left} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.7}$$

Finally, if vector $\tilde{P_1} = [wx_{left}\ 0\ E_{rad}\ 1]^{\mathrm{T}}$ denotes initial 3D position of left pupil, current 3D position of left pupil $P_1(X_1, Y_1, Y_1)$ can be computed by the following equation:

$$P_1 = T_2 * R_x * R_y * T_1 * \tilde{P_1} \tag{3.8}$$

I apply similar computation to estimate 3D position of right pupil $P_3(X_3, Y_3, Y_3)$. Note that the distance from center of right eyeball to origin of user coordinate system (wx_{right}) is

defined as:

$$wx_{right} = \frac{E_{ipd}}{2} \tag{3.9}$$

3.2.5 Gaze to screen mapping using 2D calibration

I ask the user to gaze at several 2D calibration points on screen to obtain gaze fixation on screen. Given n calibration points, I use second order polynomial calibration function to map 2D pupil coordinate of left eye (X_1, Y_1) resulted from 3D pupil position estimation and screen coordinates (X_S, Y_S):

$$X_S = f(X_1, Y_1) = a_0 + a_1X_1 + a_2Y_1 + a_3X_1Y_1 + a_4X_1^2 + a_5Y_1^2 \tag{3.10}$$

$$Y_S = g(X_1, Y_1) = b_0 + b_1X_1 + b_2Y_1 + b_3X_1Y_1 + b_4X_1^2 + b_5Y_1^2 \tag{3.11}$$

Similar computation is implemented for 2D pupil coordinate of the right eye (X_3, Y_3). The left and right gaze on screen after 2D calibration are denoted by (X_L, Y_L) and (X_R, Y_R), respective to the top left corner of the screen.

3.2.6 Extracting coordinate of binocular line-of-sight

The left and right gaze on screen related to user coordinate system can be defined as P_2 and P_4 by the following equations:

$$P_2 = \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} = \begin{bmatrix} X_L - \frac{S_{width}}{2} \\ S_{height} - E_{height} - Y_L \\ D_{screen-eye} + E_{rad} \end{bmatrix} \tag{3.12}$$

$$P_4 = \begin{bmatrix} X_4 \\ Y_4 \\ Z_4 \end{bmatrix} = \begin{bmatrix} X_R - \frac{S_{width}}{2} \\ S_{height} - E_{height} - Y_R \\ D_{screen-eye} + E_{rad} \end{bmatrix} \tag{3.13}$$

Given P_1 and P_3 as 3D position of left and right pupil, P_2 and P_4 as 3D position of left and right gaze, the line-of-sight of left and right eye are formed by P_1P_2 and P_3P_4, respectively (left panel of Fig.3.5).

3.2.7 Intersection of binocular line-of-sight in 3D space

Exact binocular line-of-sight intersection is unlikely. Instead, the nearest point between two line-of-sights is computed. Given two line-of-sights P_1P_2 and P_3P_4 that are joined by shortest line P_aP_b, a fixation point can be computed by finding the midpoint $P_m(X_m, Y_m, Z_m)$ of P_aP_b. Point P_a at line P_1P_2 and point P_b at line P_3P_4 are given by equations:

$$P_a = P_1 + \mu(P_2 - P_1) \tag{3.14}$$

$$P_b = P_3 + \eta(P_4 - P_3) \tag{3.15}$$

The shortest line between two crossing lines can be found by minimizing $|P_b - P_a|$:

$$P_b - P_a = P_3 - P_1 + \eta(P_4 - P_3) - \mu(P_2 - P_1) \tag{3.16}$$

Since P_aP_b is perpendicular to line P_1P_2 and P_3P_4, the results of dot product operation between them are zero:

$$(P_b - P_a).(P_2 - P_1) = 0 \tag{3.17}$$

$$(P_b - P_a).(P_4 - P_3) = 0 \tag{3.18}$$

Substituting Eq.3.16 to Eq.3.17 and Eq.3.18, resulting:

$$[P_3 - P_1 + \eta(P_4 - P_3) - \mu(P_2 - P_1)](P_2 - P_1) = 0 \tag{3.19}$$

$$[P_3 - P_1 + \eta(P_4 - P_3) - \mu(P_2 - P_1)](P_4 - P_3) = 0 \tag{3.20}$$

Computing Eq.3.19 and Eq.3.20 using known X, Y, and Z values of P_1, P_2, P_3, and P_4, μ, η, P_a, and P_b can be computed. The middle point of P_aP_b, is computed as follows:

$$P_m = \frac{P_a + P_b}{2} \tag{3.21}$$

3.2.8 3D gaze estimation using 3D calibration

To improve accuracy of 3D point of gaze calculation, the Z element of P_m can be corrected using 3D calibration. I use three calibration points positioned at different depths, as shown in right panel of Fig.3.5. If s_z is depth of the calibration points and z is the Z_m of 3D point of gaze P_m, polynomial mapping between s_z and z is described as follows:

$$s_z = a_0 + a_1 z + a_2 z^2 \tag{3.22}$$

or expressed as matrices:

$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 1 & z_1 & z_1^2 \\ 1 & z_2 & z_2^2 \\ 1 & z_3 & z_3^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = Ax \tag{3.23}$$

To solve vector $x = [a_0\ a_1\ a_2]^{\mathrm{T}}$ in Eq.3.23, Singular Value Decomposition (**SVD**) is used. If z_{pre} and z_{after} denote Z_m before and after correction, the following equation is used to improve the accuracy of Z_m:

$$z_{after} = a_0 + a_1 z_{pre} + a_2 z_{pre}^2 \tag{3.24}$$

The effectiveness of the proposed 3D calibration method has been simulated and validated experimentally in previous research work [7].

3.3 Experimental validation and implementation

3.3.1 Stereoscopic 3D rendering

Stereoscopic 3D is implemented to give depth impression in contents of experiment. Rendering of left and right scene should be formatted such that left-eye view and right-eye view can be separated by the graphics card system. Such rendering technique is widely known as quad-buffering [8]. Quad-buffering allows OpenGL program to render and swap buffers for each eye. OpenGL utilizes back and front buffers for right and left sides such

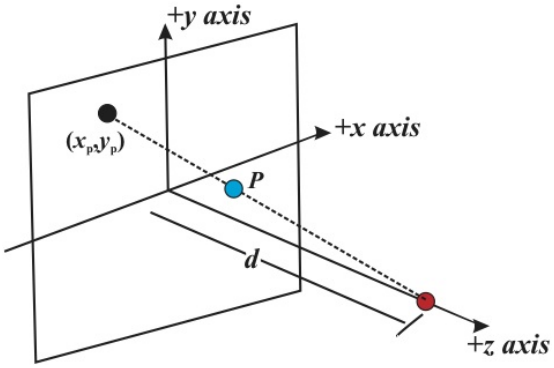


Figure 3.7: Basic perspective projection

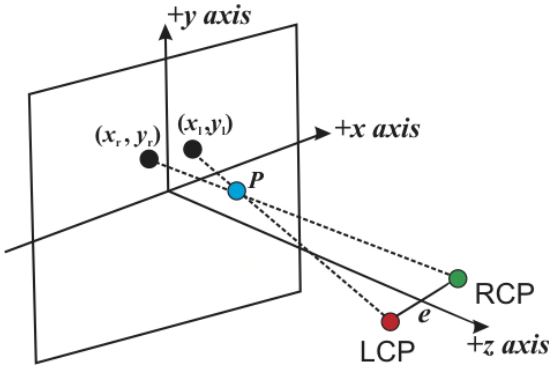


Figure 3.8: Off-axis projection

that there are four buffers in total. The usage of four buffers produces a smooth picture for each eye. Each left or right set of buffers can only be viewed through the corresponding eye by Nvidia 3D Vision® glasses.

Fig.3.7 shows basic concept perspective projection for non-stereoscopic viewing. In this research, I use right-handed coordinate system. The center of projection lies on the Z-axis at (0, 0, d). The resulting projection of a 3D object $P = (x, y, z)$ onto the projection plane is a 2D object with coordinates (x_p, y_p), described by the following equation:

$$x_p = \frac{xd}{d - z}\ ,\ y_p = \frac{yd}{d - z} \tag{3.25}$$

To produce stereoscopic images with lower discomfort level, off-axis projection is implemented (Fig.3.8). Left-eye view and right-eye view are produced by left-camera projection (**LCP**) and right-camera projection (**RCP**), respectively. LCP and RCP are horizon-

tally separated by e. For LCP, instead of offsetting the camera to the left, the entire scene is moved to the right. The 2D coordinates of the projected left-eye view can be described by:

$$x_l = \frac{(x + \frac{e}{2})d}{d - z} \ , \ y_l = \frac{yd}{d - z} \tag{3.26}$$

For the RCP, the entire screen is moved to the left, so that the coordinates of the projected right-eye view can be described by:

$$x_r = \frac{(x - \frac{e}{2})d}{d - z} \ , \ y_r = \frac{yd}{d - z} \tag{3.27}$$

In stereoscopic 3D rendering, there are three kinds of parallax that affect the position of virtual 3D object. *Negative parallax* causes virtual 3D object to be positioned in front of screen. *Positive parallax* causes the virtual 3D object to be located behind the screen. *Zero parallax* is achieved when the virtual 3D object is perfectly projected at the screen.

Recalling Eq.3.26 and Eq.3.27, arbitrary 3D objects will be projected with negative parallax, which falls in front of screen. To create well-balanced stereoscopic images, both negative and positive parallax should be used appropriately. To fix this problem, the projected object of LCP can be shifted leftward while the projected object of RCP can be shifted rightward. If the projected object is shifted by the same amount of the original camera offset, the resulting geometry will be at the original projection plane, precisely at zero parallax. By implementing this correction, scene element originally placed in front of the projection plane will be projected to negative parallax while scene element originally placed behind the projection plane will be projected to positive parallax.

The corrected equations for the projected left-eye view (x_l, y_l) and right-eye view (x_r, y_r) can be described as:

$$x_l = \frac{(x + \frac{e}{2})d}{d - z} - \frac{e}{2} \ , \ y_l = \frac{yd}{d - z} \tag{3.28}$$

$$x_r = \frac{(x - \frac{e}{2})d}{d - z} + \frac{e}{2} \ , \ y_r = \frac{yd}{d - z} \tag{3.29}$$

Algorithm 1 shows implementation of quad-buffer stereoscopic 3D rendering using off-axis projection. Instead of rendering the 3D object image on a single buffer, the program renders

two images on two buffers. Note that both left and right buffers are swapped after the 3D object has been rendered. That is to say, the left buffer is swapped with another buffer on the left and the right buffer is swapped with another buffer on the right. Left and right buffer can only be seen by their corresponding lenses in the Nvidia 3D Vision® glasses. Because of the slightly different viewpoints from each eye, an illusion of depth is achieved and thus the rendering in stereo is complete.

update projection matrix;
clear drawing buffers and z-buffers;
set background color;
if *rendering in stereo* **then**
    select left buffers;
    update left-eye projection;
    render 3D object;
    select right buffers;
    update right-eye projection;
    render 3D object;
end
forall the *buffer in buffers* **do**
    swap buffer;
end

Algorithm 1: Quad-buffer stereoscopic 3D rendering algorithm

3.3.2 Experiment procedure

The first experiment was designed to validate accuracy of the proposed system. Ten university students (8 M, 2 F; ages 21–26) participated in the study, recruited verbally on a volunteer basis. Before conducting experiment, I explained the procedure of experiment, including duration of experiment and how to perform the designated tasks. All participants had agreed with the procedure of experiment and had stated their current physical condition in written questionnaire. Only three participants wore contact lenses, while the rest wore neither contact lenses nor prescription glasses. Seven participants had previous experience with stereoscopic 3D contents. Prior to experiment, all participants were screened for depth perception by showing stereoscopic image of a 3D rotating cube. All participants were able to see the stereoscopic image correctly.

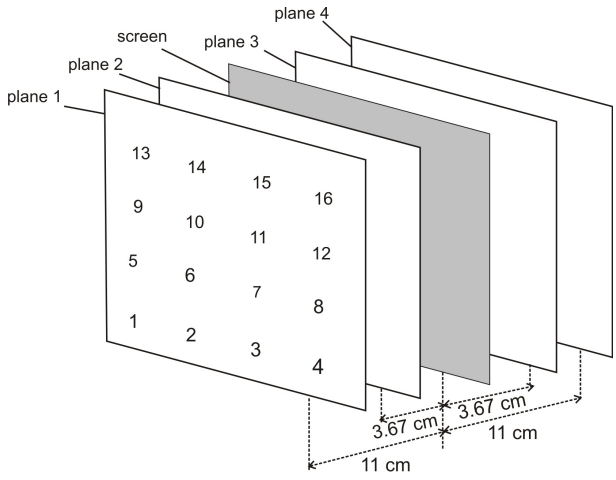


Figure 3.9: Four virtual planes in the first experiment. Plane 1 is the closest virtual plane to participant. Sixteen validation targets, denoted by number 1 to 16, were shown alternately in each virtual plane.

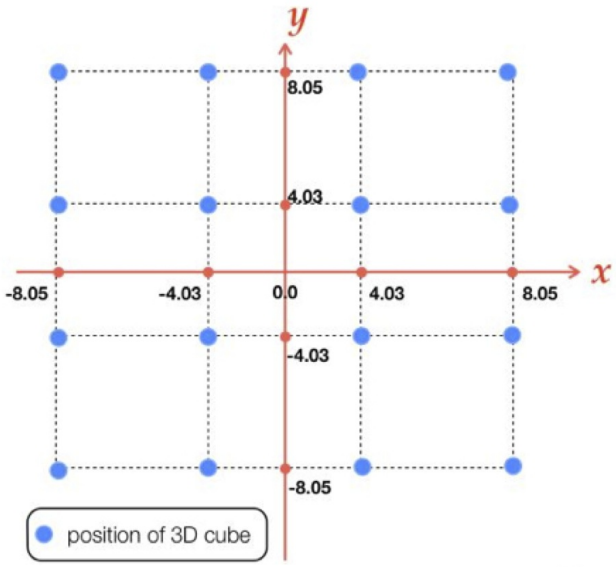


Figure 3.10: Layout of sixteen validation targets (in cm)

The participant was positioned such that the center of the screen was vertically and horizontally aligned at the middle point of both eyes. The distance between participant and screen ($D_{screen-eye}$) was set to 63 cm. The height of eye from tabletop (E_{height}) was set to 32.2 cm. The height of screen from tabletop (S_{height}) was 47 cm. The width of screen (S_{width}) was 51.5 cm. The eyeball radius (E_{rad}) and interpupillary distance (E_{ipd})

were measured digitally and saved to gaze tracking database. The averages interpupillary distance and eyeball radius of ten participants were 6.6 cm and 1.28 cm, respectively.

Next, a quick 2D gaze calibration procedure was conducted. The participant was asked to gaze at 9 calibration points in 27 x 27 cm sized square. Following the 2D gaze calibration, the 3D gaze calibration was performed. A 3.6 x 3.6 x 3.6 cm 3D cube was used as calibration target. The 3D cube was rendered stereoscopically and positioned at the center of the screen. The depth of the cube were 4 cm in front of, 0 cm from, and 4 cm behind the screen. The participant was asked to gaze at the cube and confirm their 3D gaze position by pressing space bar.

Afterwards, 3D gaze validation session using same 3D cube was carried out. Four virtual planes were provided and positioned at 11 cm and 3.67 cm in front of, and 3.67 cm and 11 cm behind the screen as shown in Fig.3.9 [9]. The sequence of the virtual planes was 1, 2, 3, and 4 as seen from participant point of view with plane 1 as the closest plane from the participant. Sixteen targets on each of four virtual planes were presented to the participants. Spatial layout of these sixteen targets is shown in Fig.3.10. The positions of these 64 points were not identical to any of 9 points in 2D gaze calibration, nor 3 points in 3D gaze calibration. The participant was asked to gaze at each of 64 validation points while pressing space bar to confirm the 3D gaze position. The average values of X, Y, and Z gaze position at each validation point were computed from five consecutive gaze data.

To compare the accuracy of the proposed system, I also recorded the result of conventional geometric method simultaneously in 3D gaze validation session. For conventional geometric method, the interpupillary distance and the eyeball radius was fixed using standard parameters: 6.5 cm [10] and 1.35 cm [11], respectively. The first experiment took about 5 minutes of time for each participant. Three minutes break was given before conducting the second experiment.

In the second experiment, the proposed system was implemented as a measurement tool for depth gaze in stereoscopic display. The accuracy of the proposed gaze tracking system in measuring depth of virtual 3D object was compared with user depth perception. The participant was positioned in the same manner as in the first experiment. Since user dependent parameters had been previously measured and saved in the gaze tracking database, only 2D and 3D calibration sessions were needed before doing experiment. To evaluate user depth

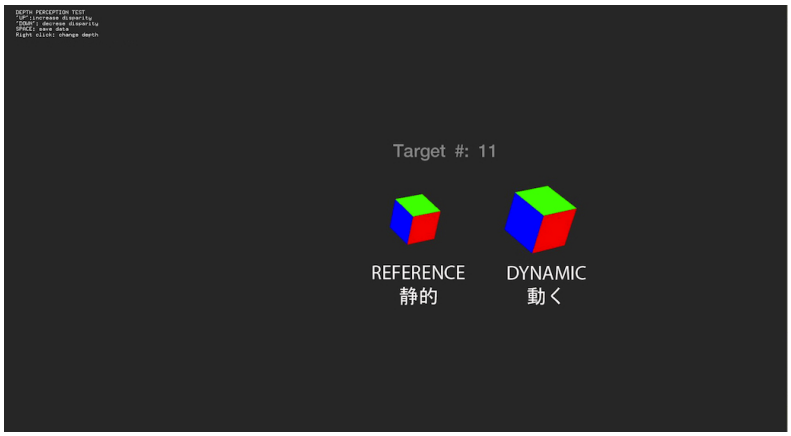


Figure 3.11: Screen capture of the third stage (R<D) in the second experiment.

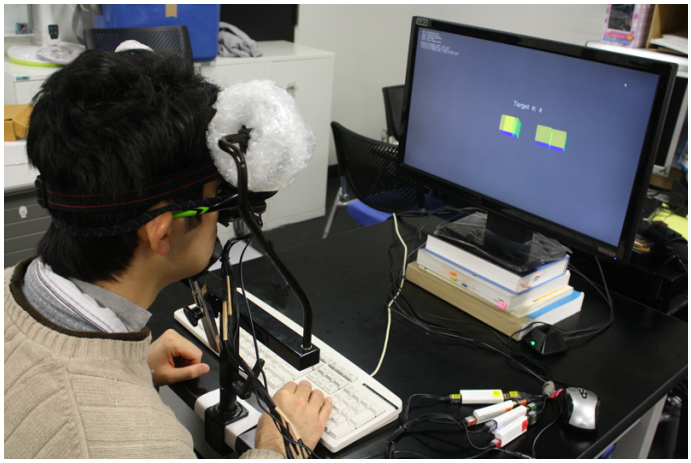


Figure 3.12: A participant adjusted the dynamic cube to the same depth as depth of reference cube while the gaze tracking system recorded Z position of participant's gaze.

perception, an interactive application for subjective depth judgment was developed.

In this interactive application, one reference cube (R) and one dynamic cube (D) were rendered on dark background. The dynamic cube was positioned on the right side of the reference cube, as shown in Fig.3.11. The application consisted of three stages. In the first stage, size of reference cube was equal to size of dynamic cube (R=D). In the second stage, size of reference cube was bigger than size of dynamic cube (R>D). In the third stage, size of reference cube was smaller than size of dynamic cube (R<D). Fig.3.11 shows a screen capture of the third stage where reference cube was smaller than dynamic cube (R<D).

In each stage, the reference cube was placed at five different virtual planes: 20 cm and 10 cm in front of, 0 cm from, 10 cm and 20 cm behind the screen. Fig.3.12 shows configuration of the second experiment. The participant was then asked to adjust position of dynamic cube to the same depth as depth of reference cube. Depth of dynamic cube was controlled using arrows keypad (UP induced farther distance, DOWN induced closer distance). Once the participant perceived that the depth of reference and dynamic cube was equal, the participant was asked to keep gazing at reference cube while pressing space bar to save Z position of 3D point of gaze and depth of both cubes. To change the depth of reference cube in each stage, the participant had to press the left button of the mouse. Duration of the second experiment was about 12 minutes for each participant.

3.4 Experimental results

In the first experiment, I evaluate effect of measurement methods and virtual plane positions to measurement error using two-way analysis of variance (**ANOVA**). The two independent variables are measurement methods (conventional vs.optimized geometric method) and virtual plane positions (11 cm and 3.67 cm in front of screen, 3.67 cm and 11 cm behind the screen). Measurement error, defined as error in Euclidean distance between virtual 3D object and 3D position of user gaze, is the dependent variable.

For the first independent variable, statistical analysis of experimental results shows that both measurement methods (conventional vs.optimized geometric method) yield result with different accuracy. Effect of measurement methods is significant ($F(1,9) = 282.95$, $p<0.001$). Comparison of average errors between conventional and optimized geometric method is shown in Table 3.1. The optimized rather than conventional geometric method shows better performance in X (mean = 0.83 cm, SD = 0.57 cm), Y (mean = 0.87 cm, SD = 0.56 cm), and Z (mean = 1.06 cm, SD = 0.70 cm) dimension.

For the second independent variable, experimental results show that effect virtual plane positions is significant ($F(3,27) = 65.53$, $p<0.001$). Furthermore, interaction between measurement methods and virtual plane positions is also significant ($F(3,27) = 6.51$, $p<0.005$). As the distance of virtual 3D object increases, average errors of conventional and optimized

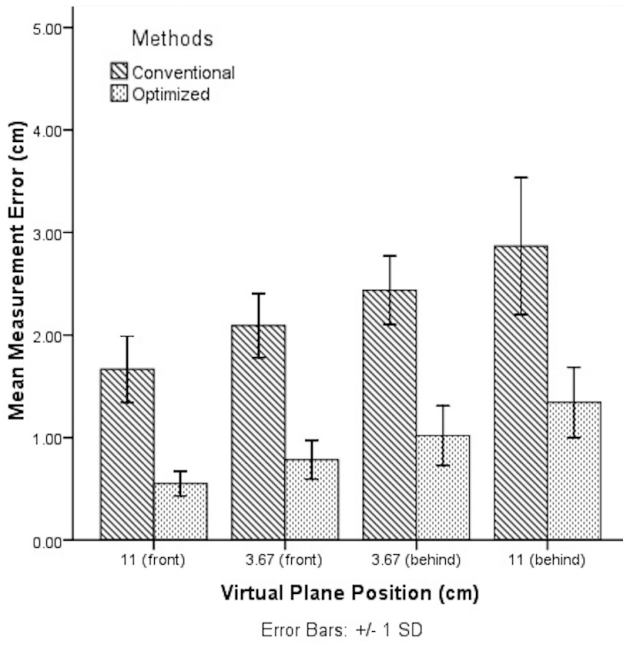


Figure 3.13: Comparison of average errors between conventional and optimized geometric methods in all virtual planes.

geometric method also increase (Fig.3.13). However, unvarying smaller statistical errors of the proposed method is observed. For example in plane 1, average errors of optimized and conventional geometric method are 0.55 (SD = 0.41) cm and 1.66 (SD = 0.81) cm, respectively. Compared with conventional geometric method, the proposed method is able to reduce average errors in each virtual plane.

In the second experiment, I evaluate the experimental results using three-way ANOVA. The three independent variables are measurement methods (gaze tracking vs. user depth perception), positions of reference cube's virtual plane (20 cm and 10 cm in front of screen, 0 cm from screen, 10 cm and 20 cm behind the screen), and sizes of cubes (R=D, R>D,

Table 3.1: Comparison of average errors between conventional and optimized geometric method for individual coordinates (in cm).

Methods	**Dimensions**		
	X	*Y*	*Z*
Conventional geometry	2.13 ± 0.96	2.17 ± 0.99	2.51 ± 1.07
Optimized geometry	0.83 ± 0.57	0.87 ± 0.56	1.06 ± 0.70

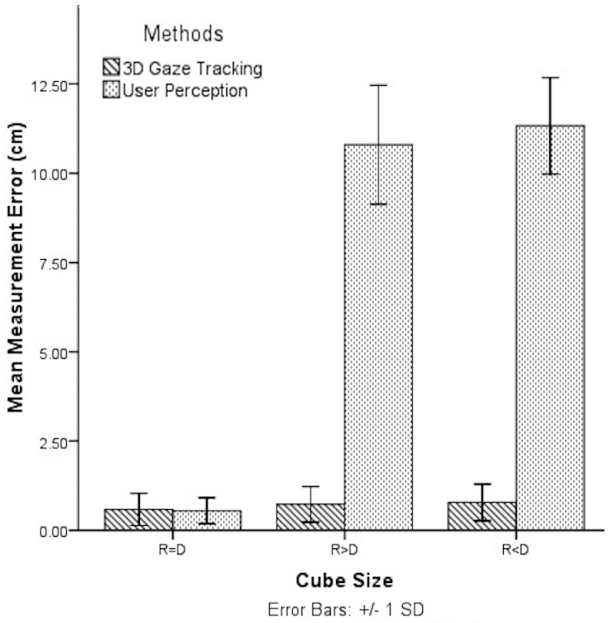


Figure 3.14: Comparison of average errors between 3D gaze tracking and user depth perception based on size of reference (R) and dynamic (D) cubes.

Table 3.2: Comparison of average errors in depth measurement between 3D gaze tracking and user depth perception for different sizes of cubes (in cm).

Methods	**Size of cubes**		
	$R = D$	$R > D$	$R < D$
3D gaze tracking	0.58 ± 0.45	0.73 ± 0.51	0.78 ± 0.52
User depth perception	0.55 ± 0.36	10.80 ± 1.66	11.33 ± 1.35

and R<D). The dependent variable is average depth measurement errors.

The analysis reveals that effect of size cue on user depth perception is significant ($F(2,120) = 2701.94$, $p<0.001$). As depicted in Fig.3.14, users tended to make high errors in adjusting dynamic cube when sizes of reference and dynamic cube were different. From experimental data, I find that dynamic cube adjustment errors are high with positive sign when size of reference cube is smaller than size of dynamic cube (R<D), while these errors are also high with negative sign when size of reference cube is bigger than dynamic cube (R>D).

On the contrary, errors of gaze tracking remain small although the sizes of cube are changed. Table 3.2 shows comparison of average errors between gaze tracking and user

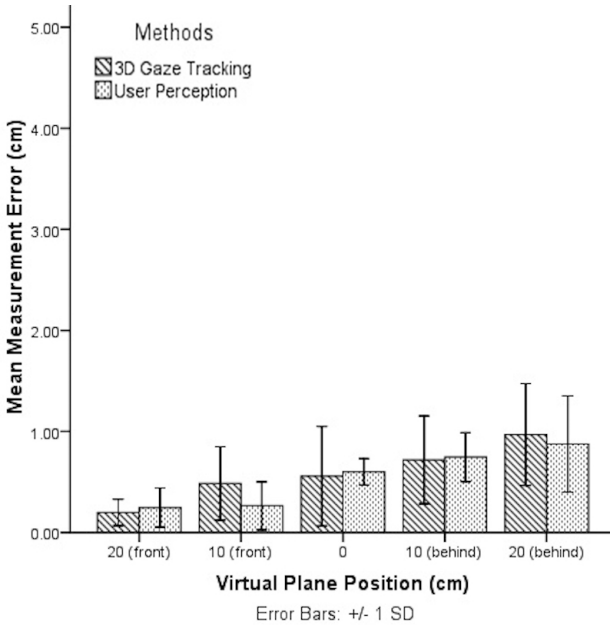


Figure 3.15: Average errors of depth measurement between 3D gaze tracking and user depth perception in all virtual planes (first stage; R=D).

depth perception for different sizes of cubes. Compared with user depth perception, the gaze tracking system is more robust to error regardless the sizes of cubes. I also compare average errors between gaze tracking and user depth perception in each virtual plane of the first stage (R=D). Fig.3.15 shows that accuracy of user depth perception and gaze tracking slightly decrease as the distance of virtual plane increases.

3.5 Discussion

In the first experiment, results have demonstrated that the proposed method is superior than conventional geometric method. Conventional geometric method is generally designed to fit average user dependent parameters, while I provide customized user dependent parameters and 3D calibration that lead to improvement of 3D gaze tracking accuracy. Errors of 3D gaze measurement in both methods are due to several assumptions of human eye, such as fix center of eyeball with no translation and simplification of geometric shape of eyeball.

Observing Table 3.1, the experimental results show that horizontal (X) gaze position is more accurate than vertical (Y) position in both conventional and optimized geometric method. This results may correlate with the nature of persistence fixation eye movement in horizontal rather than vertical direction. Moshel *et al.* [12] found that microsaccades enhance the persistence of horizontal fixation in short time presentation of stimuli, by which stereoscopic viewing is achieved. These microsaccades are not found in vertical gaze fixation.

During fixation at one attention point, human gaze should be stabilized. However, visual perception fades as a consequence of neural adaptation [13]. Thus, to keep stimulating cells in the fovea and counteract neural adaptation, human's eye performs very small eye movements during fixation. There are three kinds of small eye movements during fixation: tremor (amplitude: <0.01°), drift (amplitude: 0.01–1°), and microsaccades (amplitude: 0.16–0.67°) [14]. Those eye movements, particularly microsaccades, maintain the position of horizontal gaze fixation as well as stimulating fovea cells to avoid fading. On the other side, average errors of Z position in conventional and optimized geometric methods are higher than errors of X and Y position because Z position relies on accuracy of 2D gaze position.

The accuracy of 3D point of gaze calculation decreases as the distance of the virtual 3D object increases in both conventional and geometric method. The explanation is related with the magnitude of vergence angle. It is said that vergence angle is effective cues of 3D gaze within one meter depth range [15]. The change of vergence angle to adjust binocular gaze is greater for closer rather than farther virtual 3D object. Thus, precision of the gaze tracking system is higher for shorter depth range.

In the second experiment, size cue affects user depth perception significantly. Participants unconsciously moved dynamic cube based on size cue instead of binocular disparity cue. Participants also perceived smaller closer cube as farther away than more distant larger cube. This finding confirms previous research that proposed size-distance paradox theory in reduced cue conditions [16]. From Fig.3.15, one can see that when the size of objects is similar, participant made more mistakes in adjusting dynamic cube if the position of virtual plane is located further from the participant. These results suggest that effective distance perception is achieved by placing virtual 3D object closer to the participant.

Table 3.3: Comparison of time allocation between conventional and optimized geometric method (in milliseconds and seconds).

Sessions	**Methods**	
	Conventional	Optimized
Frame grabbing and pupil tracking	15.80 ms	15.80 ms
3D pupil position estimation	-	21.24 ms
3D point of gaze estimation	2.01 ms	3.12 ms
2D calibration using 9 target points	27 s	27 s
3D calibration using 3 target points	-	10.2 s

On the other hand, results have shown that depth perception of participants did not affect 3D gaze tracking. Instead of relying on size cue, the gaze tracking system used binocular disparity cue that induced vergence eye movements. This finding differs from previous research that proposed user depth perception influences vergence response [17]. Experimental results show that monocular cue (size cue) and binocular cue (disparity cue) contribute differently to 3D gaze and user depth perception.

Even though the proposed method shows good performance, the head movements of participants are still restricted by chinrest for experiment purpose. This restriction is not suitable if the gaze tracking system is used in virtual environment that requires large head movements, such as controlling virtual 3D object using head and eye position. Furthermore, the proposed method requires slightly longer computational time due to additional steps of 3D pupil position estimation and 3D calibration (Table 3.3.). In future, it may be possible to integrate the gaze tracking system with 3D motion sensor or depth camera to obtain 3D position of participant's head during 3D gaze tracking in virtual space.

Observing Fig.3.15, the error pattern indicates that an automatic gaze-to-perception conversion could be reliably developed, which is an important topic for research in stereoscopic 3D contents and human-computer interaction. In future, development of better user interface and stereoscopic 3D contents may be conducted by considering how user perceives depth of virtual 3D object at five different distances: within 10–20 cm in front of the screen, within 3–10 cm in front of the screen, within 3 cm of the screen, within 3–10 cm behind the screen, and within 10–20 cm behind the screen. Combining objective and subjective measurement will lead to development of more user-friendly 3D technology that considers human factors.

Bibliography

[1] S. Wibirama, S. Tungjitkusolmun, and C. Pintavirooj, "Dual-Camera Acquisition for Accurate Measurement of Three-Dimensional Eye Movements," *IEEJ Transactions on Electrical and Electronic Engineering*, vol. 8, no. 3, pp. 238–246, 2013.

[2] J. Noble, *Programming Interactivity, 2nd Edition*. O'Reilly Media, 2012.

[3] G. Bradski and A. Kaehler, *Learning OpenCV: Computer Vision with the OpenCV Library*. California: O'Reilly Media, Inc., 1st ed., 2008.

[4] Khronos, "OpenGL: The Industry Standard for High Performance Graphics," 22 January 2010.

[5] S. T. Moore, T. Haslwanter, I. S. Curthoys, and S. T. Smith, "A geometric basis for measurement of three-dimensional eye position using image processing," *Vision Research*, vol. 36, no. 3, pp. 445–459, 1996.

[6] E. Y. Dong and C. K. Joo, "Predictability for proper capsular tension ring size and intraocular lens size," *Korean J Ophthalmol*, vol. 15, no. 1, pp. 22–26, 2001.

[7] S. Wibirama and K. Hamamoto, "Error Correction in Geometric Method of 3D Gaze Measurement using Singular Value Decomposition," in *Proceeding of The 5th Indonesia Japan Joint Scientific Symposium*, pp. 554–558, 2012.

[8] R. Belleman, B. Stolk, and R. de Vries, "Immersive virtual reality on commodity hardware," in *Proceedings of the 7th annual conference of the Advanced School for Computing and Imaging*, vol. 7, pp. 297–304, 2001.

[9] K. Essig, M. Pomplun, and H. Ritter, "A neural network for 3D gaze recording with binocular eye trackers," *The International Journal of Parallel, Emergent and Distributed Systems*, vol. 21, no. 2, pp. 79–95, 2006.

[10] C. Jinjakam and K. Hamamoto, "Study on Parallax Affect on Simulator Sickness in One-screen and Three-screen Immersive Virtual Environment," *Proceedings of The*

School of Information and Telecommunication Engineering Tokai University, vol. 4, no. 1, pp. 34–39, 2011.

[11] R. Park and G. Park, "The center of ocular rotation in the horizontal plane," *American Journal of Physiology–Legacy Content*, vol. 104, no. 3, pp. 545–552, 1933.

[12] S. Moshel, A. Z. Zivotofsky, L. Jin-Rong, R. Engbert, J. Kurths, R. Kliegl, and S. Havlin, "Persistence and phase synchronisation properties of fixational eye movements," vol. 161, no. 1, pp. 207–223–, 2008.

[13] L. A. Riggs, F. Ratliff, J. C. Cornsweet, and T. N. Cornsweet, "The Disappearance of Steadily Fixated Visual Test Objects," *Journal of Optical Society America*, vol. 43, no. 6, pp. 495–500, 1953.

[14] K. Holmqvist, M. Nystrom, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press, 2011.

[15] I. P. Howard and B. J. Rogers, *Geometry of Binocular Images*, vol. 2, pp. 1–9. Toronto: I Porteous, 2002.

[16] S. K. Fisher and K. J. Ciuffreda, "Accommodation and apparent distance," *Perception*, vol. 17, no. 5, pp. 609–621, 1988.

[17] B. M. Sheliga and F. A. Miles, "Perception can influence the vergence responses associated with open-loop gaze shifts in 3D," *Journal of Vision*, vol. 3, no. 11, 2003.

Chapter 4

The Relationship among Motion Sickness, 3D Gaze Position, and Heart Rate Variability

4.1 Simulator sickness questionnaire (SSQ)

One of well-known indicator methodologies for visually induced motion sickness (**VIMS**) is self report, i.e. using a written questionnaire to investigate subjective opinion of sickness symptoms. The aims of simulator sickness questionnaire (**SSQ**) are: (1) providing valid index of overall visually induced motion sickness severity; (2) providing detail sub-scale of the total sickness score in which effect of particular simulator to the locus of sickness can be diagnosed; (3) providing a scoring approach for straightforward monitoring and cumulative tracking.

Kennedy and Lane [1] had developed SSQ based on pensacola motion sickness questionnaire (**MSQ**) by removing irrelevant questions (see *Appendix A*). The development of SSQ revealed that sickness symptoms was classified into three general types of effect or sub-scale named: (1) *nausea* (*N*); (2) *oculomotor* (*O*); (3) *disorientation* (*D*).

Scores on the nausea effect are based on symptoms that are related to gastrointestinal distress, such as stomach awareness, salivation, and burping. Scores on the oculomotor effect are based on symptoms that are related to visual disturbances, such as eyestrain,

■次の中から最もあてはまる所に○をつけて下さい。

(Please circle ○ the most appropriate score according to your feeling.)

質問(Question)	答え(Answer)			
	なし	ややある	ある	かなりある
	No	Slight	Moderate	Severe
一般的な不快感 (General discomfort)	0	1	2	3
疲労感がある (Fatigue)	0	1	2	3
頭痛がする (Headache)	0	1	2	3
眼が疲れている (Eyestrain)	0	1	2	3
眼の焦点がぼける (Difficulty focusing)	0	1	2	3
唾液の増加 (Increased salivation)	0	1	2	3
発汗する (Sweating)	0	1	2	3
吐き気がする (Nausea)	0	1	2	3
集中できない (Difficulty concentrating)	0	1	2	3
頭が重い (Fullness of head)	0	1	2	3
眼がかすむ (Blurred vision)	0	1	2	3
眩暈感がある[開眼] (Dizzy (eyes open))	0	1	2	3
眩暈感がある[閉眼] (Dizzy (eyes closed))	0	1	2	3
周囲が回転する眩暈 (Vertigo)	0	1	2	3
胃の存在感がある (Stomach awareness)	0	1	2	3
げっぷが出る (Burping)	0	1	2	3

Figure 4.1: SSQ form used in this research: explained in both japanese and english.

orbital pain, and heavy eyelids, difficulty focusing, blurred vision, and headache. Scores on disorientation effect are related to vestibular disturbances (e.g. dizziness and vertigo).

The scoring procedures presume that all individuals are in healthy condition. In this research, the measured SSQ scores were obtained from post-experiment since reports of SSQ after exposure of stereoscopic 3D content had been known to be inflated if both pre- and post-experiment questionnaire were given [2].

To evaluate VIMS by the SSQ, each participant is asked to fill in a SSQ form containing 16 symptoms with four points scale (0, 1, 2, and 3), as shown in Fig. 4.1. Zero (0) indicates no effect, one (1) indicates slight effect, two (2) indicates moderate effect, and three (3) indicates severe effect of stereoscopic 3D contents exposure. The computation of nausea, oculomotor, and disorientation effect from all symptoms are based on weight (1 and 0) contained by each symptom, as depicted in Table 4.1 [3].

The nausea effect (N), oculomotor effect (O), disorientation effect (D), and total effect (T) can be computed by the following equations:

Table 4.1: Computation and weighting of SSQ symptoms

SSQ symptom	**Score (0–3)**	**Weight**		
	x_i	N_i	O_i	D_i
General discomfort		1	1	0
Fatigue		0	1	0
Headache		0	1	0
Eyestrain		0	1	0
Difficulty focusing		0	1	1
Increased salivation		1	0	0
Sweating		1	0	0
Nausea		1	0	1
Difficulty concentrating		1	1	0
Fullness of head		0	0	1
Blurred vision		0	1	1
Dizzy (eyes open)		0	0	1
Dizzy (eyes closed)		0	0	1
Vertigo		0	0	1
Stomach awareness		1	0	0
Burping		1	0	0

$$N_s = \alpha_N \sum_{i=1}^{16} N_i x_i \tag{4.1}$$

$$O_s = \alpha_O \sum_{i=1}^{16} O_i x_i \tag{4.2}$$

$$D_s = \alpha_D \sum_{i=1}^{16} D_i x_i \tag{4.3}$$

$$T_s = \alpha_T \sum_{i=1}^{16} (N_i + O_i + D_i)\, x_i \tag{4.4}$$

with:

$$\alpha_N = 9.54; \alpha_O = 7.58; \alpha_D = 13.92; \alpha_T = 3.74 \tag{4.5}$$

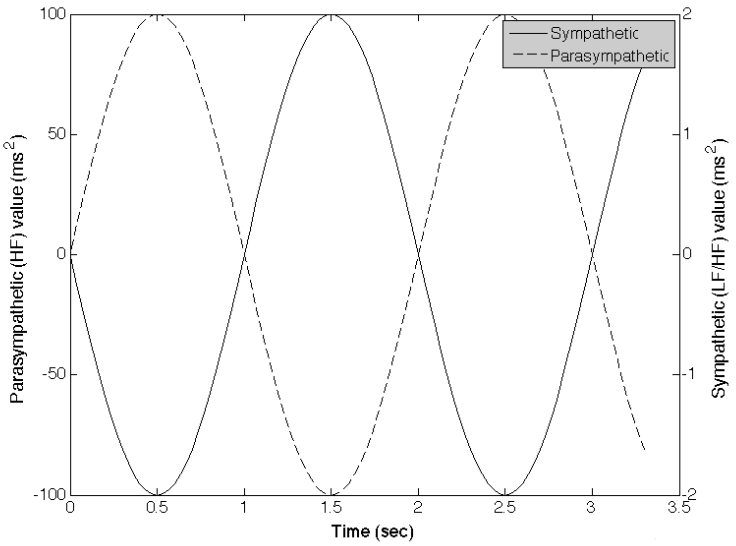


Figure 4.2: Ideal balance between sympathetic and parasympathetic nerves activity (*homeostasis*). In heart rate variability analysis, sympathetic and parasympathetic nerves are commonly observed by investigating graph of LF/HF ratio (sympathetic) and HF (parasympathetic), respectively (LF: low frequency, HF: high frequency).

4.2 Heart rate variability

Electrocardiography (**ECG**) is a method to measure and record heartbeats through attached electrodes on the chest. ECG is commonly used to investigate cardiovascular (heart-related) diseases. One of the main clinical scenarios where ECG has been found valuable includes the risk stratification of sudden cardiac death. Besides these main clinical scenarios, ECG has been studied with relation to physical exercise, occupational and psychosocial stress, gender, age, drugs, alcohol, smoking and sleep. ECG data has been used to observe the occurrence of VIMS symptoms in virtual reality-related research.

Heart rate variability (**HRV**) is variations between subsequence heartbeats. In human, the rhythm of the heart is controlled by the *sinoatrial* (**SA**) node, which is modulated by both *sympathetic* and *parasympathetic* nerves of the autonomic nervous system. Sympathetic nerves increase heart rate while parasympathetic nerves decrease heart rate. The continuous modulation of sympathetic and parasympathetic nerves result in variations of heart rate [4, 5].

The condition of HRV is commonly used as an accurate indicator of health, regularity or coherence of human body, as well as balance in autonomic nervous system. A human is expected to have harmonious condition between sympathetic and parasympathetic nerves activity in normal and healthy condition (*homeostasis*). The balance mechanism of sympathetic and parasympathetic nerves to govern cardiopulminary function is almost similar as two opposite sine waves, as depicted in Fig.4.2 [6]. On the other side, when a human is aroused, doing physical activities that need more oxygen consumption, or being mentally stressed, the ideal balance of sympathetic and parasympathetic nerves maybe changed. The sympathetic nerves are generally found to be more dominant than parasympathetic nerves. In this case, both sine waves in Fig.4.2 will have similar phase with same direction in Y-axis although the amplitudes are different.

The most distinguishable periodic component of HRV is called *respiratory sinus arrhythmia* (**RSA**). RSA ranges in high frequency (HF) from 0.15 to 0.4 Hz. This HF component of HRV is tightly related with respiration process and is argued as strong indicator of parasympathetic nerves activity. Other widely observed components of HRV is called low frequency (LF), very low frequency (VLF), and ultra low frequency (ULF), which are ranging in frequency bands of 0.04–0.15 Hz, 0.003–0.04 Hz, and 0–0.003 Hz, respectively. However, in short-term observation, ULF band is generally omitted [4].

To obtain ECG data, portable ECG system with wireless receiver and transmitter (WEB-5500; Nihon Koden Co., Tokyo) was used. I then analyzed raw ECG data with ECG analyzer provided by the system, as shown in Fig.4.3. I investigated oscillation of both HF and LF/HF values to detect increment of sympathetic nerves activity. My hypothesis is LF/HF and HF graph will be at the same phase and direction in Y-axis when the participant experiences symptoms of VIMS. By detecting this feature, I then observed scene components of 3D movie that induce VIMS.

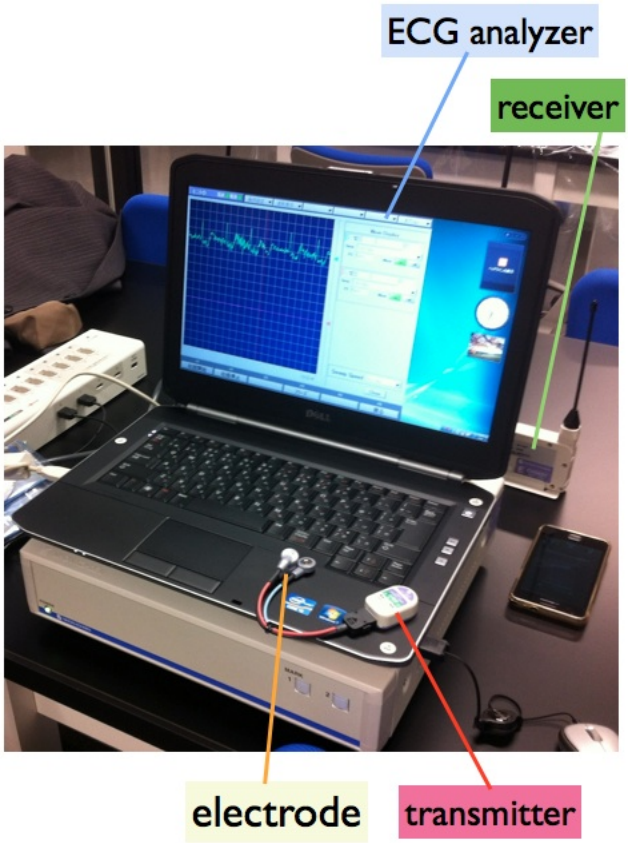


Figure 4.3: Portable ECG system that is used in this research.

4.3 Experimental setup

4.3.1 Preparation of stereoscopic contents

I implemented two experiment sessions. The first section presented low motions stimuli from a city walkthrough made of computer graphic (**CG**). The 3D Studio Max 2011 was used to develop and render the CG contents. The CG movie was developed such that participants perceiving up, down, turn left, and turn right with moderate amount as shown in Fig.4.4. Up and down motion were twice, while turn left and turn right were 7 and 8 times, respectively.

The content of the CG movie was similar as previous study [3]. The movie consisted of narrow walking road with various sizes of buildings, fences, trees, and electric poles on the left and right side. Up and down motion were produced by climbing up and down a bridge.

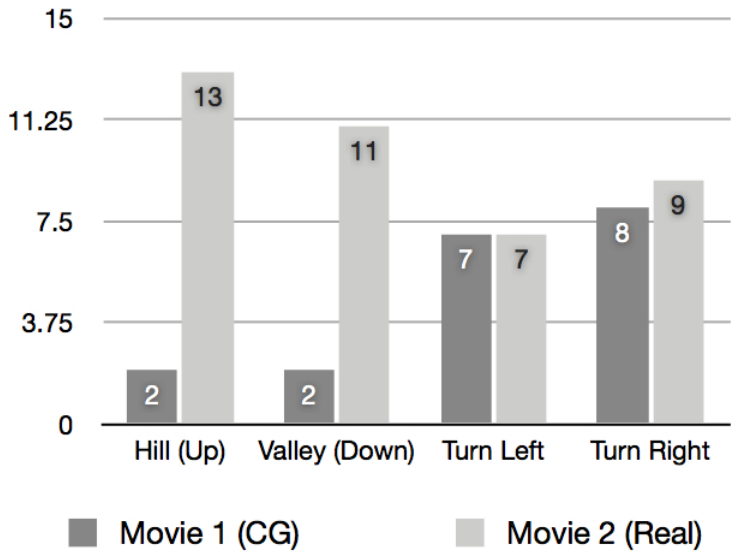


Figure 4.4: Comparison of principal motions in CG movie of gentle walking scene (low stimuli; movie 1) and real movie of roller coaster scene (high stimuli; movie 2).

The left and right scene camera in CG scene were separated by 6.5 cm of disparity (0.0325 unit in 3D Studio Max 2011) to produce stereoscopic effect. This separation produced slightly different view for left and right eye. The left and right scene were then merged using StereoMovie Maker software [7] to provide LR (left-right) type stereoscopic movie. The resolution of the merged movie was 1920 x 1080 pixel. The total duration of the CG movie was 5 minutes. The velocity of camera movement in this CG movie was 63.9 km/hour. The screenshot of CG movie used in the first session is shown in Fig.4.5.

The second session consisted of high motions stimuli from a front seat of real roller coaster scene taken at Great Yarmouth Pleasure Beach, United Kingdom [8]. Permission for usage of the roller coaster video was granted by www.CoasterForce.Com. The movie consisted of 13 up, 11 down, 7 turn left, and 9 turn right motion with high velocity of forward motion. The roller coaster movie contained stronger vection sensation than CG movie as change of motion direction in roller coaster movie was frequent. The roller coaster movie was provided in LR stereoscopic format with 1920 x 1080 pixel of resolution. The screenshot of roller coaster movie used in the second session is shown in Fig.4.6. I integrated an OpenFramework-based stereoscopic movie player [9] onto the 3D gaze tracking system to display stereoscopic movie while recording 3D gaze position.

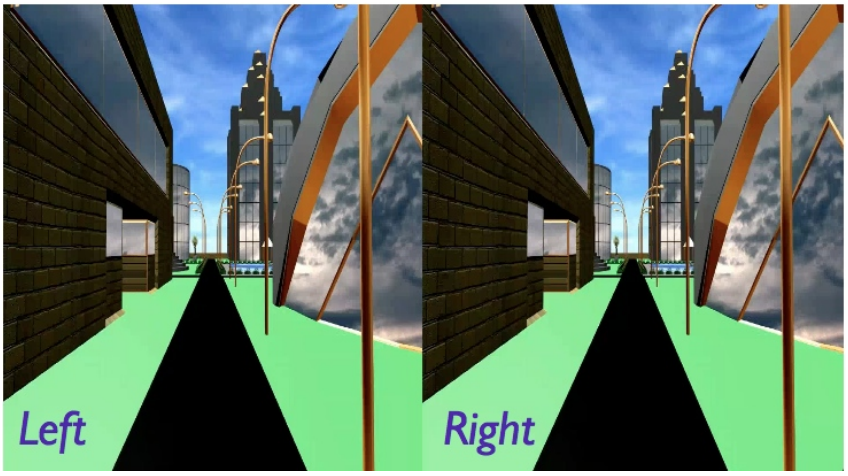


Figure 4.5: Computer graphics movie of city walkthrough. Note that "Left" and "Right" label were added for clarity, which were hidden during experiment.

Figure 4.6: Real movie of roller coaster scene. Note that "Left" and "Right" label were added for clarity, which were hidden during experiment.

4.3.2 Experiment procedure

Before conducting experiment, all experiment procedures were submitted to the ethics committee of School of Information and Telecommunication Engineering, Tokai University, Japan. The committee granted permission and all participants gave informed consents to join the study (see *Appendix B*). All participants were healthy, with normal or corrected eyes.

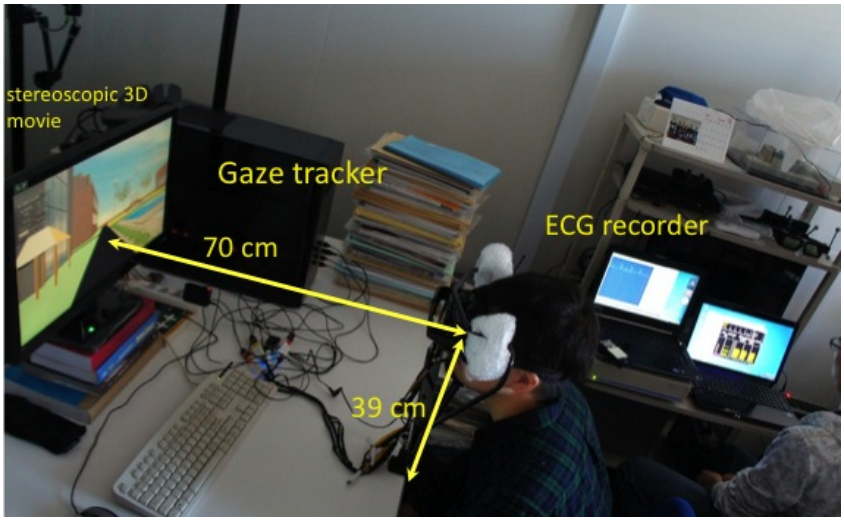


Figure 4.7: Configuration of experiment devices.

In the first session, there were 20 participants joined the experiment (15 male; 5 female; average of age 22.25 year old). The participants were randomly divided to two groups, each of which consisted of 10 students. The first group (8 male; 2 female) was asked to watch CG movie while fixating their gaze to the white point positioned at the center of the screen. The second group (7 male; 3 female) was asked to watch CG movie while freely gazing to all parts of the screen.

The participant was positioned such that the center of the screen was vertically and horizontally aligned at the middle point of both eyes. The distance between participant and screen was set to 70 cm. The height of eye from tabletop was set to 39 cm. The configuration of experiment setting is depicted in Fig.4.7.

The duration of experiment for each participant was about 15 minutes. First, three minutes was given to the participant for listening the explanation of experiment and signed an informed consent. Next, five minutes was provided to install ECG electrodes and prepare gaze tracking recording. Five minutes exposure of stereoscopic 3D content was then performed. Finally, two minutes was provided to rest and answer post-experiment SSQ. A week of rest time was given between the first and the second experiment session to avoid any learning experience and fatigue of the participants.

In the second session, there were also 20 participants joined the experiment (17 male;

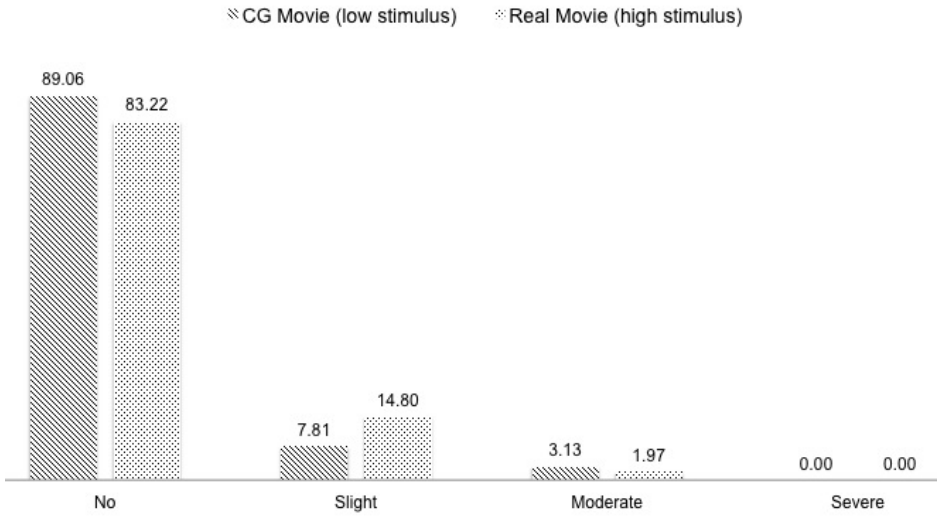


Figure 4.8: Percentage (%) of sickness level for low (CG movie) and high (real movie) motions stimuli. *No*, *Slight*, *Moderate*, and *Severe* level are represented by 0, 1, 2, and 3 in SSQ, respectively.

3 female; average of age 22.65 year old). Eighteen participants had joined the first session of experiment. Two additional male participants joined the second session of experiment to replace two female subjects who could not participate in this session. Randomization of participants was performed when dividing all participants to two groups. The first group (9 male; 1 female) was asked to watch the roller coaster scene while fixating their gaze to a white fixation point at the center of screen. The second group (8 male; 2 female) was asked to watch the roller coaster scene while freely gazing to all parts of the screen. The duration and procedure of experiment is similar to the first session.

4.4 Experimental results

Figure 4.8 shows percentage (%) of sickness level for CG walk through and real roller coaster movie. The results show that CG movie generated lower sickness level ("No effect": 89.06%) than real movie ("No effect": 83.22%). Additionally, there were more viewers of real movie that reported "Slight effect" (14.80%) than viewers of CG movie (7.81%). I found that there were only a few participants who stated "Moderate effect" in both CG movie (3.13%) and real movie (1.97%) while there was no participant felt "Severe effect".

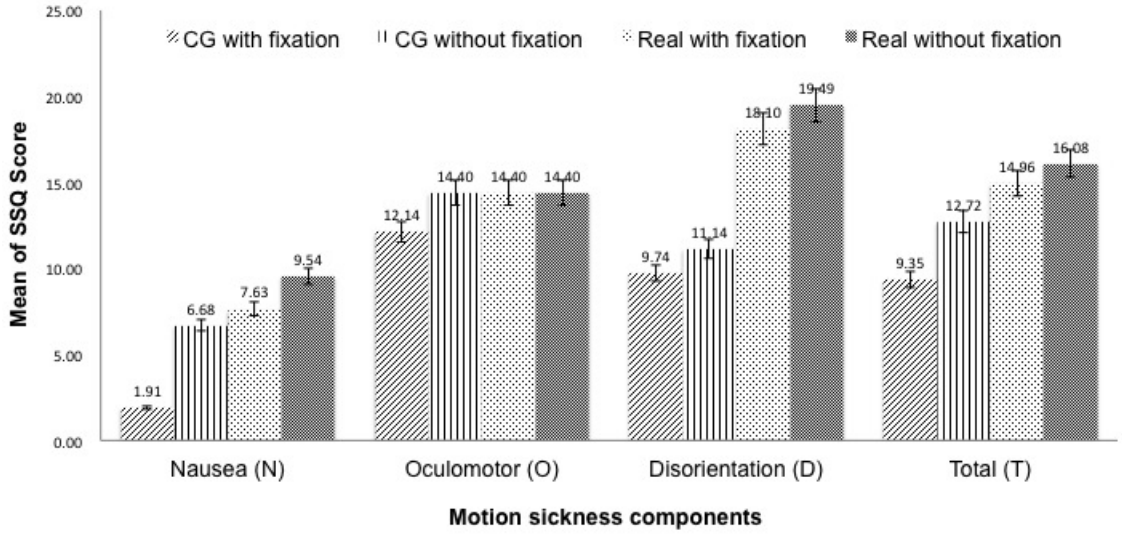
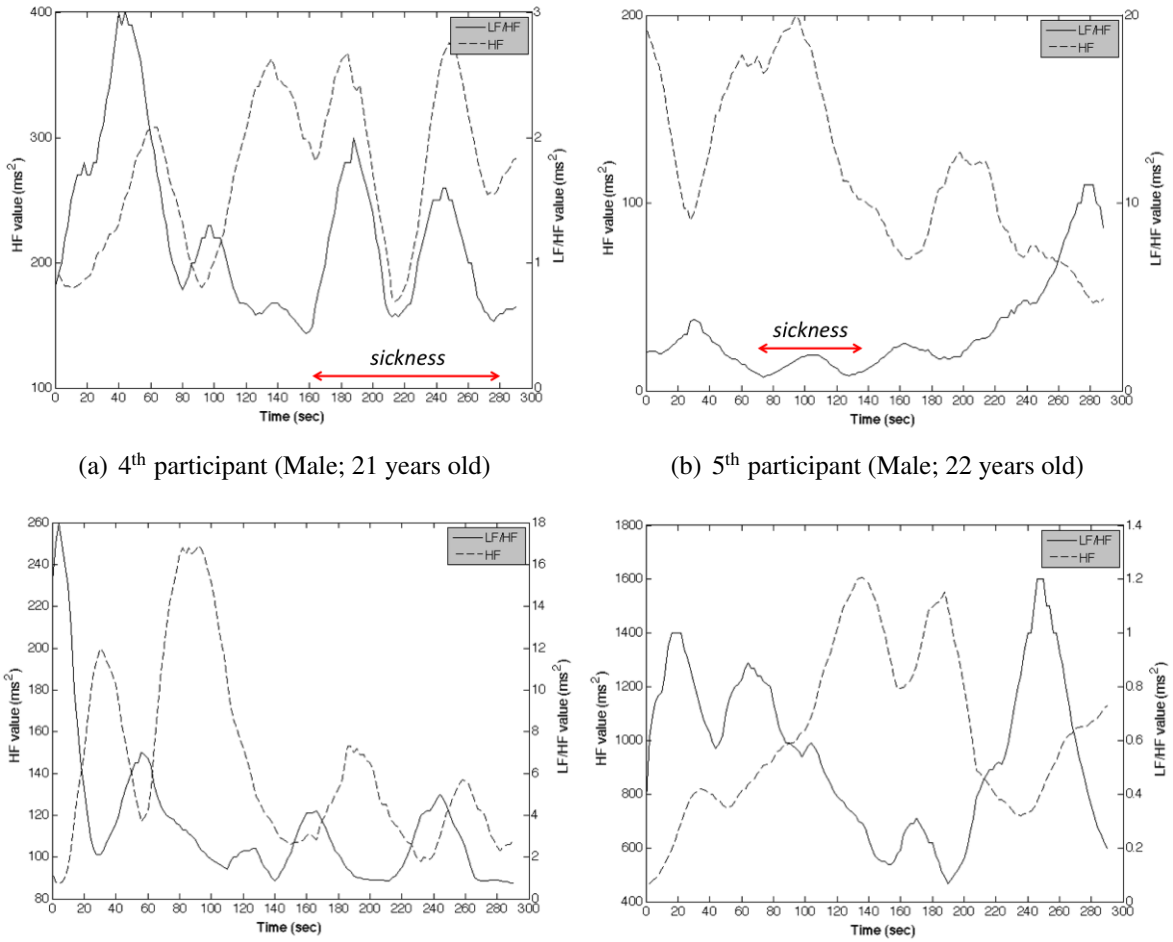


Figure 4.9: Mean of SSQ score for nausea (N), oculomotor (O), disorientation (D), and total (T) component in four experiment conditions (CG movie with fixation, CG movie without fixation, real movie with fixation, and real movie without fixation).

Figure 4.9 shows general results of SSQ score for nausea (N), oculomotor (O), disorientation (D), and total (T) component. Participants experienced smaller VIMS symptoms when watching CG movie with fixation ($N = 1.91 \pm 3.82$, $O = 12.14 \pm 13.24$, $D = 9.74 \pm 16.53$, $T = 9.35 \pm 10.21$). Increment of average SSQ score in N, D, and T component is observed in group of participants who watched CG movie without fixating their gaze at one point ($N = 6.68 \pm 11.33$, $O = 14.40 \pm 12.43$, $D = 11.14 \pm 24.74$, $T = 12.72 \pm 13.4$). Similar pattern can be observed in group of participants who watched real movie with fixation ($N = 7.63 \pm 8.32$, $O = 14.40 \pm 14.16$, $D = 18.10 \pm 22.49$, $T = 14.96 \pm 14.68$) and without fixation ($N = 9.54 \pm 11.29$, $O = 14.40 \pm 13.75$, $D = 19.49 \pm 25.06$, $T = 16.08 \pm 15.78$). It can be seen that fixation of gaze suppresses VIMS. Furthermore, real movie resulted higher SSQ score in N, D, and T component than CG movie.



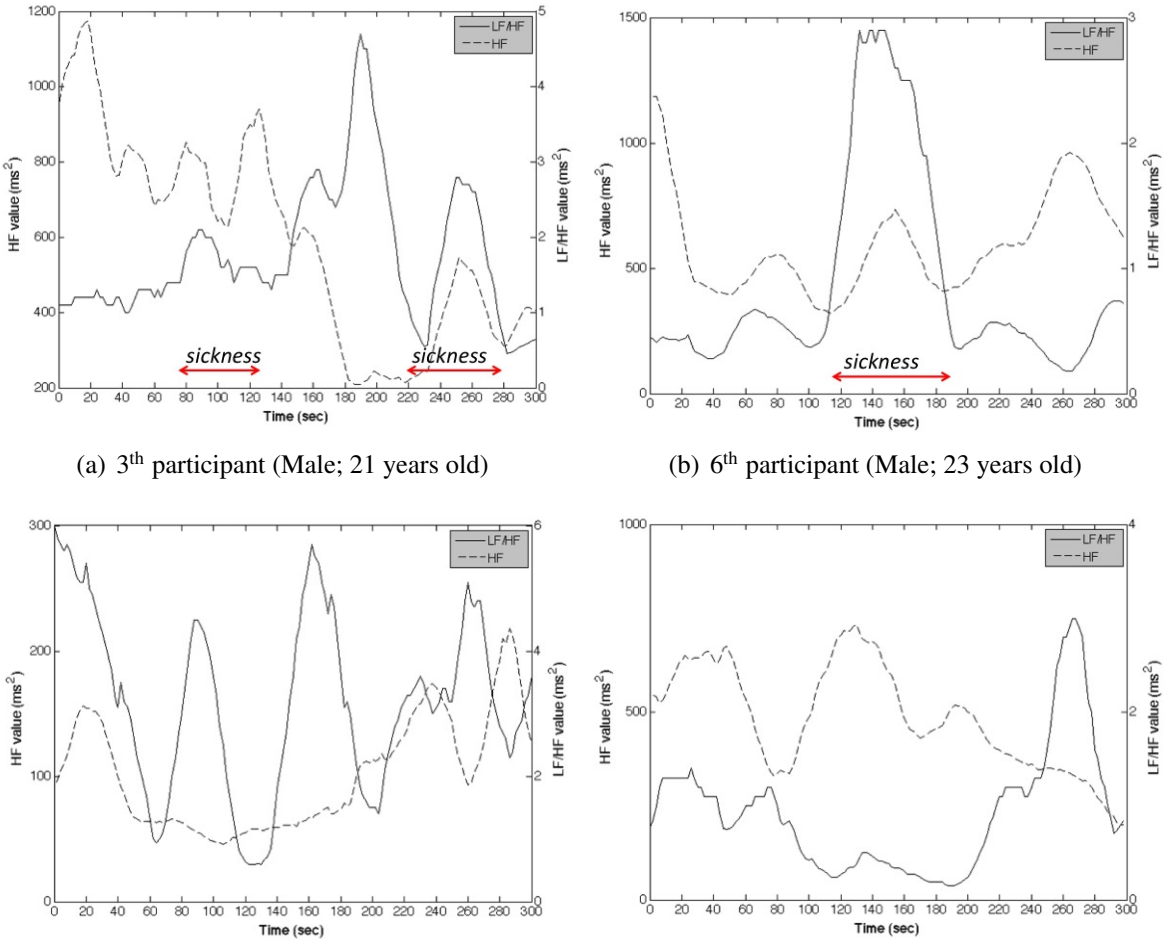
(a) 4th participant (Male; 21 years old)

(b) 5th participant (Male; 22 years old)

(c) 3rd participant (Male; 22 years old)

(d) 8th participant (Male; 21 years old)

Figure 4.10: **CG movie with fixation**: comparison of LF/HF (solid line) and HF (dashed line) data from participants with high sickness (panel (a) and (b)) and participants with low or no sickness (panel (c) and (d)). Left and right *Y*-axes show HF (ms^2) and LF/HF (ms^2) value, respectively. Horizontal axis shows time sequence (*second*). Dominant activity of sympathetic nerves is shown by red arrow.



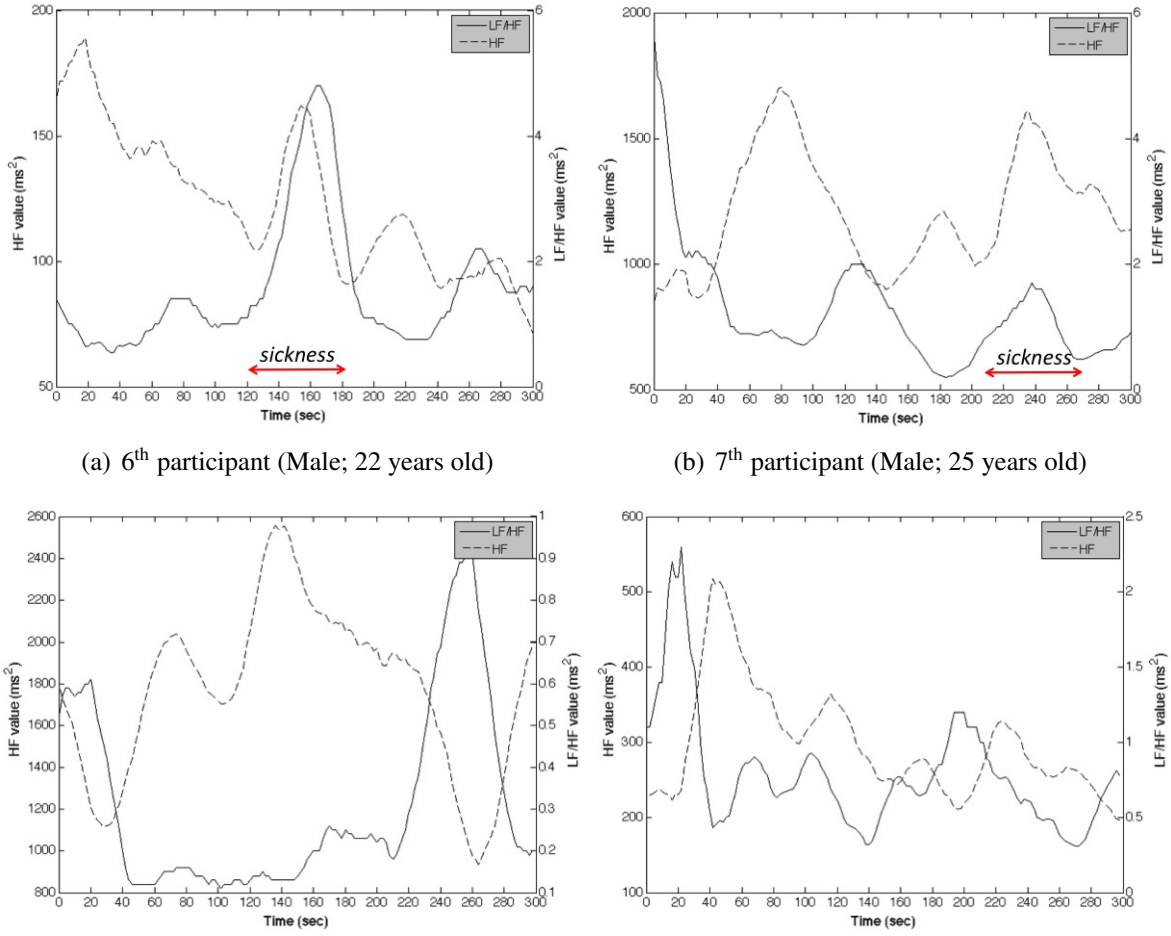
(a) 3th participant (Male; 21 years old)

(b) 6th participant (Male; 23 years old)

(c) 7th participant (Male; 25 years old)

(d) 9th participant (Male; 21 years old)

Figure 4.11: **CG movie without fixation**: comparison of LF/HF (solid line) and HF (dashed line) data from participants with high sickness (panel (a) and (b)) and participants with low or no sickness (panel (c) and (d)). Left and right Y-axes show HF (ms^2) and LF/HF (ms^2) value, respectively. Horizontal axis shows time sequence (*second*). Dominant activity of sympathetic nerves is shown by red arrow.



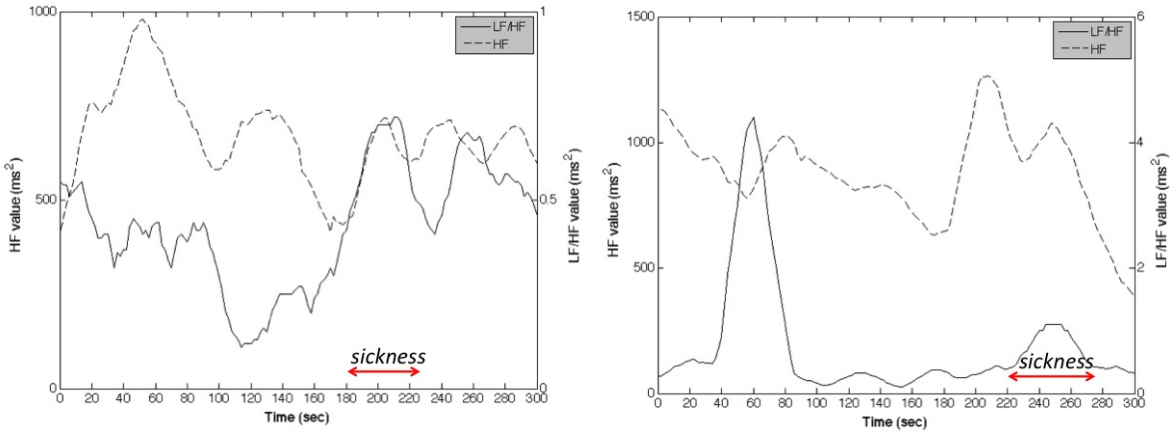
(a) 6th participant (Male; 22 years old)

(b) 7th participant (Male; 25 years old)

(c) 3th participant (Male; 25 years old)

(d) 9th participant (Male; 24 years old)

Figure 4.12: **Real movie with fixation**: comparison of LF/HF (solid line) and HF (dashed line) data from participants with high sickness (panel (a) and (b)) and participants with low or no sickness (panel (c) and (d)). Left and right Y-axes show HF (ms^2) and LF/HF (ms^2) value, respectively. Horizontal axis shows time sequence (*second*). Dominant activity of sympathetic nerves is shown by red arrow.



(a) 4th participant (Female; 21 years old)

(b) 8th participant (Female; 21 years old)

(c) 2th participant (Male; 24 years old)

(d) 7th participant (Male; 22 years old)

Figure 4.13: **Real movie without fixation**: comparison of LF/HF (solid line) and HF (dashed line) data from participants with high sickness (panel (a) and (b)) and participants with low or no sickness (panel (c) and (d)). Left and right Y-axes show HF (ms^2) and LF/HF (ms^2) value, respectively. Horizontal axis shows time sequence (*second*). Dominant activity of sympathetic nerves is shown by red arrow.

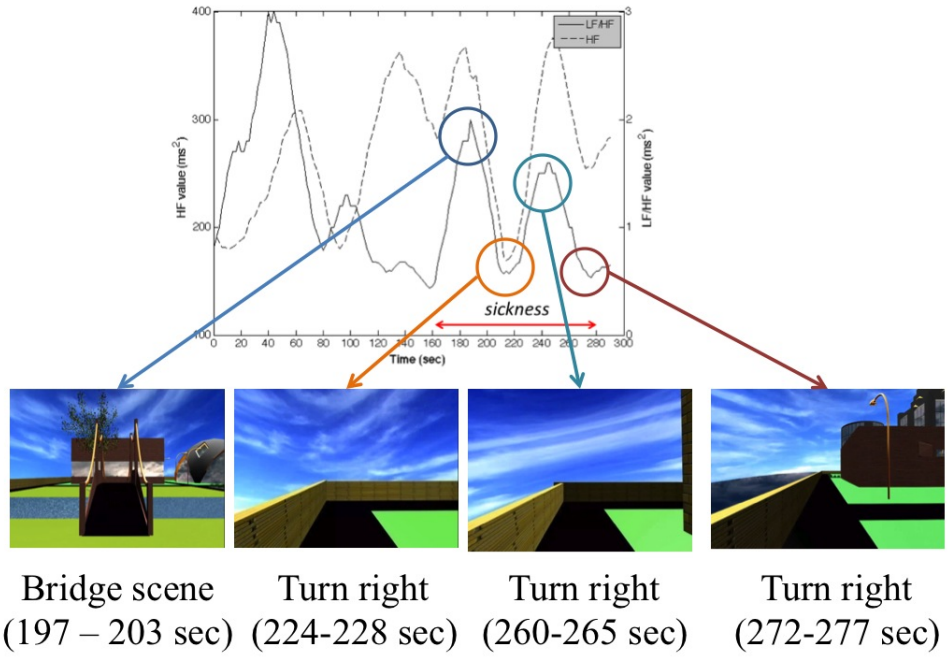
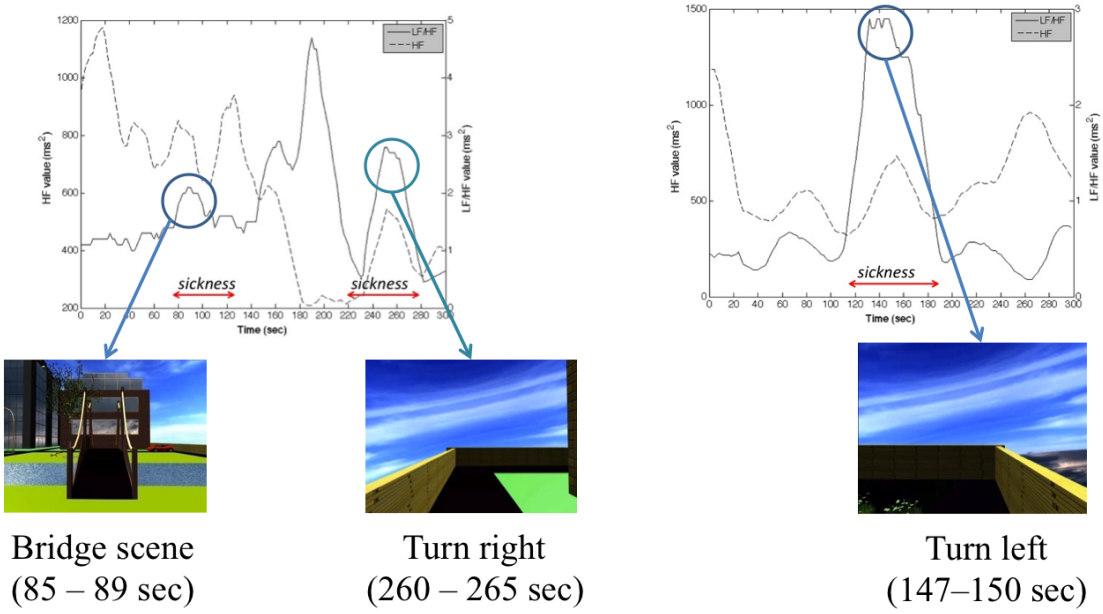
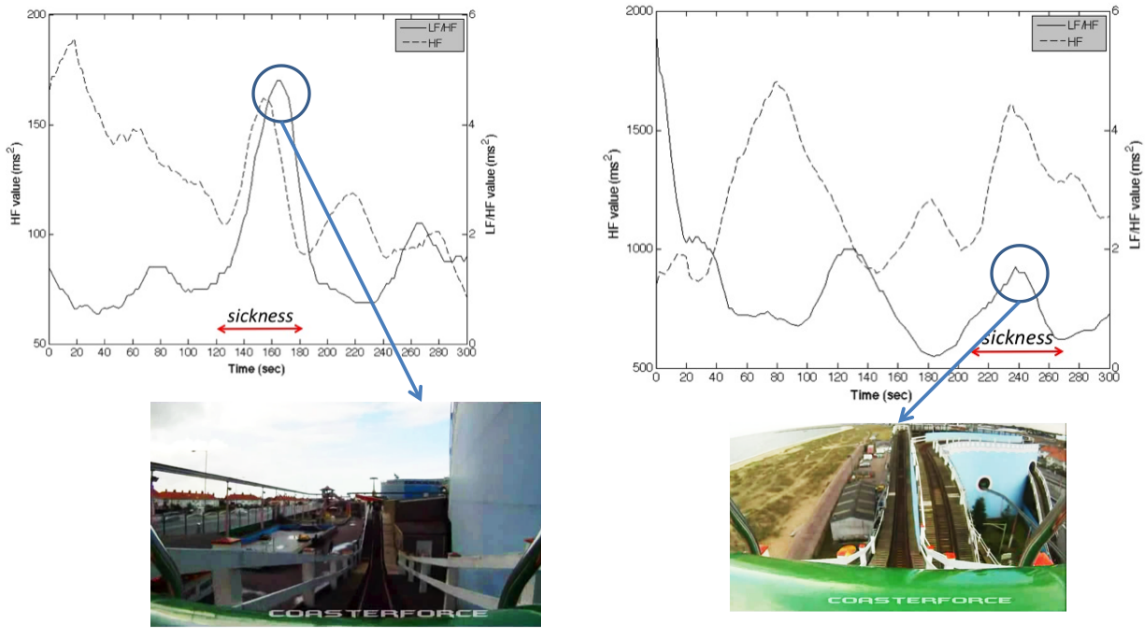


Figure 4.14: **CG movie with fixation**: sample of several scenes that contribute to dominant sympathetic nerves activity of the 4th participant



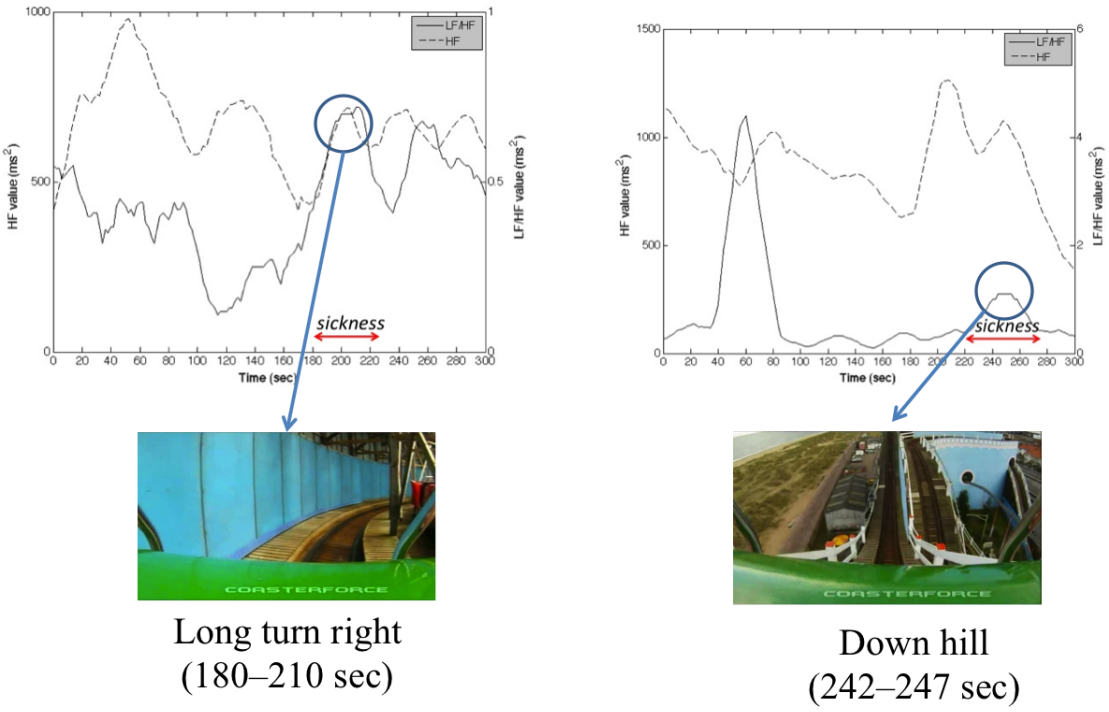
(a) 3rd participant (b) 6th participant

Figure 4.15: **CG movie without fixation**: sample of several scenes seen that contribute to dominant sympathetic nerves activity of the 3rd (panel (a)) and 6th (panel (b)) participant.



(a) 6th participant (b) 7th participant

Figure 4.16: **Real movie with fixation**: sample of several scenes seen that contribute to dominant sympathetic nerves activity of the 6th (panel (a)) and 7th (panel (b)) participant.



(a) 4th participant (b) 8th participant

Figure 4.17: **Real movie without fixation**: sample of several scenes seen that contribute to dominant sympathetic nerves activity of the 4th (panel (a)) and 8th (panel (b)) participant.

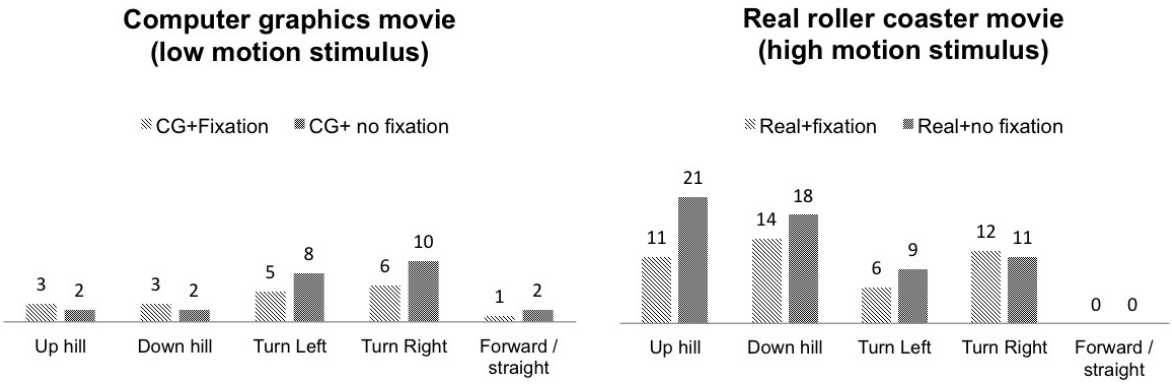


Figure 4.18: Amount of dominant sympathetic nerves activities from all participants during exposure of CG (left panel) and real roller coaster (right panel) movie.

I also evaluated the results of SSQ using two-way analysis of variance (**ANOVA**) between subjects with $\alpha = 0.05$. The two independent variables are movie (CG vs. real movie) and gaze (fixation vs. no fixation). The dependent variables are nausea, disorientation, oculomotor, and total score.

Statistical analysis shows that effect of movie to nausea ($F(1,36) = 11.161$, $p<0.005$) and disorientation ($F(1,36) = 5.886$, $p<0.05$) component is significant. No other significant effect was found in analysis.

I analyzed the results of ECG data by plotting HF and LF/HF of heart rate variability (HRV). Based on Nakagawa et al. [6], I investigated activity of sympathetic and parasympathetic nerves when both were forming unbalanced graph. I observed minimum and maximum point of LF/HF and LF graphs. Comparisons of LF/HF (sympathetic nerves) and HF (parasympathetic nerves) activity of four groups are shown in Fig.4.10–4.13.

In Fig.4.10, I show four samples of HRV during exposure taken from participants with high SSQ score (4^{th} and 5^{th} participant) and low or no sickness (3^{rd} and 8^{th} participant). The participants experienced dominant sympathetic nerves activity in several parts of scenes. For example, the 4^{th} participant experienced VIMS symptoms between 160–280 seconds while the 5^{th} participant experienced VIMS between 80–120 seconds. On the other side, two graphs of other participants (the 3^{rd} and the 8^{th} participant) show no dominant activity of sympathetic nerves. Using the same method, I investigated all ECG data of participants in other three groups (CG movie without fixation, real movie with fixation, real movie without fixation) and show them in Fig.4.11, 4.12, and 4.13.

Fig.4.14–4.17 show sample of several scenes that contribute to VIMS. I investigated scenes in movie during high sympathetic nerves activity based on maximum and minimum point found LF/HF and HF graph (Fig.4.10–4.13). Two viewers of CG movie (one participant in gaze with fixation session and the other in gaze without fixation session) experienced high sympathetic nerves activity when seeing forward motion while other viewers mostly induced by horizontal (turn left and right) and vertical (up and down hill) motion.

I found that the 4^{th} participant who viewed CG movie with fixation experienced high sympathetic nerves activity when seeing bridge scene (197–203 sec) and turn right scene (223–228 sec, 224–228 sec, and 272–277 sec), as shown in Fig.4.14. The average LF/HF and HF values during bridge scene is about 1.28 and 264.5 ms^2, respectively. During the

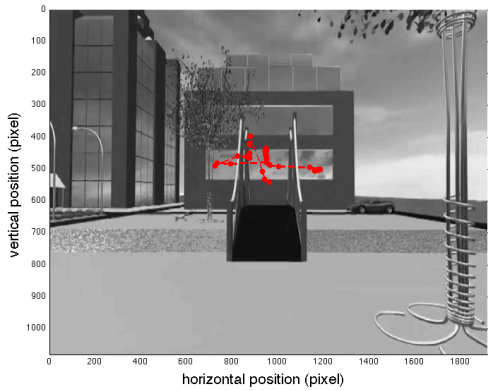
1st, 2nd, 3rd turn right, the average of LF/HF and HF values are 0.83 and 213.33 ms^2; 0.95 and 307 ms^2; 0.56 and 254.67 ms^2, respectively.

Fig.4.15 shows sample of scenes and ECG data of the 3rd and 6th participant from CG movie without fixation group. The 3rd participant experienced high sympathetic nerves activity from bridge (85–89 sec; avg.LF/HF=2.05 ms^2; avg.HF= 813.25 ms^2) and turn right scene (260–265 sec; avg.LF/HF= 2.45 ms^2; avg.HF=481 ms^2) while the 6th participant experienced high sympathetic nerves activity from turn left scene (147–150 sec; avg.LF/HF=2.87 ms^2; avg.HF=688 ms^2).

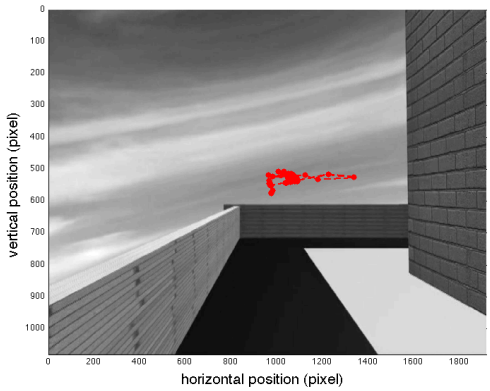
Fig.4.16 shows sample of scenes and ECG data of the 6th and 7th participant from real movie with fixation group. The 6th participant experienced high sympathetic nerve activity during downhill scene (168–172 sec; avg.LF/HF=4.6 ms^2; avg.HF=121 ms^2). The 7th participant also experienced high sympathetic nerve activity during the other downhill scene of roller coaster (242–247 sec; avg.LF/HF=1.53 ms^2; avg.HF=1515.25 ms^2).

In real movie without fixation session depicted in Fig.4.17, I found that the 4th participant experienced high sympathetic nerves activity during long turn right scene of roller coaster inside a tunnel (180–210 sec). The average LF/HF and HF values during this scene are 0.62 and 607.125 ms^2, respectively. The 8th participant experienced high sympathetic nerves activity during downhill scene (242–247 sec) with 1.08 and 1046.25 ms^2 of average LF/HF and HF values, respectively.

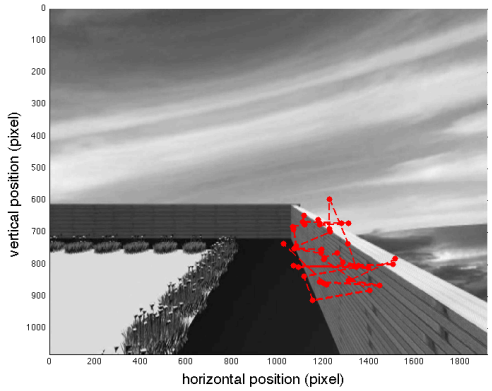
I counted all maximum and minimum point when sympathetic nerves activity is high. The results are presented in Fig.4.18, which consisted of CG movie viewers (left panel) and real roller coaster movie viewers (right panel). From the left panel of Fig.4.18, there were participants who experienced sickness when seeing long forward motion. However, the effect of long forward motion is not as strong as effect of horizontal and vertical motion. I found that horizontal motion (turn left and right) is the dominant contributing factor of VIMS in CG movie. I also found that gaze fixation reduces sympathetic nerves activity since participants who fixed their gaze experienced less sympathetic nerves activity than ones without gaze fixation.



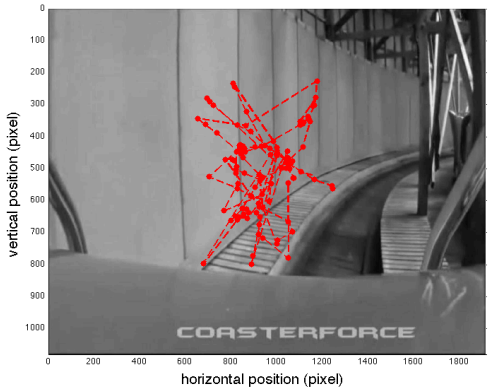
(a) Bridge scene (85–89 *sec*)



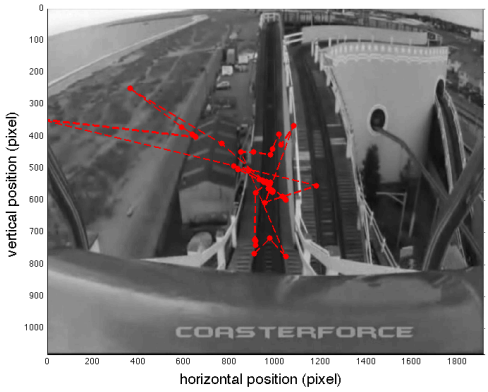
(b) Turning right scene (260–265 *sec*)



(c) Turning left scene (147–150 *sec*)



(d) Long turning right scene (180–190 *sec*)



(e) Downhill scene (242–247 *sec*)

Figure 4.19: Five samples scene of 2D point of gaze from participants experienced VIMS in no fixation group of CG and real movie. Vertical and horizontal axis are in pixel unit.

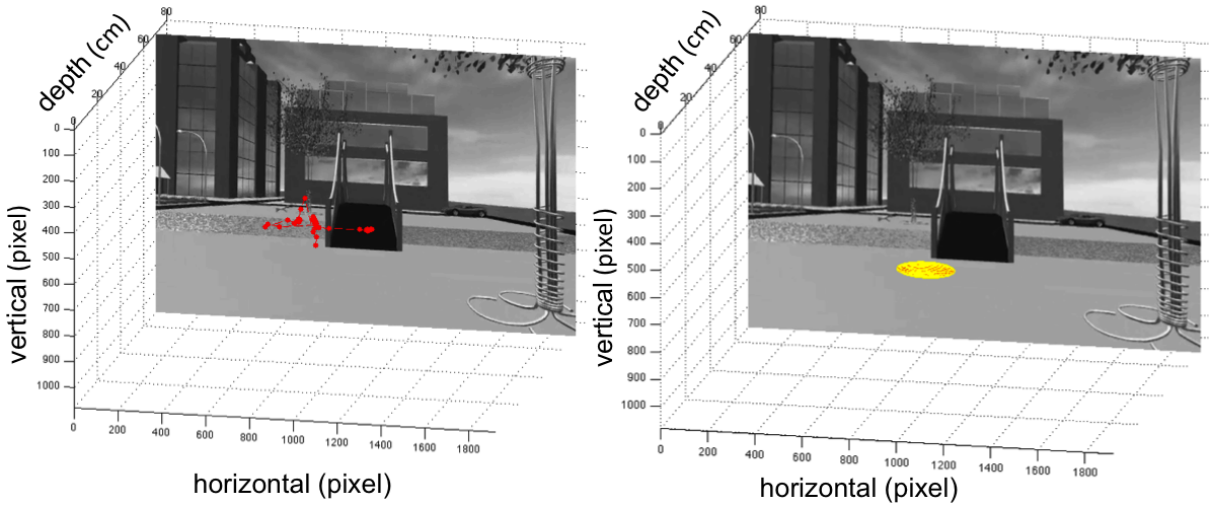


Figure 4.20: 3D point of gaze trajectory (left panel) and anisotropy map (right panel) from a participant experiencing up and down motion of bridge scene (85–89 *sec*) in CG movie without fixation. Note that the unit of horizontal and vertical axis is *pixel* while depth axis is in *cm* unit. The 2D point of gaze is shown in panel (a) Fig.4.19.

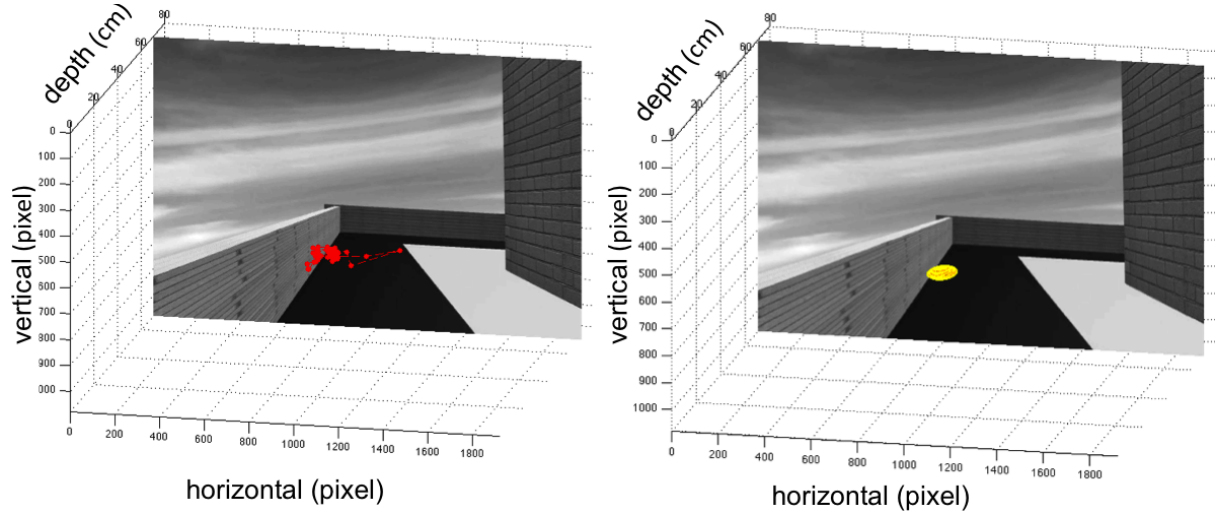


Figure 4.21: 3D point of gaze trajectory (left panel) and anisotropy map (right panel) from a participant viewing turn right scene (260–265 *sec*) in CG movie without fixation. Note that the unit of horizontal and vertical axis is *pixel* while depth axis is in *cm* unit. The 2D point of gaze is shown in panel (b) of Fig.4.19.

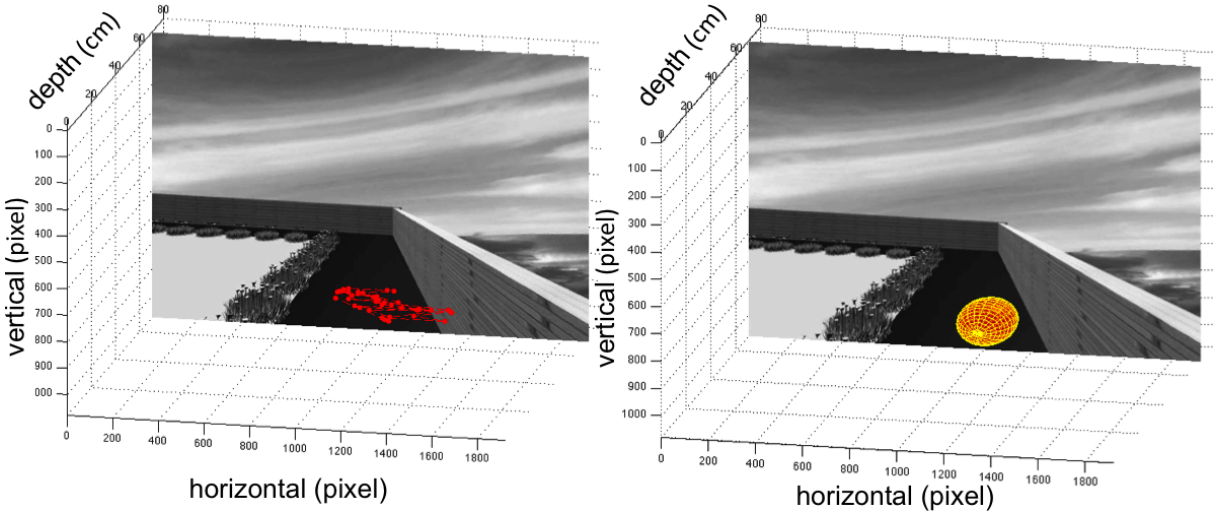


Figure 4.22: 3D point of gaze trajectory (left panel) and anisotropy map (right panel) from a participant viewing turn left scene (147–150 *sec*) in CG movie without fixation. Note that the unit of horizontal and vertical axis is *pixel* while depth axis is in *cm* unit. The 2D point of gaze is shown in panel (c) of Fig.4.19.

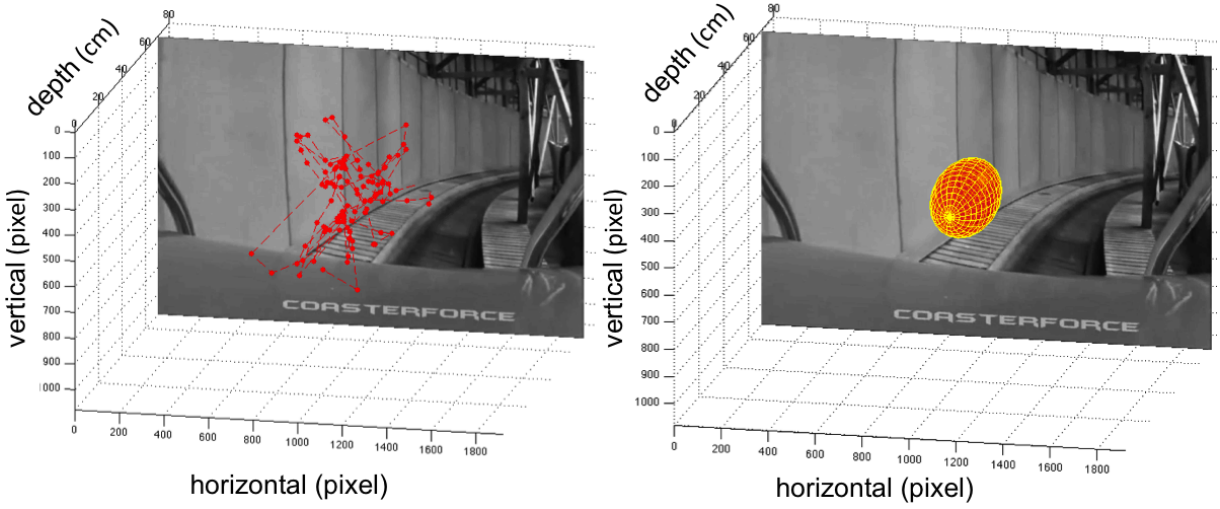


Figure 4.23: 3D point of gaze trajectory (left panel) and anisotropy map (right panel) from a participant viewing turn right scene (180–190 *sec*) in real roller coaster movie without fixation. Note that the unit of horizontal and vertical axis is *pixel* while depth axis is in *cm* unit. The 2D point of gaze is shown in panel (d) of Fig.4.19.

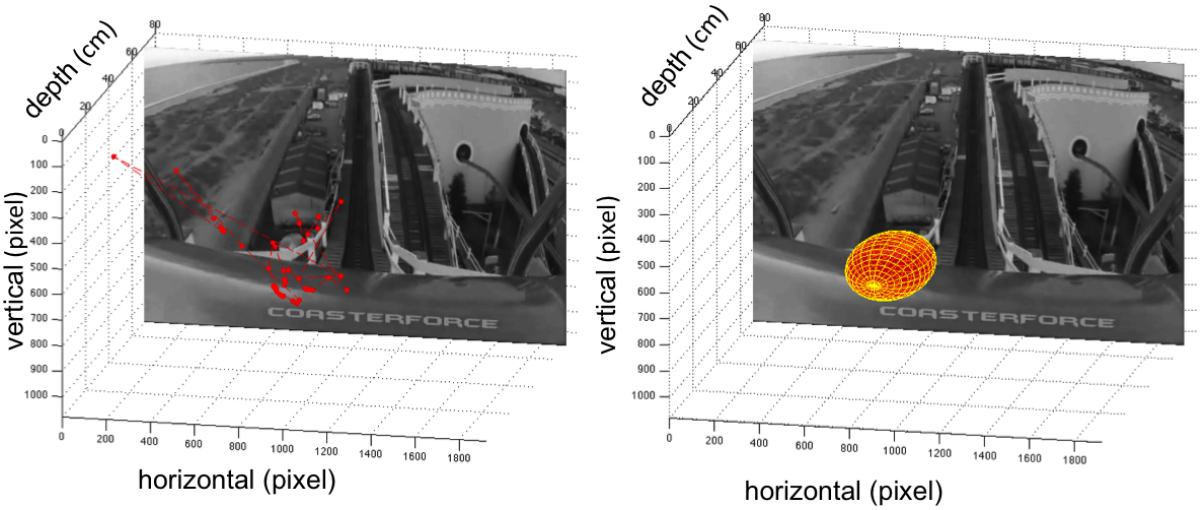


Figure 4.24: 3D point of gaze trajectory (left panel) and anisotropy map (right panel) from a participant viewing down hill scene (242–247 *sec*) in real roller coaster movie without fixation. Note that the unit of horizontal and vertical axis is *pixel* while depth axis is in *cm* unit. The 2D point of gaze is shown in panel (e) of Fig.4.19.

The right panel of Fig.4.18 shows amount of VIMS contributing factors in real roller coaster movie. I found that vertical motion (up and down hill) is the dominant contributing factor of VIMS. Interestingly, I found no participant who experienced high sympathetic nerves activity during long forward motion. Most of participants who freely gazed at screen experienced high sympathetic nerves activity except in turn right motion.

Figure 4.19 shows five samples scene of 2D point of gaze from participant experienced VIMS in no fixation group of CG and real movie. All samples were taken in provoking scenes, such as bridge scene, turning right scene, turning left scene, and downhill scene. Fig.4.20–4.24 show various trajectories of 3D point of gaze and their anisotropy maps during provoking scenes. Point and line in trajectories represent point of gaze and movement between one gaze point to other point. Anisotropy map is ellipsoid representation of mean and standard deviation of 3D point of gaze, by which tendency of 3D point of gaze can be observed. The center of ellipsoid is the mean of gaze in X, Y, and Z direction. The axes of ellipsoid are standard deviation of gaze in X, Y, and Z direction.

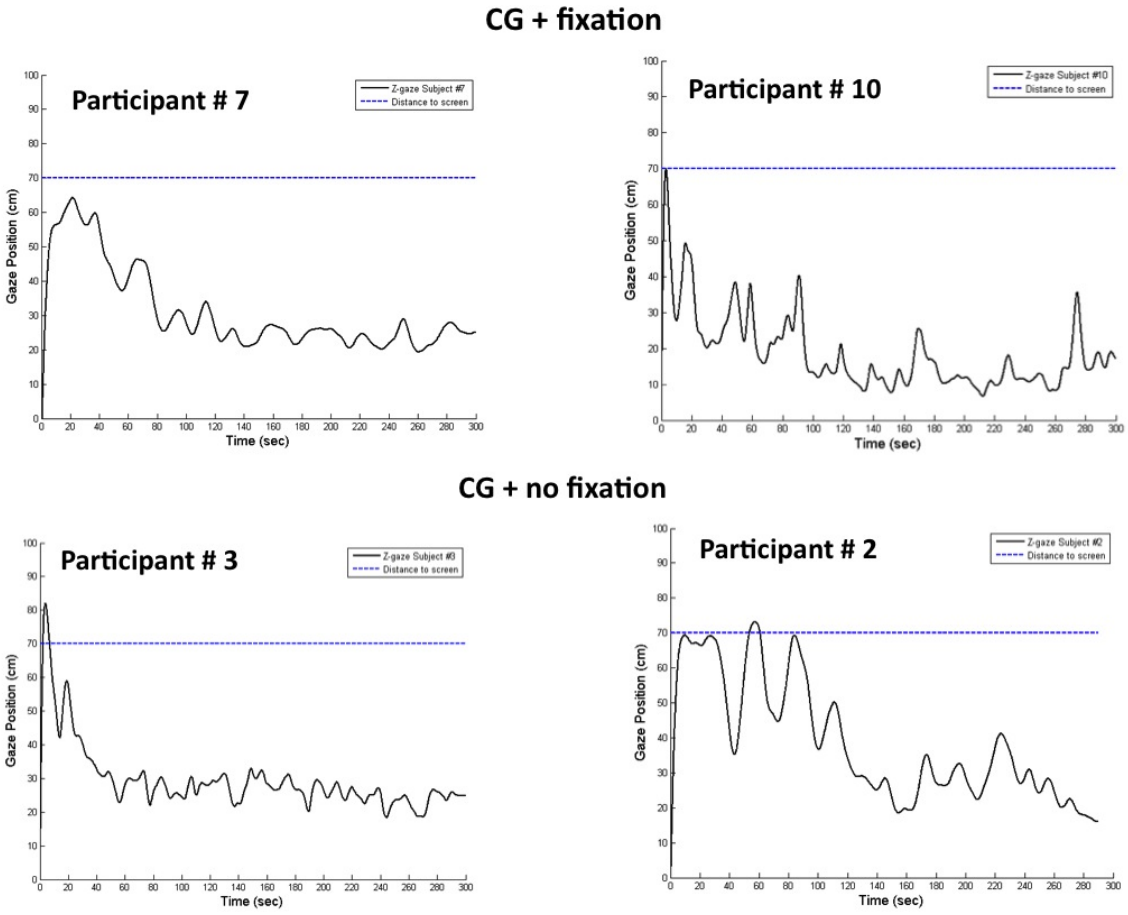


Figure 4.25: Sample of depth gaze (Z-gaze) from several participants of the first experiment session (CG movie). Top and bottom panel are CG movie with fixation group and CG movie without fixation group, respectively. Solid and dashed line are Z-gaze and distance to screen, respectively. Horizontal and vertical axis are time (*sec*) and Z-gaze position (*cm*).

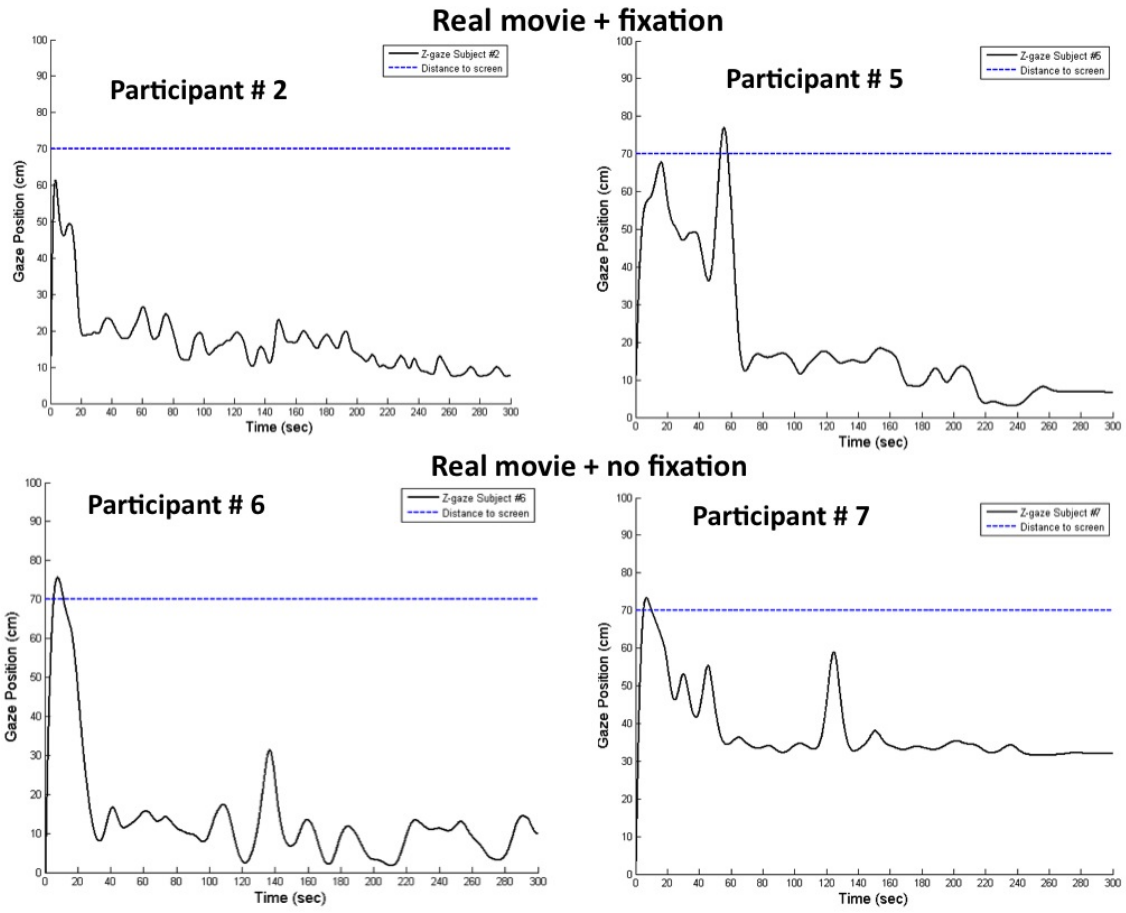
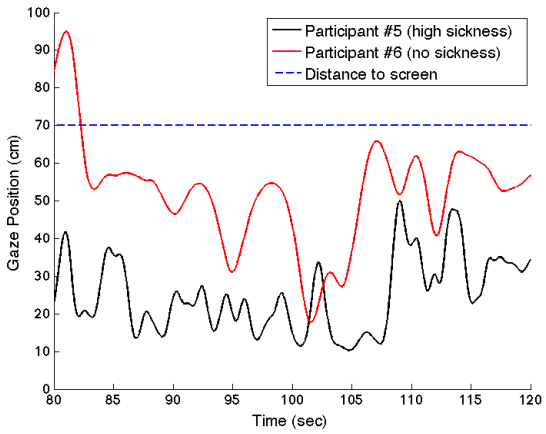
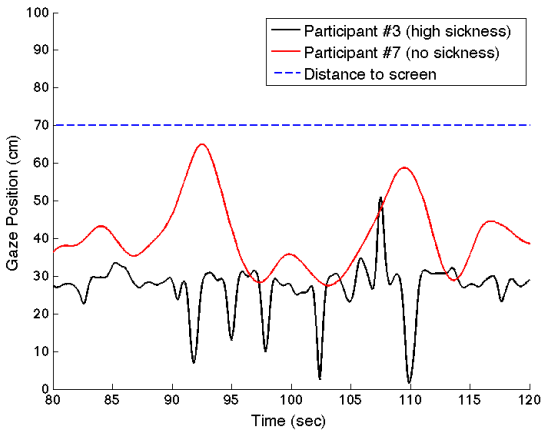


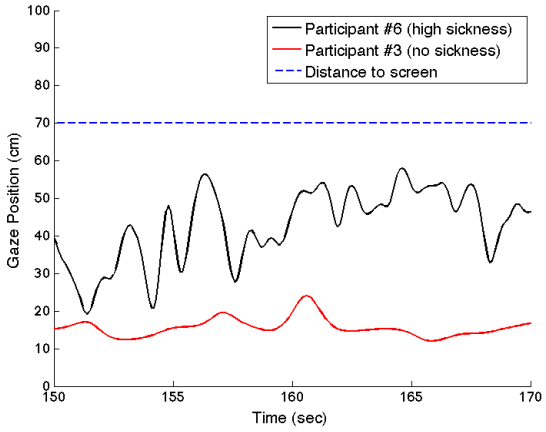
Figure 4.26: Sample of depth gaze (*Z*-gaze) from several participants of the second experiment session (real movie). Top and bottom panel are real movie with fixation group and real movie without fixation group, respectively. Solid and dashed line are Z-gaze and distance to screen, respectively. Horizontal and vertical axis are time (*sec*) and *Z*-gaze position (*cm*).



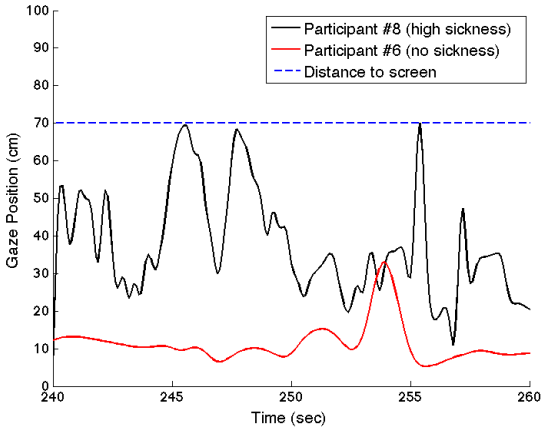
(a) CG movie with fixation: bridge scene (80-120 *sec*)



(b) CG without fixation: bridge scene (80-120 *sec*)



(c) Real movie with fixation: downhill scene (150–170 *sec*)



(d) Real movie without fixation: downhill scene (240–260 *sec*)

Figure 4.27: Comparison of depth gaze from participants with high and no sickness during provoking scene. Black, red, and blue line are participants with high sickness, participant with no sickness, and distance of screen, respectively.

Table 4.2: Center (mean in X, Y, and Z direction) and axes (standard deviation in X, Y, and Z direction) of ellipsoid representation in anisotropy maps

Anisotropy maps	**Dimensions**		
	X (*pixel*)	Y (*pixel*)	Z (*cm*)
1 (Fig.4.20)	970.03 ± 123.87	605.96 ± 27.78	29.06 ± 2.06
2 (Fig.4.21)	1056.40 ± 61.27	547.78 ± 14.54	20.29 ± 4.35
3 (Fig.4.22)	1218.40 ± 123.68	760.92 ± 77.25	31.11 ± 8.45
4 (Fig.4.23)	944.48 ± 135.80	494.92 ± 129.71	56.17 ± 13.48
5 (Fig.4.24)	899.87 ± 187.54	534.30 ± 110.48	15.42 ± 14.52

Table 4.2 shows data of ellipsoid representation in anisotropy maps. The 3D position of ellipsoids are determined by means of gaze data in X, Y, and Z dimension. The size of ellipsoids are determined by standard deviations of gaze data in X, Y, and Z dimension. Table 4.2 shows that participants tended to gaze in horizontal rather than vertical direction. Standard deviation of gaze data in X dimension is larger than in Y dimension. Note that the unit of horizontal (X) and vertical (Y) axis is *pixel* while depth axis (Z) is in *cm* unit. *Pixel* unit is used in X and Y dimension as the selected provoking scenes in anisotropy maps were displayed in *pixel* unit on the computer screen. On the other side, the depth gaze (Z) was measured in *cm* unit using 3D gaze tracking.

Figure 4.25 and 4.26 show samples of depth gaze from CG and real movie experiment session, respectively. I found that depth gaze behavior of participants formed similar trend. At the beginning of the movie, the depth gaze was positioned nearly as far as distance of screen from participant (70 *cm*). As the movie progressed, the depth gaze tended to be below 70 *cm* and sustained during the movie. This finding occurred on both fixation and no fixation task. The average deviation error of depth gaze is 43.95 ± 6.54 *cm*, i.e. depth gaze was approximately shortened from 70 *cm* to approximately 26 *cm* distance from participant.

Figure 4.27 shows comparison of depth gaze from participants with high sickness and no sickness in several provoking scenes. Participants with high sickness and no sickness are shown in black and red line, respectively. Distance of screen from participant is shown in dashed blue line. Similar trends across group of participants were observed. I found that as participants experienced VIMS, their depth gaze oscillated more frequently than participants who did not experience VIMS.

4.5 Discussions

Given two different types of movie with low and high dynamic motions stimuli, SSQ results show that movie with high motions stimuli (real roller coaster movie) produces more sickness effect than movie with low motions stimuli (CG city walkthrough movie). Increment of sickness effect was observed with significant effect in nausea ($p<0.005$) and disorientation ($p<0.05$) factor of SSQ. Furthermore, gaze fixation during exposure was found to be an effective way to reduce VIMS. This finding was confirmed by ECG data from both CG and real movie (Fig.4.18). The results agree with previous research work [10].

ECG data shows that horizontal motion is the dominant contributing factor of VIMS in CG movie while vertical motion is the dominant contributing factor of VIMS in real roller movie (Fig.4.18). One possible reason is different composition of motions in the movie as shown in Fig.4.4. Since the roller coaster movie contains more vertical motion, VIMS is mostly induced by vertical motion. Compared with forward motion, vertical (pitch) and horizontal (yaw) motion have significant effect for the occurrence of VIMS. This finding confirms previous research work [11].

Figure 4.19 shows several samples scene of 2D point of gaze from no fixation groups in CG and real movie. The visualization of 3D point of gaze and anisotropy maps of Fig.4.19 are shown in Fig.4.20–4.24. The position and size of ellipsoids in anisotropy maps are presented in Table 4.2. From these samples, I found that variability of 3D gaze area in CG movie is smaller than real roller coaster movie. This finding may relate with higher amount of dynamic motion stimuli in real roller coaster movie rather than CG movie. As the level of VIMS increases, the variability of 3D gaze area also increases.

The anisotropy maps show that participant tended to gaze in horizontal rather than vertical direction. This finding may correlate with a fact that horizontal and vertical components of saccades are controlled by different part of brainstem nuclei [12, 13]. Since vertical and horizontal components are independently controlled, vertical eye movements of left and right eye are normally coupled while horizontal eye movements of left and right eye are disconjugated [14].

The disconjugation of horizontal eye movements of left and right eye is due to requirement to match stereoscopic image in left and right retina for different depth of virtual 3D

object. Various small eye movements also support the stereoscopic matching to maintain persistence of horizontal gaze fixation [14, 15]. However, the experimental results show that disconjugation of horizontal eye movements is not the cause of visually induced motion sickness.

Observation of depth gaze data shows that vergence is not only induced by binocular disparity, but also forward and backward motion [16, 17]. When human sees a movie with forward motion, the velocity of the retinal flow at any point depends on distance of viewing in that location. This condition implies faster movement of nearby object on the retina while slower movement of farther object. To stabilize the retinal image of a nearby object, an eye movement is needed. Since forward motion is related with changes in viewing distance, the participants should converge their eyes during forward motion. Closer object induces bigger amplitude of vergence eye movement. Since vergence eye movement affect depth gaze, prolonged forward motion compresses depth gaze (Fig.4.25 and 4.26). Given that participant always focus their view at the screen, this depth gaze compression strengthen unnatural decoupling between vergence and accommodation, which is one cause of visual fatigue and VIMS [18, 19].

Investigation to the depth gaze of participants who experienced high sickness reveals that there is more variability and oscillation of depth gaze during provoking scene (Fig.4.27). This finding shows that intense oscillation of depth gaze during provoking scene in dynamic 3D contents can be used as indicator of VIMS occurrence. The result also shows that instead of applying scleral search coil (**SSC**) or high frame rate binocular eye tracker (>1000 Hz), detection of visually induced motion sickness in stereoscopic environment is sufficient with low frame rate consumer-grade cameras (± 25 Hz).

Oscillation of depth gaze may relate with the presence of intense *optokinetic nystagmus* (**OKN**) in participants with high susceptibility of VIMS [20]. Previous research works found that radial optic flow stimuli induced OKN with version and vergence components [21–24]. OKN is involuntary eye movement following visual stimuli. OKN allows participants to follow visual stimuli and reduce retinal slip. Horizontal and vertical OKN particularly affect stabilization of depth gaze. However, previous research work also found that OKN was one component of eye movement responsible for VIMS as OKN increased velocity of eye movement [25].

The 3D gaze tracking data suggests that movie with forward motion produces similar result as optical flow, i.e. forward motion induces OKN with version and vergence components. The center of screen is the area with the lowest local image velocity. Local image velocity increases toward the periphery of the scene. As long as participants gaze at the center of screen, image of scene remains stable on the retina. If participants fixate their gaze away from the center of screen, retinal image velocity in central vision also increases. The eyes compensate the increment of velocity to stabilize image on the retina. In this case, gaze fixation is useful for participants with high VIMS susceptibility to reduce the velocity of eye movement, which also reduces the occurrence of VIMS.

Bibliography

[1] R. Kennedy and N. Lane, "Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness," *The International Journal of Aviation Psychology*, vol. 3, no. 3, pp. 203–220, 1993.

[2] S. D. Young, B. D. Adelstein, and S. R. Ellis, "Demand Characteristics in Assessing Motion Sickness in a Virtual Environment: Or Does Taking a Motion Sickness Questionnaire Make You Sick?," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 3, pp. 422–428, 2007.

[3] C. Jinjakam, *Study on Simulator Sickness in Immersive Virtual Environment and Proposal for Safety Guidelines*. PhD thesis, Graduate School of Science and Technology Tokai University, 2013.

[4] A. Camm, M. Malik, J. Bigger, G. Breithardt, S. Cerutti, R. Cohen, P. Coumel, E. Fallen, H. Kennedy, R. Kleiger, *et al.*, "Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology," *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.

[5] G. G. Berntson, J. T. Bigger Jr, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone, and M. W. Van der Molen, "Heart rate variability: origins, methods, and interpretive caveats.," *Psychophysiology*, vol. 34, no. 6, pp. 623–648, 1997.

[6] M. Nakagawa, T. Iwao, S. Ishida, H. Yonemochi, T. Fujino, T. Saikawa, and M. Ito, "Circadian rhythm of the signal averaged electrocardiogram and its relation to heart rate variability in healthy subjects," *Heart*, vol. 79, no. 5, pp. 493–496, 1998.

[7] M. Suto and D. Sykes, "StereoMovie Maker - www.stereo.jpn.org," December 2013.

[8] I. Bell, "Roller Coaster 3D Front Seat on-ride HD POV Great Yarmouth Pleasure Beach - www.CoasterForce.Com," December 2013.

[9] R. Fazio, "Element.map stereoscopic project - www.robertofazio.com," December 2013.

[10] M. B. Flanagan, J. G. May, and T. G. Dobie, "The role of vection, eye movements and postural instability in the etiology of motion sickness," *Journal of Vestibular Research*, vol. 14, pp. 335–346, Jan. 2004.

[11] H. Ujike, "Effects of Global Motion Included in Video Movie Provoking an Incident on Visually Induced Motion Sickness," in *Lecture Notes in Computer Science* (R. Shumaker, ed.), vol. 4563, pp. 392–396, Springer Berlin Heidelberg, 2007.

[12] D. L. Sparks, "The brainstem control of saccadic eye movements," *Nature Reviews Neuroscience*, vol. 3, pp. 952–964, Dec. 2002.

[13] O. Hikosaka, Y. Takikawa, and R. Kawagoe, "Role of the Basal Ganglia in the Control of Purposive Saccadic Eye Movements," *Physiological Reviews*, vol. 80, pp. 953–978, July 2000.

[14] S. Moshel, A. Z. Zivotofsky, L. Jin-Rong, R. Engbert, J. Kurths, R. Kliegl, and S. Havlin, "Persistence and phase synchronisation properties of fixational eye movements," vol. 161, no. 1, pp. 207–223–, 2008.

[15] K. Holmqvist, M. Nystrom, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press, 2011.

[16] C. Busettini, G. S. Masson, and F. A. Miles, "Radial optic flow induces vergence eye movements with ultra-short latencies," *Nature*, vol. 390, pp. 512–515, Dec. 1997.

[17] D. E. Angelaki and B. J. M. Hess, "Self-motion-induced eye movements: effects on visual acuity and navigation," *Nat Rev Neurosci*, vol. 6, pp. 966–976, Dec. 2005.

[18] 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. 33,1–30, 2008.

[19] H. Takada and M. Miyao, "Visual Fatigue and Motion Sickness Induced by 3D Video Clip," *Forma*, vol. 27, pp. S67–S76, 2012.

[20] V. K. Gupta, "Motion sickness is linked to nystagmus-related trigeminal brain stem input: a new hypothesis," *Medical Hypotheses*, vol. 64, no. 6, pp. 1177–1181, 2005.

[21] T. Niemann, M. Lappe, A. Buscher, and K.-P. Hoffmann, "Ocular responses to radial optic flow and single accelerated targets in humans," *Vision Research*, vol. 39, no. 7, pp. 1359–1371, 1999.

[22] D. Yang, M. Zhu, and R. W. Hertle, "Version and vergence eye movements in optokinetic nystagmus induced by optic flow," *Journal of Vision*, vol. 6, no. 6, p. 1, 2006.

[23] D. Yang, M. Zhu, C. H. Kim, and R. W. Hertle, "Vergence nystagmus induced by motion in the ground plane: Normal response characteristics," *Vision Research*, vol. 47, no. 9, pp. 1145–1152, 2007.

[24] M. Zhu, R. W. Hertle, and D. Yang, "Relationships between versional and vergent quick phases of the involuntary version-vergence nystagmus," *Journal of Vision*, vol. 8, no. 9, pp. 1–11, 2008.

[25] J. Yang, C. Guo, R. So, and R. Cheung, "Effects of eye fixation on visually induced motion sickness: are they caused by changes in retinal slip velocity?," in *Proceedings of the human factors and ergonomics society 55th annual meething*, pp. 1220–1224, 2011.

Chapter 5

Conclusions

Stereoscopic environment becomes a ubiquitous technology in various areas. As the usage of SE increases, concerns on safety issues and human factors in 3D image technology also increase. One important safety issue in SE is visually induced motion sickness (VIMS).

Previous research works in VIMS used simulator sickness questionnaire (SSQ) as subjective measurement, while objective measurement was observed from electrocardiography (ECG) and 2D point of gaze. Therefore, there was no information about relationship between VIMS, ECG data, and 3D point of gaze in active stereoscopic environment. Furthermore, gaze tracking systems used in previous works were not compatible with active shutter glasses.

In this study, I present a novel 3D gaze tracking system based on optimized geometric method. The 3D gaze tracking system has been validated experimentally. A gaze tracking headgear consisting of dual-camera systems and hot mirrors has been designed and is compatible with consumer-level active shutter glasses. Customized user dependent parameters and 3D calibration using only three calibration points are implemented to estimate 3D point of gaze on stereoscopic display. The experimental results show that the proposed system yields better precision compared to conventional geometric method as empirically proven with average errors 0.83, 0.87, and 1.06 cm in X, Y, and Z dimension, respectively. Compared with subjective depth judgment, the proposed 3D gaze tracking system is more robust in measuring depth of virtual 3D object with various sizes. Statistical analysis of the

experimental results show that the 3D gaze tracking system and user depth perception are more accurate in closer rather than farther depth range.

I also present a novel investigation of visually induced motion sickness (VIMS) in stereoscopic environment using the proposed 3D gaze tracking, electrocardiography (ECG), and simulator sickness questionnaire (SSQ). SSQ is used to obtain general information of VIMS occurrence. ECG and 3D gaze tracking are used to investigate detail and duration of VIMS during exposure of dynamic 3D contents. The investigation of VIMS is performed in two different types of 3D movie with low and high dynamic motions stimuli, respectively.

Two-way ANOVA on SSQ data shows that nausea and disorientation symptoms increase as amount of dynamic motions increase (nausea: $p<0.005$; disorientation: $p<0.05$). Analysis of ECG data shows that horizontal motion (turning left and right) and vertical motion (up and down hill) are effective contributing factors of VIMS. To reduce VIMS, ECG data suggests that viewers of dynamic 3D contents should perform voluntary gaze fixation at one point when experiencing vertical and horizontal motion. The 3D gaze tracking data shows that depth gaze is compressed by sustained forward motion. This unnatural behavior may strengthen visual fatigue and VIMS caused by decoupling of accommodation and vergence in dynamic 3D contents. Furthermore, participant who experiences VIMS tends to have unstable depth gaze than ones who does not experience VIMS. Intense oscillation of depth gaze during provoking scene in dynamic 3D contents can be used as indicator of VIMS occurrence. This result also shows that instead of applying scleral search coil (SSC) or high frame rate binocular eye tracker (>1000 Hz), detection of visually induced motion sickness in stereoscopic environment is sufficient with low frame rate consumer-grade cameras (± 25 Hz).

5.1 Implications of study

The most important contribution of this study for general society is promoting methods for development of user-friendly 3D contents that considers safety issues and human factors. Results of this study suggest that minimizing amount of vertical and horizontal motion with high velocity in 3D contents maybe useful to reduce VIMS. Furthermore,

continuous forward motion for long time should be avoided to prevent strong decoupling between accommodation and vergence. At some crucial scenes (up, down, turn left, turn right), asking viewers of dynamic 3D contents to gaze at one fixation point maybe useful to prevent VIMS. Safety guidelines for stereoscopic 3D technology should include not only contributing factors of visual fatigue, but also visually induced motion sickness. Developers of 3D contents and entertainment companies are then encouraged to adhere those safety guidelines.

5.2 Future works

In future, development of mobile 3D gaze tracking that allows various head movements is important. Based on this study, the proposed 3D gaze tracking system maybe incorporated with optical head tracker or time-of-flight camera to overcome limitation of head movements. Minimizing the size of mobile 3D gaze tracking system using embedded system is also possible. Furthermore, the usage of corneal reflection and pupil position to reduce inaccuracy caused by headgear movement is important if such 3D gaze tracking system is used in standing or moving condition. Mobile 3D gaze tracking is useful to investigate VIMS in large screen immersive virtual environment (IVE).

Prediction and estimation of VIMS occurrence in particular 3D contents are important issues. Such problems can be avoided by development of software that can analyze motions in dynamic 3D contents using computer vision or image processing algorithm. Investigation of maximum rotational angle of horizontal-vertical motion and maximum velocity of forward motion that are safe for viewers of 3D contents are also important. Comparison of horizontal and vertical motion in more structured experiment is also important to understand which motion is more potential as contributing factor of VIMS.

List of Publications

Journal

1. S. Wibirama and K. Hamamoto, "3D Gaze Tracking on Stereoscopic Display using Optimized Geometric Method", *IEEJ Transactions on Electronics, Information, and Systems*, Vol.134, No.3, 2014. pp.345-352. (This paper is related to Chapter 3, section 3.3–3.5)

International Conferences

1. S. Wibirama and K. Hamamoto, "Design and Implementation of Gaze Tracking Headgear for Nvidia 3D Vision®", in *Proceeding of The 5th International Conference on Information Technology and Electrical Engineering (ICITEE) of IEEE*, Yogyakarta, Indonesia, 7-8 October 2013, pp. 84-87. (This paper is related to Chapter 3, section 3.1)

2. S. Wibirama and K. Hamamoto, "3D Gaze Tracking System for Nvidia 3D Vision®", in *Proceeding of The 35th Annual International Conference of the IEEE EMBS*, Osaka, Japan, 3-7 July 2013, pp. 3194-3197. (This paper is related to Chapter 3, section 3.2)

Appendix A

Simulator Sickness Questionnaire

The simulator sickness questionnaire (SSQ) is adapted from:

[1] R.S. Kennedy and N.E. Lane, "Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness", *The International Journal of Aviation Psychology*, Vol.3, No.3, pp.203–220, 1993.

Simulator Sickness Questionnaire

Post-test assessment

Subject # : _____ Date : _____ / _____ / _____
Session : CG / Real Movie Gaze Fixation : Yes / No

■次の中から最もあてはまる所に○をつけて下さい。

(Please circle ○ the most appropriate score according to your feeling.)

質問(Question)	答え(Answer)			
	なし No	ややある Slight	ある Moderate	かなりある Severe
一般的な不快感 (General discomfort)	0	1	2	3
疲労感がある (Fatigue)	0	1	2	3
頭痛がする (Headache)	0	1	2	3
眼が疲れている (Eyestrain)	0	1	2	3
眼の焦点がぼける (Difficulty focusing)	0	1	2	3
唾液の増加 (Increased salivation)	0	1	2	3
発汗する (Sweating)	0	1	2	3
吐き気がする (Nausea)	0	1	2	3
集中できない (Difficulty concentrating)	0	1	2	3
頭が重い (Fullness of head)	0	1	2	3
眼がかすむ (Blurred vision)	0	1	2	3
眩暈感がある[開眼] (Dizzy (eyes open))	0	1	2	3
眩暈感がある[閉眼] (Dizzy (eyes closed))	0	1	2	3
周囲が回転する眩暈 (Vertigo)	0	1	2	3
胃の存在感がある (Stomach awareness)	0	1	2	3
げっぷが出る (Burping)	0	1	2	3

1. If you felt some sickness symptom above, from what reason do you think it came ?
(e.g. : *I did not sleep well last night, I had headache last week, I was hungry, etc)*

_____

_____

_____

2. Do you have any unusual events that occurred during experiment ?

_____

_____

3. Do you have any comment for future improvement ?

_____

_____

********** STOP HERE **********

Appendix B

Informed consent form

Informed consent form was given to all participants before joining experiment. All forms were in Japanese language, explaining procedure of experiment and asking whether participants agree to perform several tasks during experiment. All experiment procedures had been submitted to ethics committee of School of Information and Telecommunication Engineering, Tokai University, Japan.

研究参加への説明書

1. 本研究は，立体映像視認時の注視点を三次元的に計測する装置の開発と，それを用いたシミュレータ酔いと注視点の関連を分析することを目的としています。立体映像視認時のシミュレータ酔いの原因を分析し，より安全な立体映像を作成するための方法を明らかにするための研究です。今回の実験では，CG および実写による立体映像視認時の注視点を計測すると共に，心電図を計測し交感神経・副交感神経の働きを分析します。さらに主観的な酔いの評価としてアンケート調査を行います。この結果から，より安全な立体映像の条件を明確にすることを目指します。

2. 本研究の対象者は大学生を考えており，あなたにご協力を戴く必要があります。実験は，東海大学高輪校舎 1514 研究室で行います。約 5 分間の立体映像視認時の注視点の動きを計測し，立体映像提示前後の酔いの状況をアンケート調査により回答して戴きます。実験の所要時間は 15 分程度です。

3. この研究へのご協力は全くの任意ですので，拒否されてもかまいません。また，研究開始後でも，お申し出頂ければ，いつでもご辞退戴けます。参加の拒否，ご辞退されたことにより，あなたに不利益が生ずることは一切ありません。また，本研究では，心身に悪影響を来すことは無いと考えていますが，不測の事態が生じた場合は直ちに研究担当者に申し出ていただければ，誠意をもって対処いたします。

4. 得られたデータは研究目的以外に使用されることはなく，あなたの氏名等個人が特定できない形で学会に発表し，論文とすることを考えております。全てのデータは個人情報と切り離し，厳重に管理し，研究が全て終了するまで保管いたします。

5. 個人情報管理責任者は東海大学情報通信学部情報メディア学科教授濱本和彦です。他者に漏洩する事がないように細心の注意を払います。

6. 研究実施期間　2013 年 12 月 12 日～2014 年 3 月 31 日

以上，ご説明申しあげましたが，同意後であっても，疑問点があれはいつでも下記の研究責任者にご連絡ください。

本研究に関する説明により，自由意思に基づく同意を得たことを確認いたします。

研究責任者：濱本　和彦　　印
※説明者が研究責任者と異なる場合
説明者氏名：　　印
連絡先
住所：〒108-8619　東京都港区高輪 2-3-23
電話番号：03-3441-1171

研究参加についての同意書

私は，研究計画「立体映像視認時の注視点の三次元計測とそのシミュレータ酔いとの関連に関する研究」に関する以下の事項について，詳しく説明を受けました。

【研究内容】
立体映像視認時の注視点の動きとその時の心電図の変化およびアンケート調査により，シミュレータ酔いの度合いと立体映像，注視点の動きの関係について明らかにする。
【研究実施期間】
2013 年 12 月 12 日～2014 年 3 月 31 日

【説明事項】
《研究実施期間中》
□ 研究への参加は任意であること
□ この研究の目的
□ 研究の方法
□ 研究の実施場所
□ 研究における危険および不快な状況
□ 情報およびデータ等の収集方法について
□ あなたに研究参加をお願いする理由
□ 研究に対する問い合わせ・苦情等の連絡先

《収集したデータ等について》
【保管期間】
研究終了時（2014 年度）

□ 個人情報の取り扱いについて
□ データ等の取り扱いについて（保管期間等）
□ データ等の利用について
これらの事項を確認したうえで，自らの自由意思により本研究への参加に同意いたします。

年　　月　　日

参加者氏名：＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿