REAL-TIME 3D EYE MOVEMENTS TRACKING AND VISUALIZATION USING DUAL CAMERAS ACQUISITION

SUNU WIBIRAMA

A THESIS SUBMITTED IN PARTIAL FULFILLMENT

OF THE REQUIREMENT FOR THE DEGREE OF

MASTER OF ENGINEERING IN ELECTRONICS ENGINEERING

FACULTY OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

2010

KMITL-2010-EN-M-040-035

หัวข้อวิทยานิพนธ์	ระบบติดตามการเคลื่อนไหวของดวงตาแบบสามมิติโดยแสดงภาพเป็นเวลาจริง ด้วยระบบกล้องจับภาพ 2 ตัว
นักศึกษา	นายซูนุ วีบิรามะ
รหัสนักศึกษา	51060420
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมอิเล็กทรอนิกส์
พ.ศ.	2553
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผศ.ดร.สุพันธุ์ ตั้งจิตกุศลมั่น
อาจารย์ที่ปรึกษาวิทยานิพนธ์ร่วม	รศ.ดร.ชูชาติ ปิณฑวิรุจน์

บทคัดย่อ

ในปัจจุบันมีงานวิจัยที่ศึกษาเกี่ยวกับระบบติดตามการเคลื่อนไหวของดวงตาแบบสามมิติโดยการจับภาพวิดีโออย่างแพร่หลาย แต่ทว่าอุปกรณ์ที่ใช้ในการติดตามการเคลื่อนไหวของดวงตานั้นยังคงมีราคาแพงอยู่มาก อีกทั้งยังคงพบปัญหาในด้านความถูกต้องแม่นยำของอุปกรณ์และความสามารถในการแสดงภาพแบบเวลาจริง ดังนั้น งานวิจัยนี้จึงมุ่งเน้นไปที่การพัฒนาระบบติดตามการเคลื่อนไหวของดวงตาที่สามารถแสดงภาพเป็น เวลาจริง และมีต้นทุนการผลิตอุปกรณ์ต่ำ โดยใช้กล้องซีซีดีจำนวน 2 ตัวงานวิจัยนี้ได้-ประยุกต์ใช้ Center of Mass (COM) และ Template Matching Algorithm ในการหาพิกัดสองมิติของจุดศูนย์กลางของรูม่านตาและม่านตา ตามลำดับ แล้วทำการหาพิกัดสามมิติด้วยเทคนิค Direct Linear Transformation (DLT) จากนั้น ได้ทำการแก้ไขปรับปรุงความผิดเพี้ยนของภาพที่เกิดขึ้นจากความบิดเบือนของเลนส์กล้องด้วยเทคนิค Lens Undistortion Algorithm เพื่อให้ได้ผลการวัดตำแหน่งของดวงตาในทั้งสามระนาบ ได้แก่ แนวนอน แนวตั้ง และมุมในการบิดหมุนได้อย่างถูกต้องแม่นยำนั่นเอง จากผลการทดลองพบว่า สามารถติดตามการเคลื่อนไหวของดวงตาได้อย่างแม่นยำโดยมีความผิดพลาดที่เกิดขึ้นเนื่องจากการวัดระยะรัศมีของดวงตาคลาดเคลื่อนในระนาบ แนวนอน แนวตั้ง และมุมบิดหมุน เฉลี่ยเพียง 0.15, 0.14, และ 0.20 องศา ตามลำดับ และเมื่อทำการปรับปรุงคุณภาพของภาพเนื่องจากความบิดเบือนของเลนส์แล้วจะสามารถลดความผิดพลาดในแต่ละระนาบได้ถึงร้อยละ 53.13, 63.16, และ 57.45 ตามลำดับ นอกจากนั้น ระบบยังสามารถแสดงการเคลื่อนไหวของลูกตาเป็นแบบเวลาจริงได้อีกด้วย ซึ่งจะช่วยอำนวยความสะดวกให้แพทย์ในการวิเคราะห์การเคลื่อนไหวของดวงตาได้ง่ายขึ้น อีกทั้งยังสามารถนำอุปกรณ์ที่ประดิษฐ์ขึ้นในงานวิจัยนี้ไปใช้ในการตรวจวินิจฉัยความผิดปกติในคนไข้ที่เป็น โรค Benign Paroxysmal Positional Vertigo ได้อย่างถูกต้องแม่นยำอีกด้วย

Thesis Title	Real-Time 3D Eye Movements Tracking and Visualization using Dual Cameras Acquisition
Student	Mr. Sunu Wibirama
Student I.D	51060420
Degree	Master of Engineering
Program	Electronics Engineering
Year	2010
Thesis Advisor	Asst. Prof. Dr. Supan Tungjitkusolmun
Thesis Co-Advisor	Assoc. Prof. Dr. Chuchart Pintavirooj

ABSTRACT

Nowadays, growing research developments on diagnostic eye tracker have been focused on the video-based three-dimensional (3D) eye movements tracking. Nevertheless, acquisition system's cost factor, accuracy of measurements, and 3D visualization of tracking results are challenging issues needed to be overcome in practical applications. In this thesis, an inexpensive real-time 3D eye movements tracking and visualization system using two charge-coupled device (CCD) cameras has been proposed. Center of mass (COM) and template matching algorithms have been employed to detect two-dimensional (2D) coordinates of pupil and iris striation center. Direct Linear Transformation (DLT) and lens undistortion algorithms are utilized to improve the accuracy of tracking results. Real-time 3D visualization based on tracking results can then help clinicians understand the characteristics of eye movements. Experimental results demonstrated that our system had high accuracy as the average errors in horizontal, vertical, and torsional angular positions were confined to 0.15°, 0.14°, and 0.20°, respectively. Additionally, the lens undistortion algorithm was able to reduce the errors in horizontal, vertical, and torsional angular positions up to 53.13%, 63.16%, and 57.45%, respectively. Inaccurately measuring radius of the eye affects horizontal and vertical angular positions while torsional angular position is considered more robust to radius estimation error. Real-time implementation shows that our system can be used in clinical routine to detect either voluntary or involuntary human eye movements, particularly symptom of Benign Paroxysmal Positional Vertigo.

ACKNOWLEDGMENTS

All praise, glory, and gratitude only for Allah who said in the Holy Qur'an that "He who taught (the use of) the pen. Taught the man that which he knew not". Peace be upon the Prophet Muhammad, his family, his companions, and all those who follow him until the Judgement Day.

First of all, I would like to express my deepest gratitude to my thesis advisor Assistant Professor Dr. Supan Tungjitkusolmun and my thesis co-advisor Associate Professor Dr. Chuchart Pintavirooj for all invaluable advice, personal attention, and continuous encouragement throughput my master program at KMITL. It was a great experience learning with both of them. I am also deeply grateful to my thesis committee members for their involvement and the time they spared to review this thesis.

My sincere gratitude goes also to Japanese Professor Support, Professor Kazuhiko Hamamoto, Tokai University, Japan, for giving me suggestion and encouragement to improve my research skill.

Special thanks to Japan International Cooperation Agency (JICA) under AUN/Seed-Net Project for their financial support and chances to publish, present, and contribute in international conferences. My appreciation is extended to Head of Electronics Department of KMITL, all lecturers, professors, and supporting staff for their strong support and encouragement to complete the study.

Finally, I would like to send my best regards to my parents, my younger sister, all BIOSIS laboratory members, AUN/Seed-Net officers and students, Thais and Indonesians colleagues, and everyone who helped me to get this work done.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure

Page

Figure Page

Figure Page

LIST OF ABBREVIATIONS

AC	Alternating Current
BPPV	Benign Paroxysmal Positional Vertigo
CCD	Charge-Coupled Device
COM	Center of Mass
CPU	Central Processing Unit
CRT	Cathode Ray Tube
DC	Direct Current
DLT	Direct Linear Transformation
DOF	Degree of Freedom
EOG	Electro-Oculography
FFT	Fast Fourier Transform
HVT	Horizontal-Vertical-Torsional
IO	Inferior Oblique
IR	Inferior Rectus
LCD	Liquid Crystal Display
LED	Light-Emitting Diode
LR	Lateral Rectus
MOA	Minute of Arc

MR	Medial Rectus
NTSC	National Television System Committee
OPD	Outpatient Department
PAL	Phase Alternation Line
RAM	Random Access Memory
ROI	Region of Interest
RPE	Retinal Pigment Epithelium
SO	Superior Oblique
SR	Superior Rectus
SVD	Singular Value Decomposition
USB	Universal Serial Bus
VOG	Video-Oculography
VOR	Vestibulo-Ocular Reflex
VR	Virtual Reality
VSR	Vestibulo-Spinal Reflex
VTM	Video Torsion Measurement

Chapter 1

Introduction

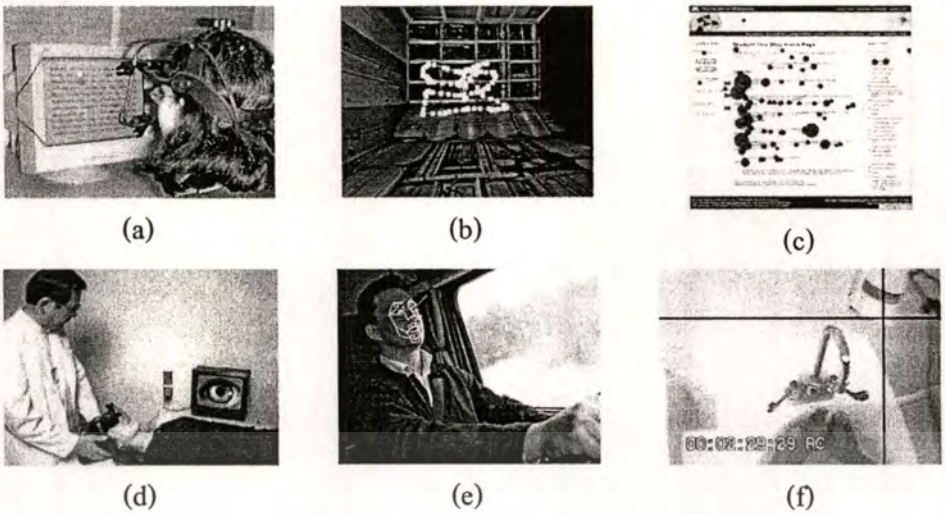
The measurement device most often used for measuring or tracking eye movements is generally known as an eye tracker. Nowadays, eye tracking research is entering its fourth era. Rayner [1] summarizes the characteristics of eye tracking's first three eras: (1879–1920) defined by the discovery of many basic eye movements facts, including saccadic suppression, saccade latency, and the size of the perceptual span; (1930–1958) characterized by a more applied research focus, coinciding with the behaviorist movement in experimental psychology; (1970–1998) marked by improvements in eye movement recording systems facilitating increasingly accurate and easily obtained measurements. A wide variety of eye tracking applications now exist, which can be broadly categorized from a system analysis point of view as interactive tool or diagnostic tool.

Supported by advancements in computational power, richness of graphical displays, and robustness of interactive styles-as an interface modality, the eye tracker serves as a powerful input device in a host of visually-mediated applications. An interactive system must respond to or interact with the user based on the observed eye movements. Such interactive systems may fall into two application subtypes: selective and gaze-contingents. Selective systems use the point of gaze as a pointing device such as the mouse, whereas gaze-contingent systems exploit knowledge of the user's gaze to facilitate the rapid rendering of complex displays (e.g., graphical environments).

In its diagnostic role, eye tracker provides objective and quantitative evidence of the user's visual process. In this capacity, eye movements are generally recorded to ascertain the user's attentional patterns over a given stimulus. Moreover, the stimulus being displayed does not usually need to change or react to the viewer's gaze. In this scenario, the eye tracker is used to record eye movements for offline assessment of the viewer's gaze during the experiment.

Figure 1.1 shows several research areas related to eye movements tracking:

- Human and computer interaction: eye tracker is used as a replacement of common controlling device such as mouse, keyboard, and remote control (Figure 1.1 (a)).
- Virtual Reality (VR): eye tracker is used as a visual inspection tool in virtual environment (Figure 1.1 (b))



(a) (b) (c)

(d) (e) (f)

Figure 1.1 Example of eye tracking-based research areas

- Web usability research: eye movements tracking offers the ability to analyze user interaction between clicks. (Figure 1.1 (c))
- Vestibulo-ocular research: eye movements tracking is a convenient clinical way to analyze and diagnose vestibular and oculomotor system behaviour in human or animal (Figure 1.1 (d)).
- Smart vehicle: combined with face tracking, eye movements tracking acts as a tool for driver monitoring and smart car controlling (Figure 1.1 (e)).
- Natural tasks: eye movements tracking is used as a tool to analyze natural human behaviour during several tasks and observe the monitoring action that the eyes perform, such as hand-washing or tea-making activities (Figure 1.1 (f)).

1.1 Motivation

As a diagnostic tool, eye movements tracking can be used in vestibulo-ocular research. Several eye movements tracking methods have been proposed and put into commercial use. Nevertheless, there are still some issues regarding availability, portability, measurement accuracy and data visualization which inspire researchers to find more convenient methods.

Current markets for diagnostic eye movements tracking include: hospital and medical experts; corporations; universities; and others specialized in the field of engineering, medical science, animal science and neuroscience. While there have been significant advances in diagnostic eye movement tracking technology, the cost of commercial systems remains high. Thus, eye tracker device is often not used simply because of the cost factor, making it a technology that is used only for niche applications. The high cost of commercial systems (US$1000 to US$30,000) [2] have led to numerous

efforts to build user-consumer grade eye movements tracking as seen in [3-6]. However, some low-cost methods require complicated installation procedure and considered impossible to use in portable mode. This thesis proposes an inexpensive portable head mounted eye movements tracking system which can gain wider acceptance.

A precise and accurate measurement of eye movements is also a challenging issue in vestibulo-ocular research. Current method of eye movements tracking using video processing suffers from inaccurate measurement caused by lens aberration and geometric distortion. Some researchers introduce calibration procedures to eliminate inaccuracy of measurement by manual adjustment of camera position related to defined reference position on the eye [5, 7] or asking the patient to fixate the gaze on several calibration points in order to get correct parameter for geometric compensation algorithm in every experiment process [8]. Those proposed calibration procedures, although yielding accurate measurement, will require distinctive setup for each patient which can be impractical and time consuming. Developing a simple procedure to eliminate the measurement error is therefore another motivation for this study.

Presently, all reported eye movements tracking methods do not provide a real-time 3D visualization information based on measurement results which is essential for accurate diagnosis and understanding of how the voluntary and spontaneous eye movements function. In addition, huge advantages in the use of 3D visualization language as a tool to create a 3D model have increased the interest in simulating and modeling human vision system for clinical and educational purposes [9-11]. Hence, a real-time 3D eye movements visualization based on tracking result is also incorporated into our proposed system.

1.2 Research Objectives and Methodologies

Recently, growing research developments on diagnostic eye tracker have been focused on 3D eye movements tracking using video processing. The horizontal and vertical eye movements can be tracked by pupil tracking method and/or corneal reflection method. To determine ocular torsion, most currently available Video-Oculography (VOG) systems either track two or more landmarks on the eye or employ some variations of pattern matching which involves measuring and tracking changes in iris striation images. Nevertheless, the portability and performance of these methods still need to be improved. The main objective of this study is to develop a new precise real-time 3D eye movements tracking and visualization using dual cameras acquisition. The head-mounted model is chosen to address with practicality and portability issues.

In this study, we have designed a dual cameras system mounted on a consumer-grade welding goggle to track a human eye (Chapter 3). The dual cameras system consists of two parallel mini CCD cameras. The analog to digital image converter has been used on each camera to convert captured frame for image processing purpose. The algorithm for image processing can be explained as follows: for first time usage, camera calibration procedure using chessboard pattern must be applied on each camera to gain camera parameters and lens distortion coefficients. Next, eyeball radius is obtained by direct measurement. Iris striation is gained from captured eye image. The dual cameras system is used to capture simultaneously two images of one eye. Pupil detection based on center of mass (COM) algorithm is implemented to obtain the 2D coordinates of pupil center. Template matching based on correlation-coefficient is then implemented to track iris striation location on each image frame. The 2D coordinates of pupil center and iris striation from each camera are then submitted to 3D coordinates extraction algorithm based on Direct Linear Transformation (DLT). The results of 3D coordinates extraction process are used to determine horizontal, vertical, torsional movements of the eye. Finally, we represent the real-time data in both eye movements trajectory and 3D visualization. All steps of our algorithm are depicted in Figure 1.2.

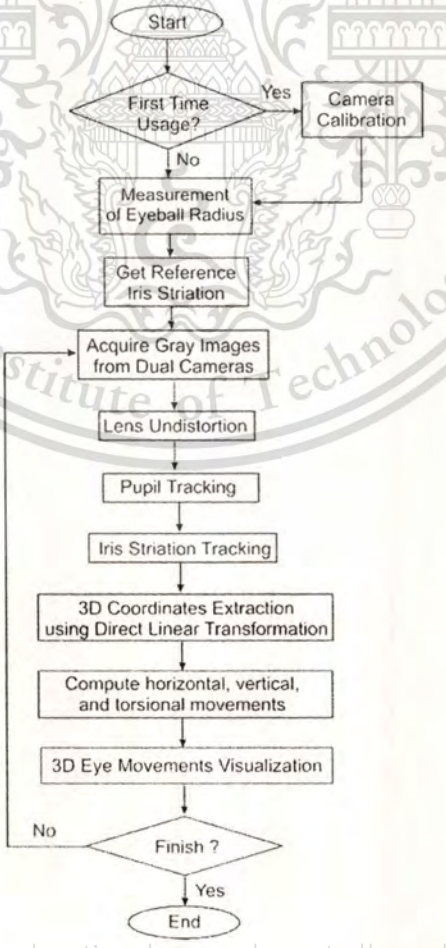


Figure 1.2 Flowchart of 3D eye movements tracking and visualization

To validate the system, we measure the accuracy using artificial eye mounted on gimbal system [5]. The artificial eye is moved on several positions based on 5 x 5 calibration points. Horizontal, vertical and torsional coordinates are then measured for each position. The computational results and artificial eye positions are then compared to observe the system's accuracy. The system is also tested in human subjects to show real-time and practical implementation.

1.3 Thesis Outline

The remainder of this thesis is organized as follows:

Chapter 2 presents an overview of human eye anatomy, vestibular system, several types of eye movements disorder, and reviews of diagnostic eye movements tracking. The reviews include hardware and algorithms for detecting the eye movements, advantages and disadvantages of each method.

Chapter 3 describes the eye movements tracking methods. In this chapter, we present hardware configuration of our eye movements tracking system and also image processing algorithm used to compute 3D eye movements. Experiment procedure will be described briefly in this chapter.

Chapter 4 provides experimental results. The performance of the system is observed using artificial eye and gimbal system. We also evaluate the effect of lens aberration toward the accuracy of eye movements measurement. The system latency will be discussed in this chapter. Experimental results on real human subjects are presented to show real-time and practical implementation.

Chapter 5 summarizes the main contributions of this thesis and suggestion for future work.

Chapter 2

Literature Reviews

This chapter presents an introduction about vestibulo-ocular system, eye movements system, eye movements disorder, and reviews of diagnostic eye movements tracking. For further reading and more information, the reader can refer to the book clinical ophthalmology [12], eye movements disorder [13], and eye tracking methodology [14].

2.1 An Overview of Vestibulo-Ocular System

Figure 2.1 presents an overview of the inputs and outputs characterizing human movements. Part A consists of inputs for human movements generated by hearing system, vision system, balance system, proprioception (the sense of the relative position of neighbouring parts of the body), smell, and taste. These inputs will give stimulations to central processing system (part B). Human brain will produce necessary outputs transformed into head, eye, and limb movements (part C). Furthermore, these outputs will help part A produce the next inputs for human brain. In this study, we limit our review on the relation between balance (vestibular) system and eye (ocular) movements.

Eye movements are one of the most accessible human characteristics for scientific investigation. The eye is an object with negligible inertia and only three pairs of extraocular muscles control its orientation and movement.

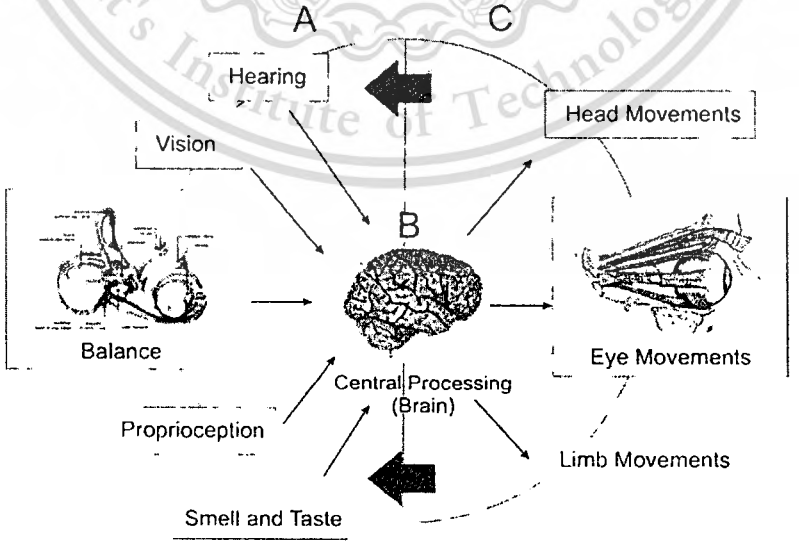


Figure 2.1 Diagram of inputs and outputs characterizing human movements

An eye essentially displays a ball-in-socket behavior and is thus restricted to only 3DOFs (rotatory degree of freedoms). By comparison, 24 muscles and tendons have to be taken into consideration when simple walking movements in the 2D sagittal plane are modeled [15].

The second reason for concentrating on the vestibulo-ocular system is the tremendous importance of eye movements in the diagnosis of neurological, ontological, and ophthalmologic disorders. Since eye movements control involves many areas of the brain, any improvement in our understanding of the balance system and of the execution of eye movements can have a direct effect on the understanding, diagnosis, and treatment of a number of diseases.

2.1.1 Anatomy of Human Eye

The eye is suspended in a cone-shaped orbit by a fibrous sac of fascia called Tenon's capsule. The eye is rotated by six muscles: four rectus muscles and two oblique muscles, as shown in Figure 2.2. The four recti and the superior oblique (SO) arise from the apex of the orbit. The inferior oblique (IO) muscle arises from the inferior nasal aspect of the orbit. The four rectus muscles insert into the sclera anterior to the equator of the globe: the medial rectus (MR) muscle on the nasal side, the lateral rectus (LR) muscle on the temporal side, the superior rectus (SR) muscle on the superior side and the inferior rectus (IR) muscle on the inferior side. The SO and IO muscles approach the globe from its anterior and medial aspect and insert posterior to the equator of the globe. The SO muscle first passes through the trochlea before inserting on the superior side of the globe. The IO inserts on the temporal side. The tendons of the rectus extraocular muscles pass through sleeve-like pulleys that lie within the peripheral Tenon's capsule. These pulleys limit side-slip movement of the rectus muscles during eye rotations [16].

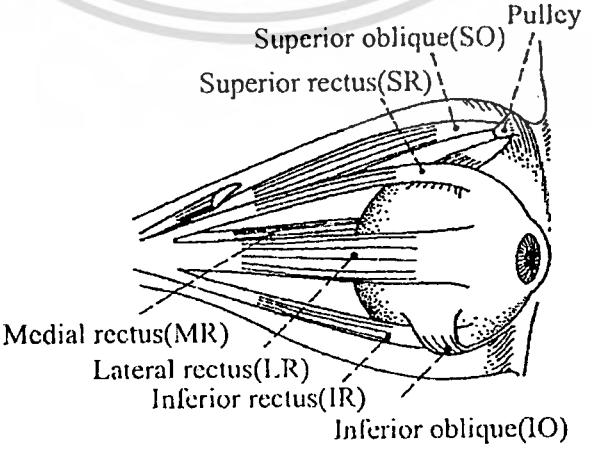


Figure 2.2 Eyeball with six muscles [17]

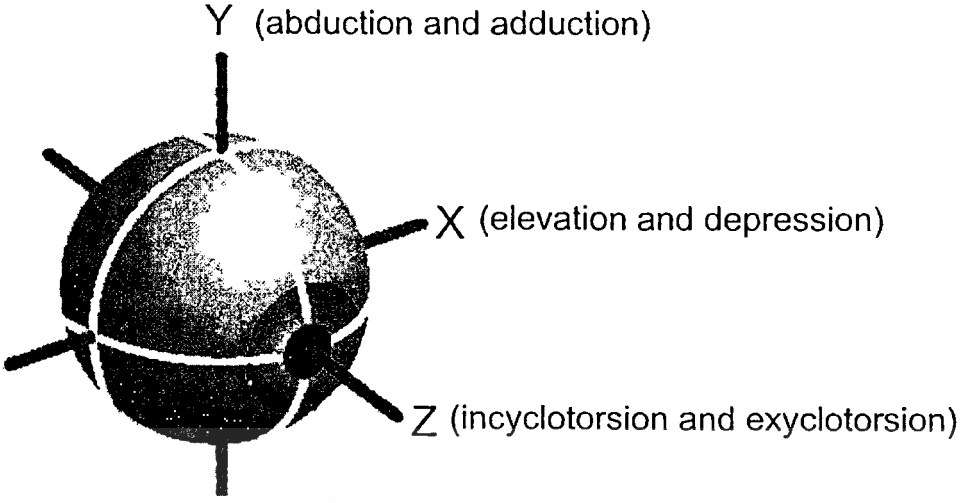


Figure 2.3 Eye movements rotation on three axes [13]

Figure 2.3 shows how the single eye rotates related to three axes. Abduction occurs about the *y*-axis and is away from the median plane. Adduction occurs about the *y*-axis and is toward the median plane. Elevation occurs about the *x*-axis and is an upward rotation of the eye. Depression occurs about the *x*-axis and is a downward rotation of the eye. Incyclotorsion occurs about the *z*-axis so that the upper pole of the eye rotates toward the median plane. Excyclotorsion occurs about the *z*-axis so that the upper pole of the eye rotates away from the median plane. Table 2.1. explain several eye movements produced by the six extraocular muscles. The LR and MR muscles will produce abduction and adduction about the *y*-axis. The SR and IR muscles generate elevation and depression about the *x*-axis. Finally, the SO and IO produce incyclotorsion and exyclotorsion related to *z*-axis.

Table 2.1 Eye movements produced by six extraocular muscles

Extraocular Muscles	**Eye Movements**
Lateral rectus (LR)	Abduction
Medial rectus (MR)	Adduction
Superior rectus (SR)	Elevation
Inferior rectus (IR)	Depression
Superior Oblique (SO)	Incyclotorsion
Inferior oblique (IO)	Excyclotorsion

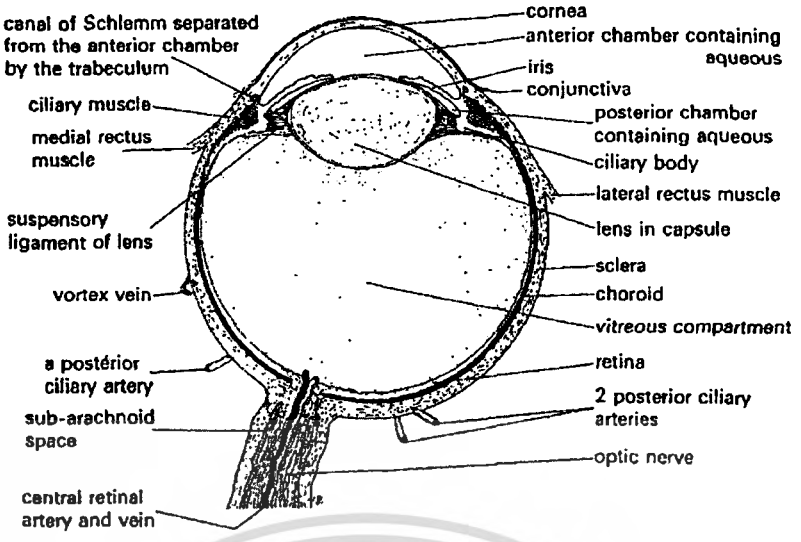


Figure 2.4 The horizontal view of eyeball

The horizontal view of eyeball is depicted in Figure 2.4 [12]. The axial length of the normally sighted eye is approximately 24mm. Measuring posteriorly along the surface from the *limbus*, the anterior termination of the sensory *retina* lies at 8mm, the equator at 16mm and the posterior pole at 32mm. The eyeball has three main layers. The outer fibrous supporting coat in front, is the clear *cornea* which is continuous with the white opaque *sclera* behind. The middle vascular coat or *uvea* consists of the *choroid,* the *ciliary body* and the *iris* which has a central opening or *pupil.* The inner sensory coat, *the retina,* has a multi-cell layered neural membrane and a single celled outer membrane, the *pigment epithelium.* A fenestrated opening in the *sclera* 1.5mm in diameter and 3mm medial to the posterior pole transmits the fibres of the optic nerve, mainly the axons of the ganglion cells of the retina. The *lens* is a transparent structure, suspended immediately behind the iris by fine fibres, forming the *zonule* or suspensory ligament, which runs from the surface of the ciliary body to the periphery of the lens [12].

2.1.2 The Vestibular System

The vestibular system provides a short latency connection between our head movements and the compensatory eye movements to keep a clear vision [13]. The central processing system, composed of vestibular nuclei and cerebellum, processes information from vestibular system in conjunction with other sensory inputs (somatosensory and visual) providing accurate information about the position and movement of the head in space. The outputs of central processing system go to the six ocular muscles and to the spinal cord to serve two vestibular reflexes: the Vestibulo-Ocular Reflex (VOR) and the Vestibulo-Spinal Reflex (VSR). The primary function of the VOR is to maintain gaze stability during head movements while the VSR affects whole-body equilibrium by facilitating and inhibiting

skeletal extensor muscle activity. Through very short latency connections between the vestibular system, central processing system, and the six extraocular muscles, head movements can be compensated within as little as 5-8ms.

Figure 2.5 shows the anatomy of vestibular system. The vestibular system is composed of specialized hair cells, three semicircular canals, and otoliths. The hair cell is the basic sensory element of vestibular system. Present in the sacculus, utricle, and the cristae of the semicircular canals, the hair cells transduce mechanical force into electrical nerve action potentials. The semicircular canals are approximately orthogonal to each other, and are called the horizontal lateral semicircular canal, the superior semicircular canal, and the posterior semicircular canal. The primary function of the semicircular canals is to sense angular acceleration of the head. The semicircular canals are maximally sensitive to high frequency motion, such as motion that occurs during locomotion. Each canal is maximally sensitive to movement in the plane of that canal. The otoliths consist of utricle and sacculus. The otoliths are responsible for detecting linear acceleration of the head. The sacculus detects vertical linear accelerations of the head in the sagittal plane and the utricle responds to horizontal plane linear accelerations. The information from the utricle dominates the eye movement response, whereas information from the saccule seems to be mainly used for postural control [18]. The otoliths are also sensitive to tilts of the head with respect to gravity. The otoliths are most sensitive to lower-frequency motion, such as motion that occurs while standing in place.

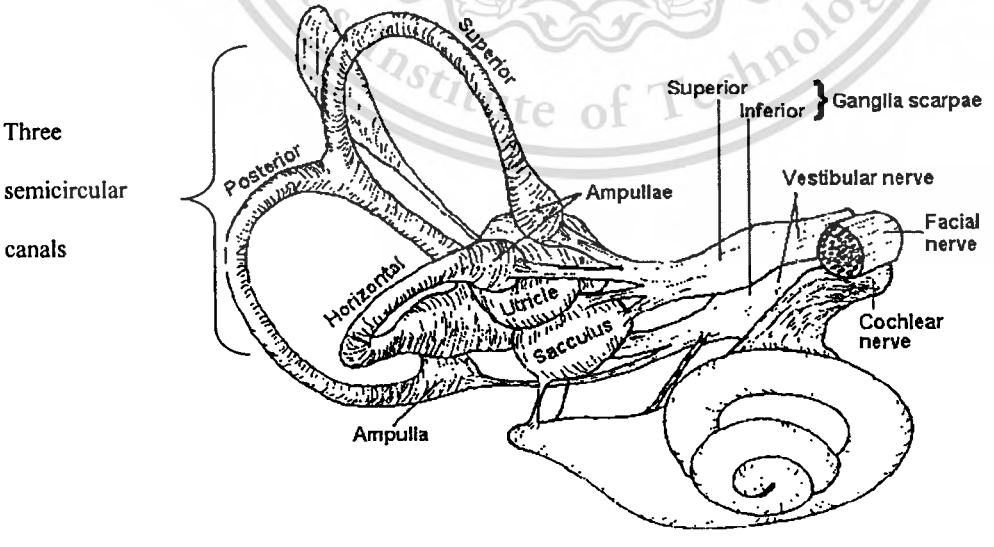


Figure 2.5 Anatomy of vestibular system

2.2 The Eye Movements System

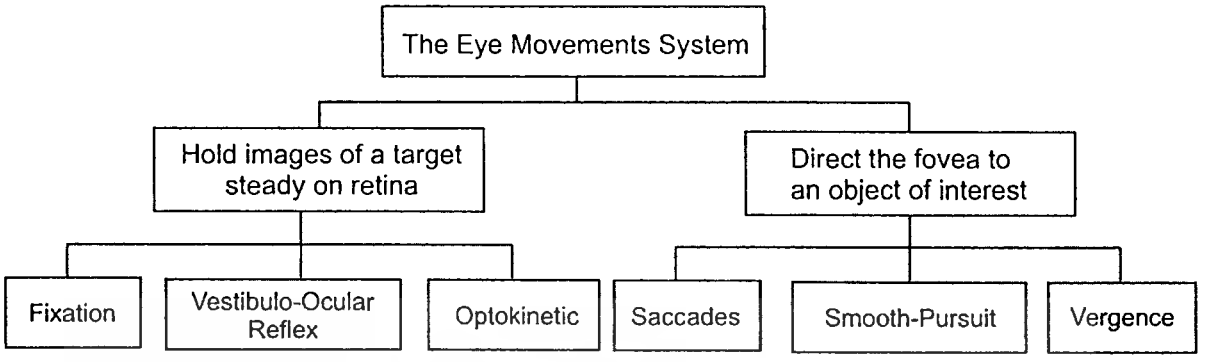


Figure 2.6 The eye movements system

Figure 2.6 shows the eye movements system that can be functionally divided into those that hold images of a target steady on the retina and those that direct the fovea onto an object of interest. The former category includes (1) the fixation system; (2) the Vestibulo-Ocular Reflex (VOR); and (3) the optokinetic system. The latter category includes (1) the saccadic system; (2) the smooth pursuit system; and (3) the vergence system. Small amplitude of eye movement can be measured by *a minute of arc* or *arcminute* (MOA) unit. MOA is a unit of angular measurement, equal to one sixtieth (1/60) of one degree. Since one degree is one three hundred sixtieth (1/360) of a circle, 1MOA is 1/21,600 of the amount of arc in a closed circle.

The fixation system holds the image of a stationary object on the fovea when the head is immobile. Reading involves fixating on a successive location across the page or screen. Visual fixation is never perfectly steady. Fixational eye movements occur involuntarily. The term "fixation" can also be used to refer to the point in time and space of focus rather than to the act of fixating. Normal fixation consists of three distinct types of physiological miniature movements that are not detectable by the naked eye: microsaccades, microdrift, and microtremor. Microsaccades are miniature saccades that have amplitude of less than 26MOA, with average amplitude of 6MOA. They occur at a mean frequency of approximately 120Hz. Microsaccades have no known function and are considered superfluous to visual perception. Microdrift consists of smooth eye movements that occur at a velocity of less than 20MOA per second. They are necessary to prevent the image of a stable object from fading. Microtremor is continuous, high-frequency ocular motor activity that underlies both microdrift and microsaccades. Microtremor occurs at a frequency of 50-100Hz. Its average amplitude is <1MOA (usually 5-30sec of arc) and is much smaller than the amplitude of microsaccades.

The Vestibulo-Ocular Reflex (VOR) stabilizes retinal images during head motion by counter-rotating the eyes at the same speed as the head but in the opposite direction. Information about head motion passes from the vestibular sensors in the inner ear to the VOR circuitry within the brainstem, which computes an appropriate eye velocity command. The eyes, confined in their bony orbits, normally do not change position, and their motion relative to the head is restricted to a change in orientation. However, the head can both change position and orientation relative to space. Thus, the function of the VOR is to generate eye orientation that best compensates for changes in position and orientation of the head. Because the drive for this reflex is vestibular rather than visual, it operates even in darkness.

Optokinetic eye movement is induced reflexively by motion of a large visual scene, which causes an illusionary sensation of self-rotation (circularvection) in the opposite direction. The function of the optokinetic system is to supplement the angular VOR. Whereas the angular VOR responds best to brief, high-frequency head rotation, the optokinetic system maintains retinal image stability during sustained, low-frequency rotation.

Saccades are fast conjugate eye movements that move both eyes quickly in the same direction, so that the image of an object of interest is brought on the fovea. Saccades can be made not only toward visual targets, but also toward auditory and tactile stimuli, as well as toward memorized targets. Saccades can be generated reflexively, and they are responsible for resetting the eyes back to the mid-orbital position during vestibulo-ocular or optokinetic stimulation. Saccades need to be fast to get the eyes on the target as soon as possible. Saccades are the fastest type of eye movements and they are among the fastest movements that the body can make. Saccade speed is not under voluntary control but depends on the size of the movement, with larger saccades attaining higher peak velocities. It has been estimated that human eye yields more than 100,000 saccades per day.

Smooth pursuit consists of conjugate eye movements that allow both eyes to smoothly track a slowly moving object so that its image is kept on the foveae. For example, smooth pursuit eye movements are used when the eyes track a child on a swing. Humans have both smooth pursuit and optokinetic eye movements, but pursuit predominates. When the eyes track a small, moving object against a detailed stationary background, such as a bird flying against a background of leaves, the optokinetic system will try to hold the gaze on the stationary background, but it is overridden by pursuit.

Vergence eye movements shift the gaze point between near and far, such that the image of a target is maintained simultaneously on both fovea. Unlike other eye movement systems, vergence movements are disjunctive, meaning that the eyes move in opposite directions. To move from a far to a near target, the eyes converge (i.e., rotate toward the nose) so that the lines of sight of the two eyes intersect at the target. To aim at a target farther away, the eyes diverge (i.e., rotate toward the temples). When the target is located at optical infinity, the lines of sight are parallel. During deep sleep, deep anesthesia, and coma, the eyes diverge beyond parallel, indicating that eye alignment is normally actively maintained by the brain because the orbits, in which the eyeballs are located, are divergent. Vergence eye movements are very slow, lasting 1s or longer. One reason for this may be that vergence, unlike saccades, is driven by visual feedback, which normally takes at least 80ms. Another reason may be that the speed of vergence movements is limited by how fast the lenses change shape (accommodation) and how fast the pupils constrict [13].

2.3 The Eye Movements Disorder

This section explains some eye movements disorders. The eye movements disorders comprise of nystagmus and saccadic dyskinesia. Nystagmus consists of involuntary oscillations of the eyes that are initiated by slow eye movements. Figure 2.7 shows several types of nystagmus. If both phases are slow eye movements, it is called pendular nystagmus, as shown in Figure 2.7 (A). If one phase is a saccade (quick phase), which alternates with a slow phase, it is called jerk nystagmus, as shown in Figure 2.7 (B, C, and D). Although the fundamental defect in nystagmus is an imbalance of slow eye movements that drives the eyes off their target, the direction of nystagmus is conventionally named in the direction of the corrective quick phases that return the eyes toward their target. The slow-eye-movement imbalance may due to defects in the vestibulo-ocular, smooth pursuit, optokinetic, or gaze-holding systems, or rarely in the vergence system.

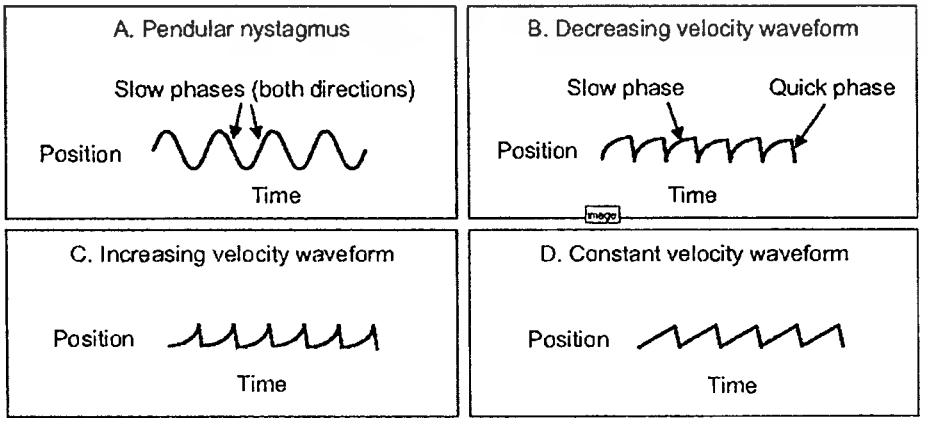


Figure 2.7 Several types of nystagmus [13]

Saccadic dyskinesia includes saccadic intrusions and oscillations. These are abnormal fast eye movements that drive the eyes away from their target and disrupt visual fixation. Involuntary fast eye movements that intermittently take the fovea off a target are called saccadic intrusions, whereas sustained abnormal fast eye movements are called saccadic oscillations. Some saccadic intrusions and oscillations have an interval between sequential saccades (e.g., square wave jerks, macro-square wave jerks, square wave oscillations, macrosaccadic oscillations), whereas others do not. The presence of an intersaccadic interval signifies the integrity of both pause cells that stop saccades and the neural integrator that sustains eye position between saccades. Saccadic dyskinesia usually occurs with lesions of the cerebellum or brainstem.

Figure 2.8 shows several types of saccadic intrusions and oscillations. Figure 2.8 (A) shows square wave jerks. This type consists of a small saccade of 0.5-3^{o} that takes the eye away from fixation, followed by a saccade that returns the eye back to the fixation after about 200ms. Figure 2.8 (B) shows saccadic pulse. This type of saccadic dyskinesia has some burst of saccades with defective steps of innervations. Figure 2.8 (C) shows square wave oscillations. This type is similar to square wave jerks, but they occur continuously rather than sporadically. Figure 2.8 (D) shows macrosaccadic oscillations. This type consists of a series of large saccades that straddle fixation (i.e., passing from one side to the other side of the target), overshooting it each time without foveation of the target. The intersaccadic interval of macrosaccadic oscillations is 200ms when the eye is stationary in the eccentric position.

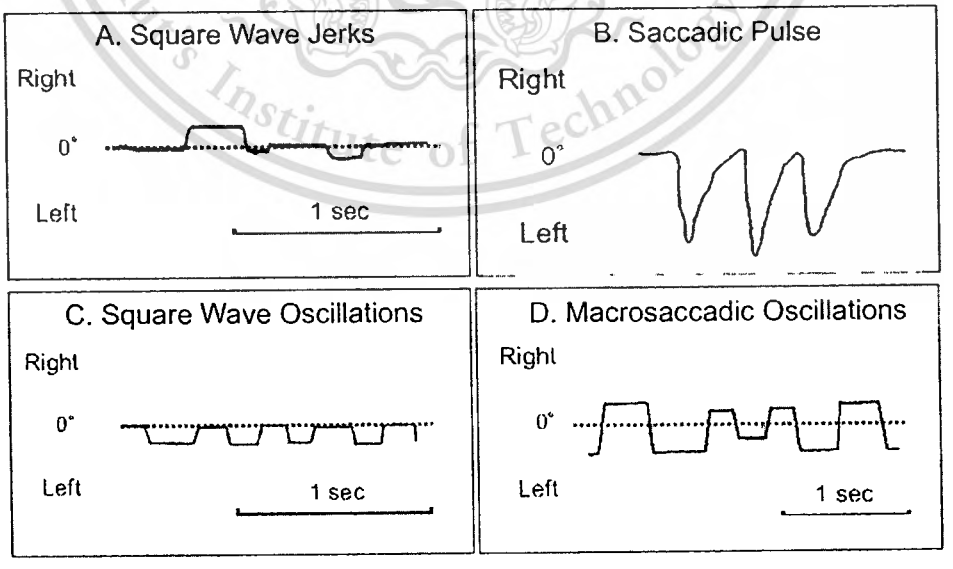


Figure 2.8 Several types of saccadic dyskinesia [13]

2.4 Reviews of Diagnostic Eye Movements Tracking

This section presents reviews of several widely-used diagnostic eye movements tracking. The reviews include the usage of Electro-Oculography as the oldest method to measure the eye movements, followed by Scleral Search Coil method, and Video-Oculography. We also explain a brief introduction of our proposed method in the Video-Oculography sub-section.

2.4.1 Electro-Oculography

The Electro-Oculography (EOG) is an electrophysiological test of function of the outer retina and retinal pigment epithelium (RPE) in which the change in the electrical potential between the cornea and ocular fundus is recorded during successive periods of dark and light adaptation, by placing sensors around the eye [19]. Figure 2.9 shows the measurement method of EOG by placing electrodes around the eye.

In 1877, Dewar [20] used dilute sodium chloride solution held by a small trough of paraffin attached around the edge of the orbit of human eye to measure eye movements. Within this solution, the terminal of a nonpolarizable electrode was inserted and connected to a large gutta-percha trough containing salt solution into which the subject inserted his hand. By training the subject, it was possible to reduce large electrical variation caused by involuntary eye movements, and by fixing the eye concentratedly on a specific point while another observer checked galvanometer and altered the intensity of the light, electrical variation can be detected.

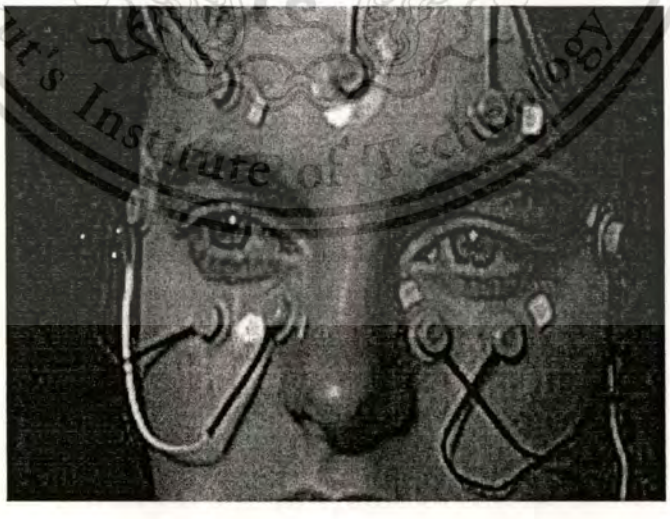


Figure 2.9 Eye movements measurement using EOG [14]

In 1922, Schott [21] used two electrodes of copper wire 0.4 to 0.5mm thick which were mounted nasally and temporally on the left side of a horn rim eyeglasses frame. The wires were arranged into the form of a spring. The nasally installed electrode ended in a spherical button 2mm thick of polished copper and laid on the caruncle of the cocainized eye. The temporally installed electrode end was bent back into a loop 2 to 3mm in diameter, which put on the bulbar conjunctiva in the outer canthus. These electrodes were wired directly to a string galvanometer of an electrocardiograph. Meyers [22] used an electrocardiograph to study the movements of nystagmus. He claimed that his technic required horseshoe-shaped tin electrodes on gauze saturated in saline solution and held on by tight bandages. Meyers regarded his records as originating from the action potentials of the extraocular muscles. Mowrer et al. [23] were able to obtain the results of Meyers, but their explanation was not one of action currents of eye muscles. Using a direct current amplifier and a ballistic galvanometer, they picked up potentials with silver-silver chloride electrodes in isotonic sodium chloride solution. They noticed that eye movements in the plane of the electrodes gave galvanoinetric changes, whereas movements at right angles to this plane gave none. Based on Mowrer et al., it is now accepted that the generated electrical potentials arise due to the permanent potential difference of between 10 to 30mV that exists between the cornea and the ocular fundus. This is commonly referred to as the cornea-retinal potential with the cornea being positive. An electrical field is set up in the tissues surrounding the eye and rotation of the eye causes a corresponding rotation of the field vector. For this reason, it is possible to detect eye movement with the appropriate placement of electrodes on the skin surrounding the eyes.

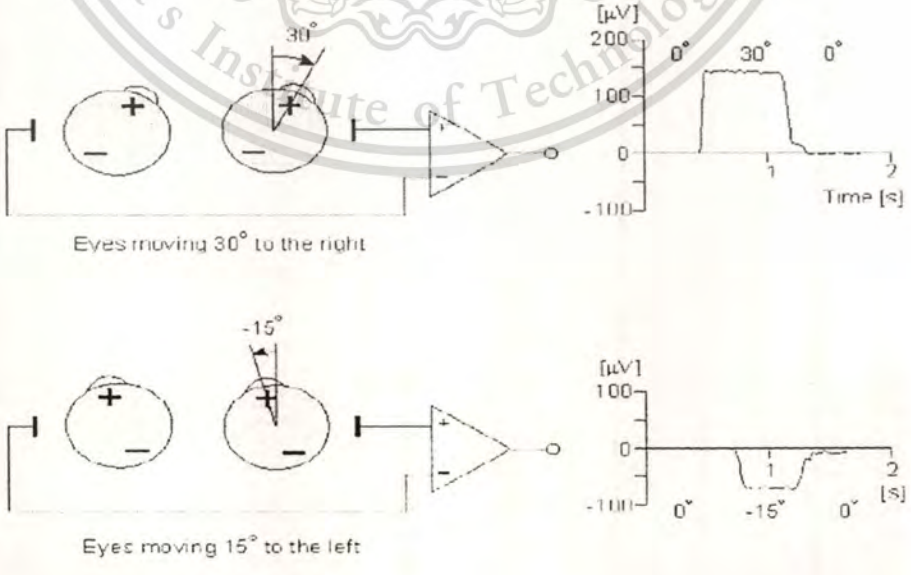


Figure 2.10 Measurement of horizontal eye movement using EOG.

Figure 2.10 illustrates the measurement of horizontal eye movements by the placement of a pair of electrodes at the outside of the left and right eye (outer canthi). With the eye at rest the electrodes are effectively at the same potential and no voltage is recorded. The rotation of the eye to the right results in a difference of potential, with the electrode in the direction of movement (i.e., the right canthus) becoming positive relative to the second electrode. (Ideally the difference in potential should be proportional to the sine of the angle). The opposite effect results from a rotation to the left, as illustrated. The calibration of the signal may be achieved by having the patient look consecutively at two different fixation points located a known angle apart and recording the appropriate EOG. Typical signal magnitudes range from 5-20μV/°.

The EOG is one of the very few methods for recording eye movements that does not require a direct attachment to the eye itself. For this reason, the EOG technique is preferred for recording eye movements in sleep and dream research and when recording eye movements in infants. However, the resulted signal is not stable and may vary under different light condition. Moreover, this method only measures 2D eye movements, which are horizontal and vertical movements. In case of measuring 3D eye movements, the researcher and medical experts commonly use Scleral Search Coil or VOG.

2.4.2 Scleral Search Coil

One of the most important eye movements measurement method was introduced by Robinson [24] who adapted the search coil magnometer to "scleral" eye movements tracking tool.

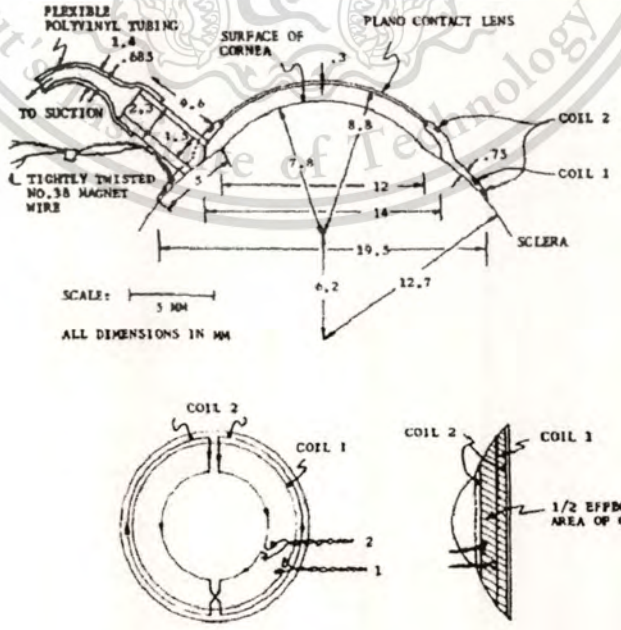
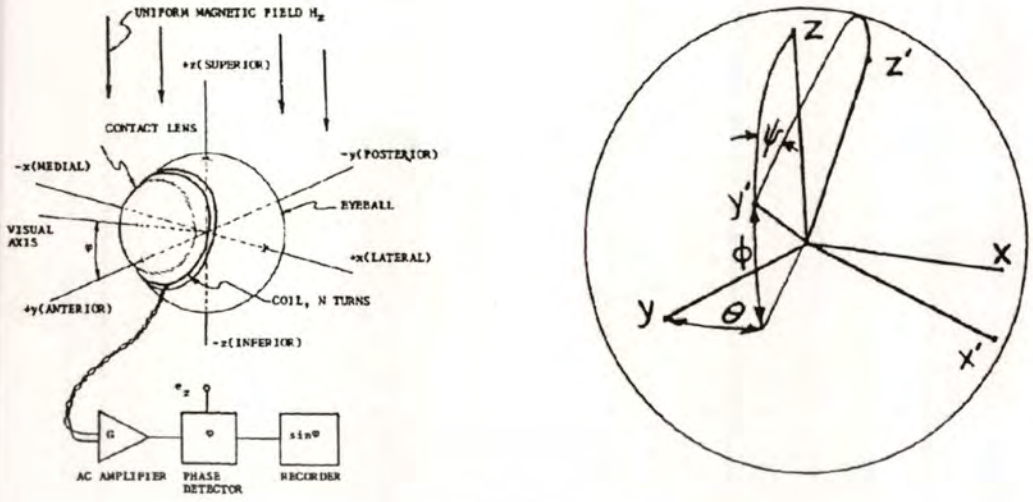


Figure 2.11 Original design of suction contact lens of Robinson



(A) (B)

Figure 2.12 (A) a simplified schematic of the use of a sclera field coil for obtaining the vertical (ϕ) component of eye position; (B) The coordinate system of the orbit and the globe defining the angles of the horizontal gaze (θ), vertical gaze (ϕ) and torsion (ψ)

Two induction coils were mounted on a scleral contact lens, as shown in Figure 2.11. The optical portion of the contact lens is an afocal or plano lens with a radius of curvature of 8.8mm and 0.3mm thick. The scleral portion is machined from methyl methacrylate plastic (Lucite) into which the lens is cemented. The coils are 10 turns of No. 44 magnet wire cemented in slots in the sclera portion. The air pocket over the cornea is 1mm thick and is vented by a laterally placed pipe leading to a very soft thin polyvinyl tubing. The scleral skirt is limited radially to 5mm to prevent interference with the fornices, the superior being the least accommodating. The scleral radius of curvature is already in general clinical use and in view of the corneal clearances allowed this lens should be applicable to the adult human eye without the need of special fitting. The lens is applied with the aid of a corneal anesthetic.

The scleral search coil was mounted on the eyeball, as depicted in Figure 2.12 (A). The coordinate system convention used in the measurement was depicted briefly in Figure 2.12 (B). To adhere the coil firmly on the eyeball, a vacuum was made by suction through polyvinyl tube. This suction method tended to be a source of annoyance for some users. The eyeball was surrounded by alternating-current magnetic field generated from external field coils. In the primary position of the gaze, the eye looked straight ahead along the *y* axis, the coil lies in the *x*-*z* plane and no voltage was induced. If the eye looked up by the angle ϕ, the voltage induced was given by the following equation:

$$e_1 = +NA \sin\phi H_z \omega \sin(\omega t) \times 10^{-8} \quad (2.1)$$

where A is the area of coil (2.55cm^2), ω is the radian frequency (2π X 5000) and A sin ϕ is the area of the coil projected onto the *x-y* plane which links the flux. H_z is the vertical alternating magnetic field (2.19 gauss peak). *N* is the coil magnet wire. By suitable amplification and using phase-sensitive detection, the coil voltage was changed into analog voltage, proportional to the sine of the angular eye position. One coil on the lens and one magnetic field would permit the measurement of only one degree of freedom. By generating a horizontal and vertical magnetic field in phase-quadrature of at different frequencies, simultaneous measurements of horizontal (θ) and vertical eye movements (ϕ) were accomplished. Finally, a specially configured second coil that installed on the lens in the *y-z* plane allowed the system to record the torsional movement of the eye.

In 1975, Collewijn et al. [25] introduced a new carrier of the search coil that enables more general use. The carrier was a flexible ring, which fits on the limbic area. concentric with the cornea. The ring was cast out of silicone rubber in a special mold. At the ocular side, it was concave with a radius of curvature (6.0mm) smaller than that of the globe. The induction coil (9 windings of magnet wire of 0.05mm diameter) was embedded within the material of the ring. The twisted ends of the wire leaved the ring and were connected to the detection device. For application of the ring, the surface of the eye was anaesthetized with one or two drops of an anaesthetic. The ring was wetted with Ringer's solution, put on the limbus with the leads at the side of the outer canthus and pressed firmly upon the eye. Once placed, the ring adheres to the eye with surprising force and can only be removed by lifting its edge, e.g. with a blunt forceps. The mounting procedure is facilitated by handling the ring with a special suction device in the shape of a hollow ring.

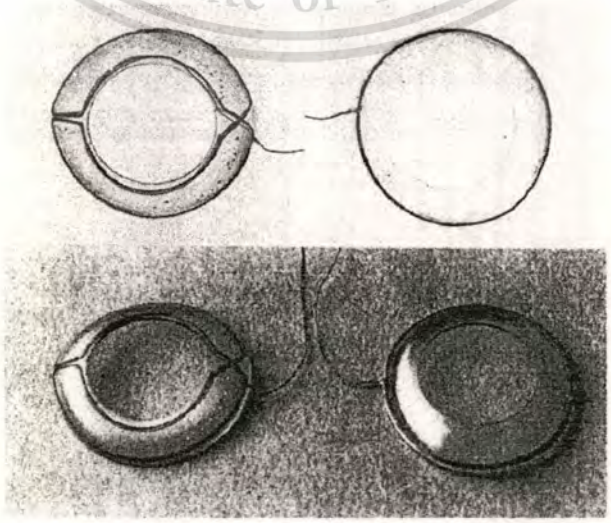


Figure 2.13 Cyclorotational coil (left) and standard coil (right) attached on silicon rubber

Figure 2.13 shows the proposed new carrier with the coil inside [26]. There were two types of coil, standard coil (right pictures) and cyclorotational coil (left pictures). The frontal view of the cyclorotational and standard coil in upright position was depicted in Figure 2.13 (top), as they would be seen when facing the a subject with the coils mounted. The oblique frontal view was shown in Figure 2.13 (bottom). The standard coil was used to measure the horizontal and vertical movements, while the cyclorotational coil was used to measure the torsional movement. To measure the 3D eye movements simultaneously, the coils can be worn separately on left eye and right eye. These changes were considerable improvement and the technique has become the standard for eye measurements in the laboratory, and increasingly in the clinic.

Currently, scleral search coil is still considered as the most precise eye movements measurement tool since the measuring results are directly yielded from the eye movements. This method is also considered as a flexible method, insensitive to illumination change and head movements. In the other side, some researcher found that this method also has several disadvantages. Technically, this method suffers from coil slipping and possibility of electrical leads breakage [24, 25]. The coil slipping and electrical leads breakage occur due to eye blink and force exerted by the eye lid. 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 [27, 28]. Generally, lid irritation and corneal staining tended to be associated with improper procedure of coil installation. Figure 2.14 shows the topography image of corneal staining during coil wear on a subject. The picture was taken 24 minutes after insertion of the coil.

Figure 2.14 Topography image showing corneal staining

2.4.3 Video-Oculography

The drawbacks of scleral search coil method mentioned in previous sub section and advances in digital image processing technology over the past decade, have encouraged the researcher to develop video-based systems (VOG) for measurement of three-dimensional eye movements. The small number of existing image processing systems for eye movements measurement fall into two broad categories: those based iris striations tracking using polar cross correlation method [29-32] and those employing natural or artificial landmark tracking [4, 5, 7, 33-36].

In 1983, Hatamian and Anderson [29] presented a technique to measure horizontal, vertical, and torsional eye movements based on video image analysis. The system consisted of infrared sensitive video camera, Z80-based camera control unit, TV monitor, memory, and LS-11/23 processor. Horizontal and vertical eye movements were measured by tracking the center of pupil. The algorithm that was being used for detecting the center of pupil was based on boundary scanning and centroid computation. Firstly, a threshold operation at a level was chosen to highlight the black pupil against the background. Then, boundary scanning algorithm was implemented. When a black point was passed, its coordinates were saved and if for a certain number of following pixels were continuously sensed as a black region, then it was most probably the pupil area. The saved point was taken as a point on the boundary of the pupil. Another boundary point was found in a similar way for the far edge of the pupil. The x and y coordinates of these edge points for all horizontal lines were averaged, yielding the coordinates for the pupil center.

The center of pupil was then used as a sampling center for torsional eye movement measurement. Promising features that could be used for detection of torsional eye movements were iris landmarks. These landmarks were stable patterns, and therefore, under rotation they would contain information about the angular position of the eye with respect to a reference position.

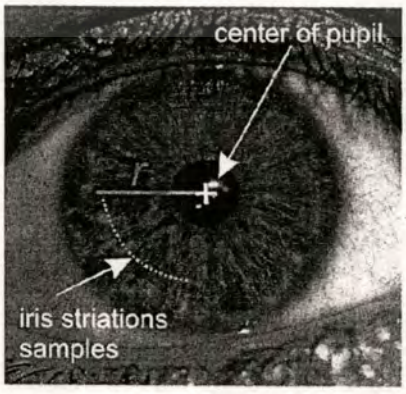


Figure 2.15 Photograph of human eye showing iris patterns

Figure 2.15 shows a photograph of a human iris and a pupil. Let $I(r, \vartheta)$ represented this image in polar coordinates where r was the constant radius from center of pupil and ϑ was the angle of rotation. Suppose that $I_r(r, \vartheta)$ was the rotated version of this image, and it was assumed that the center of rotation was coincident with the origin of our polar coordinate system. To detect the angle of rotation between the two images, an angular cross correlation was performed as

$$c_{ii_r}(\tau) = \int_{r=0}^{R} \int_{\vartheta=0}^{2\pi} I(r,\vartheta) I_r(r,\vartheta+\tau)\, dr\, d\vartheta \qquad 0 \le \tau < 2\pi \tag{2.2}$$

where R was the radius of the largest disk that fits the frame of the image. The value of τ for which $c_{ii_r}(\tau)$ attained its maximum was the estimate of measured torsional shift. In practical, the implementation of (2.2) could be done by evaluating annular region ($R_1 < r < R_2$) or even a single ring ($r = R_0$). To improve the measurement resolution, a fast second degree local least square interpolation was used by computing maximum value of fitted polynomial from several sample points. The system was then validated by comparing rotated human eye picture about the center of pupil and the computed value from video image analysis. This system, although it is considered as the first generation of VOG which uses polar cross correlation to measure torsional movement, can only be used in research laboratory due to complicated measurement procedure.

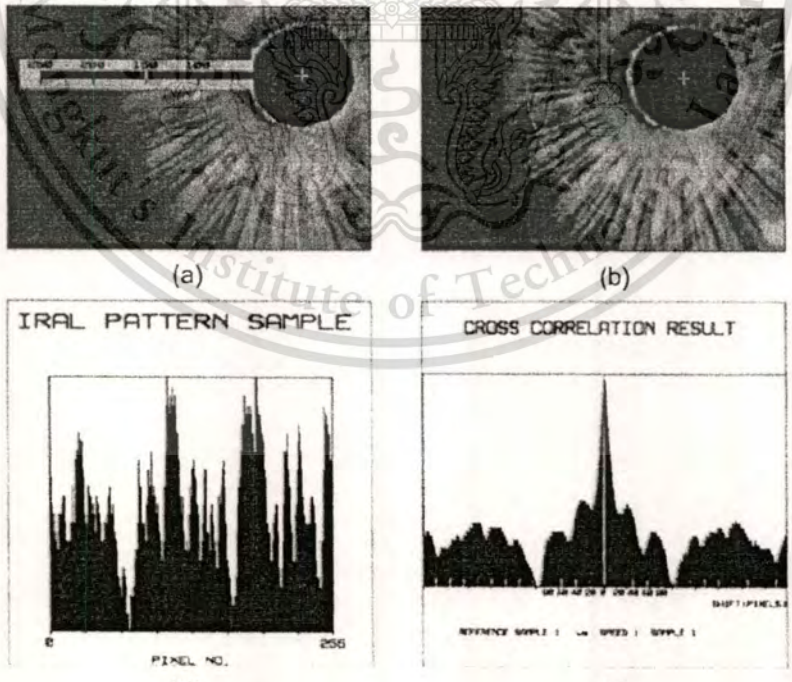


Figure 2.16 The procedure of iris striation tracking using VTM

In 1991, Moore et al. [31] proposed a real-time and interactive eye movements video system which was called Video Torsion Measurement (VTM). Horizontal and vertical eye movements were computed by center of mass algorithm. The method for measuring torsional eye movement was to cross correlate the two grey-level distributions of an arc of the iris from two separate images using an image processor card interfaced to an IBM-AT compatible computer. The card (Matrox MVP-AT) was supplied with a library of low-level functions for controlling the hardware operations of the board and the VTM system software, which was written in the C programming language, incorporating these low-level functions to interface with the MVP-AT board as well as carrying out the data-acquisition and processing algorithms. The reported frame acquisition time was 40 ms/frame.

The procedure of iris striation tracking using VTM is shown in Figure 2.16. This program acquired an image of an iris illuminated by a single infrared (IR) light source; threshold this image in order to identify the pupil and calculated the pupil area and located the centre of the pupil using a centre-of-gravity algorithm (Figure 2.16 (a)); recorded the grey-level distribution along an arc 256 pixels long at a selected radius from the pupil center (Figure 2.16 (b) and (c)); carried out a Fast Fourier Transform (FFT) on this interpolated grey level distribution; stored the parameters of this reference FFT and cross correlated the comparable iral grey-level distribution from other test images of the same eye in order to determine the amount of torsional rotation of the test images relative to the reference image (Figure 2.16 (d)). This first generation of VTM system was considered suitable for real clinical operation compared to the previous system by Hatamian and Anderson [29] although it could only measure the torsional movement within a small range of horizontal and vertical eye positions (approx. $\pm 5^{O}$) [37].

Simpler algorithm was presented by Yamanobe et al.[34] by tracking natural landmark of the human eye or Kim [4] by tracking artificial landmark created from direct painting on the eye.

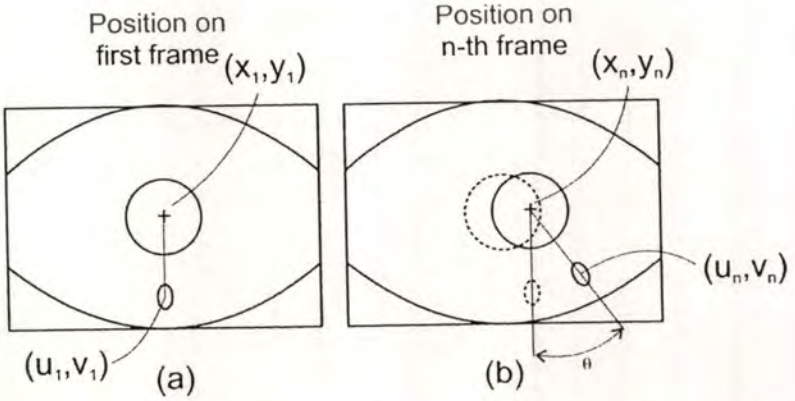


Figure 2.17 Proposed torsional movement measurement by Yamanobe et al.

Yamanobe et al. used similar algorithm as Moore et al. [31] to detect the center of pupil whereas one iris striation was used to measure the torsional movement of the eye. The landmark was manually defined by operator and saved in the memory. A cross correlation was then implemented to track the landmark in successive images. The torsion angle was calculated to observe the relation between the center of pupil and the striation with each image, as shown in Figure 2.17. The torsion angle was computed by observing the following equation:

$$\theta_n = \tan^{-1}\frac{u_n - x_n}{v_n - y_n} - \tan^{-1}\frac{u_1 - x_1}{v_1 - y_1} \tag{2.3}$$

where θ_n was the measured angle, (u_1, v_1) and (x_1, y_1) were the iris striation and center of pupil coordinates on the reference frame, respectively. (u_n, v_n) and (x_n, y_n) were the iris striation and center of pupil coordinates on the n-th frame.

Kim [4] proposed a simple method to detect horizontal, vertical, and torsional eye movements using two artificial landmarks placed on the eye. The landmarks were created by painting the eye using sterilized fine paintbrush, as shown in Figure 2.18 (A). In this research, the experiment was implemented to guinea pig. For human subject, the usage of contact lens as proposed by Ott [35] can be considered as well, as shown in Figure 2.18 (B). The landmarks were than detected by center of mass (COM) algorithm or the other tracking methods. By representing the landmarks coordinates in spherical coordinate system and computing the angle by general trigonometric functions, the three-dimensional eye movements measurement can be performed. Although this method was considered more simple and inexpensive, the usage of contact lens on human subject has similar drawbacks as scleral search coil method. Special treatment under medical expert should be considered if this method is used in clinical application. Therefore, most commercial diagnostic eye movements tracking applications used in clinical routine tend to use iris striation or natural landmark tracking.

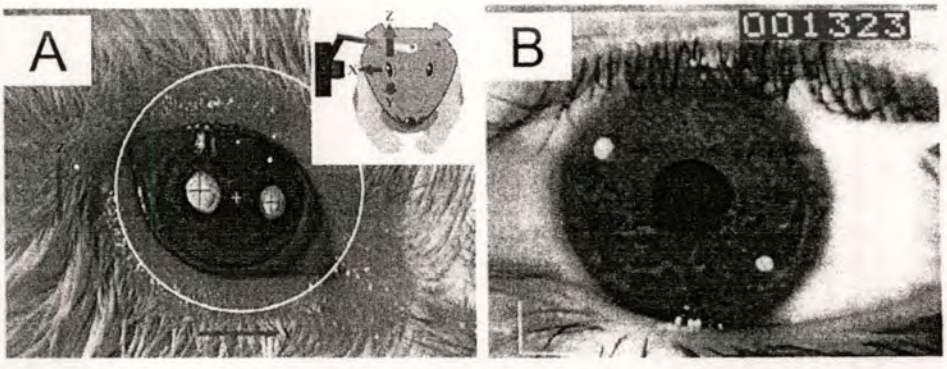


Figure 2.18 Eye movements measurement using two artificial landmarks

Both current eye movements tracking method using iris striation or natural landmarks yield inaccurate measurement results when these methods have to deal with geometric distortion due to 3D point projection from eyeball to image plane of camera sensor and lens aberration. These distortions result in erroneous measurement of ocular torsion given by the angle formed between a line (bisecting two arbitrary points projected from the eye onto a two-dimensional video image) and the reference horizontal or vertical line.

Moore et al.[32, 37, 38] proposed a new generation of VTM system by including geometric-compensation algorithm to obtain accurate measurement of torsion angle. Moore et al. found that when the eye was looking directly into the camera, the sampling window formed in the image plane was a segment of a circle, and the projection of this arc onto the spherical surface of the eye formed a similarly shaped arc. As the eye position changed, the shape and location of this sampling window would vary depending on the orientation of the eye. To compensate this misalignment, a mathematical formula used in 3D point projection from eyeball to image plane of camera sensor was improved by considering the rotation matrices of pupil and sampling windows from reference position to new position. To measure user-dependent parameters, such as the radius of eyeball, the user had to undergo a complicated calibration procedure by fixating the gaze to several points and approximating the eyeball radius by adopting formula of Fick system [8, 37]. This improved VTM system was able to measure eye position over a larger range of horizontal and vertical eye positions (approximately ± 20^{o}). Figure 2.19 shows the projection of iris striations into the image plane. Light dashed lines indicate the original sampling windows when the eye in reference position. Heavy dashed lines indicate the uncompensated sampling windows, the solid lines show location and shape of the properly compensated arcs.

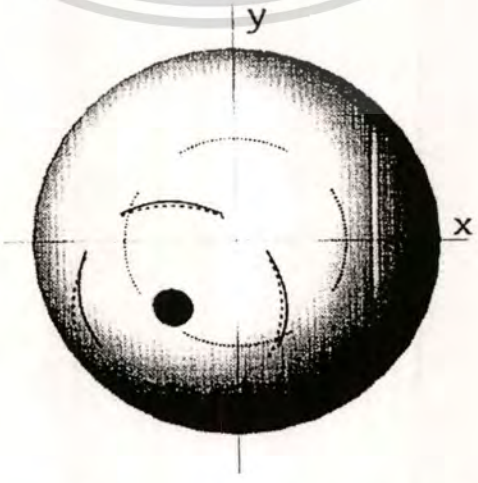


Figure 2.19 Diagram of projection of iris striations into the image plane

An eye tracking system using single camera and fluorescent marker array in [5] proposed a camera alignment method to reduce geometric distortion. The proposed method was based on the use of four 3mm diameter white LEDs facing the eye positioned around each camera. The camera's optic axis was considered directed through the surface's center of eyeball curvature when the LED reflections were equidistant from the center pixel of the image. If the camera was tilted with respect to the eyeball surface, the LED reflections were no longer equidistant from each other, and the center of the cross they define was no longer aligned with the center pixel of the captured image. Although the camera alignment method yielded precise measurement for camera alignment, it required distinctive camera setup for each patient which could be impractical and time consuming. Furthermore, all reported eye movements tracking methods do not provide real-time 3D visualization information based on measurement results which is very important to diagnose and understand how the voluntary and spontaneous eye movements work. In the other side, huge advantages in the usage of 3D visualization language as a tool to create a 3D model have increased the interest of simulating and modeling human vision system for clinical and educational purposes [9-11].

In this study, we have developed a new VOG system for 3D eye movements tracking and visualization using dual cameras acquisition. The system is noninvasive, real-time, inexpensive, and based on natural landmark tracking. Two parallel mini CCD cameras were used to capture two simultaneous images of one eye. Center of mass (COM) and template matching algorithms were utilized to obtain 2D positions of the center of pupil and iris striation. The geometric distortion and lens aberration were solved using 3D coordinates extraction based on Direct Linear Transformation (DLT) and lens undistortion algorithms that have never been reported before in the literature. These geometric distortion removal algorithms required only one preliminary calibration procedure and were considered more efficient than others described in the previous methods [5, 37]. The camera parameters and distortion coefficients were obtained through camera calibration. The DLT and lens undistortion algorithms were then used to generate accurate 3D coordinates of pupil and iris landmark. The 3D eye rotations were computed from 3D coordinates. The tracking results were then presented as eye movements trajectories and real-time 3D visualization of eyeball. The system was validated by standard 5 x 5 calibration angles and artificial eyeball as used in [5]. Our system has high accuracy due to small values of mean of error: 0.15^{o}, 0.14^{o}, and 0.20^{o} for horizontal, vertical, and torsional movements, respectively.

Chapter 3

Methodology

As mentioned in chapter 2, several techniques have been examined to measure 3D eye movements. A good technique should be noninvasive, real-time, inexpensive, and based on natural landmark tracking. In this thesis, a new Video-Oculography (VOG) system for 3D eye movements tracking and visualization is proposed.

In order to measure 3D eye movements, a dual cameras system mounted on a consumer-grade welding goggle had been designed. The hardware configuration is described in section 3.1. The algorithms of eye movements tracking software to compute 3D eye movements and display the measurement result in real-time 3D visualization will be explained in section 3.2. Section 3.3 explains experiment procedure for system validation and real-time implementation.

3.1 Hardware Configuration

Figure 3.1 shows several core components for dual cameras system. Two special mini color CCD cameras (Shenzhen Xinzhilin Electronic Co., Ltd, Shenzhen, China) for video surveillance (Figure 3.1 (a)) were adopted as main capturing devices. The camera size is 12mm x 12mm. Each camera was equipped by a 1/3" Sony Color CCD sensor. The focal length of the pinhole lens was 3.7mm. The maximum resolutions of the CCD camera were 510 x 492 in NTSC video format or 582 x 500 in PAL video format.

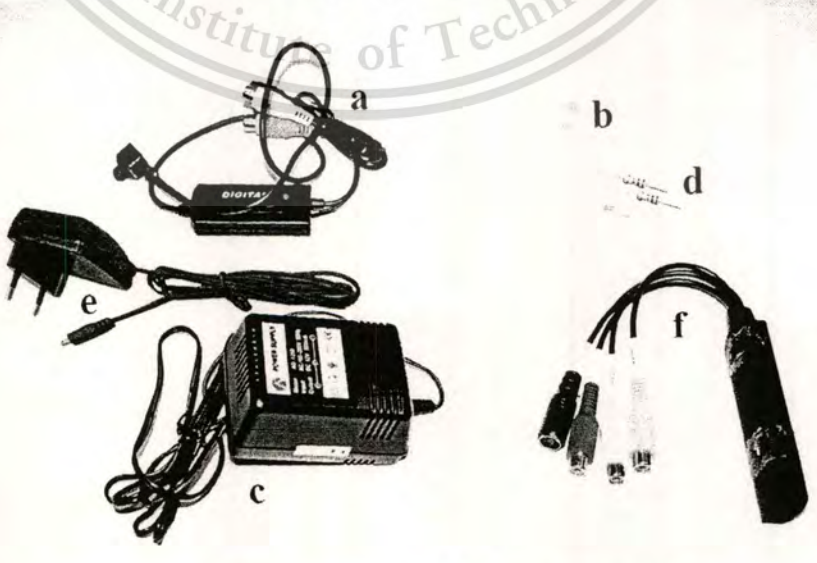


Figure 3.1 Core components for dual cameras system

Table 3.1 Specifications of the mini CCD camera

Model	XL-3010CP4
Image Sensor	1/3" Sony Color CCD sensor
Max. Resolution	NTSC: 510 x 492 ; PAL: 582 x 500
Dimension	12mm x 12mm
Lens	Pinhole camera (fixed), Focal length: 3.7mm
Power Consumption	DC 12 V, 100mA
White Balance	Auto
Minimum illumination	0.5Lux
Cover color	Black

The camera was then manually modified into infrared camera in order to be able capturing clear iris striation by removing the infrared filter (Figure 3.1 (b)) installed inside the camera. Each camera was powered by an AC/DC adaptor (Figure 3.1 (c)) with AC 180-260V, 50Hz input and DC 12V, 100mA output. The complete specifications of the camera are explained in Table 3.1.

Four 5 mm infrared LEDs (940nm) (Figure 3.1 (d)) were added to create illumination and reveal the iris texture of the eye. Each infrared LED was coupled in serial circuit with 500Ω ± 5% resistor (Figure 3.1 (d)) to avoid short circuit. The LED ran at ± 1V, supplied by special AC/DC adapter (Figure 3.1 (e)) with AC 100-240V, 50Hz input and DC 12V, 100mA output. The power supply separation of mini CCD camera and LED was conducted to avoid noise that influencing the quality of images. EasyCap® Video Adapter (Shenzhen Forwardvideo Industrial Co., Ltd., Shenzhen, China) (Figure 3.1(f)) was used on each camera to convert analog image into digital image. The maximum sampling rate of EasyCap® was 30frames/second. The cameras and infrared LEDs were then installed into 40mm x 55mm special plastic box as shown in Figure 3.2. The horizontal and vertical distances between LEDs were approximately 26mm and 35mm, respectively. This positioning was critical. Simply placing the infrared LED directly in front of the eye produced a featureless, low-contrast image of the iris which was unsuitable for measurement. In order to produce clear, high-contrast images of iris striation it was necessary to place the light source so it had a grazing angle of incidence [31]. It was found through experiment that placing four LEDs in this rectangular position had given enough illumination to reveal iris pattern of the subject.

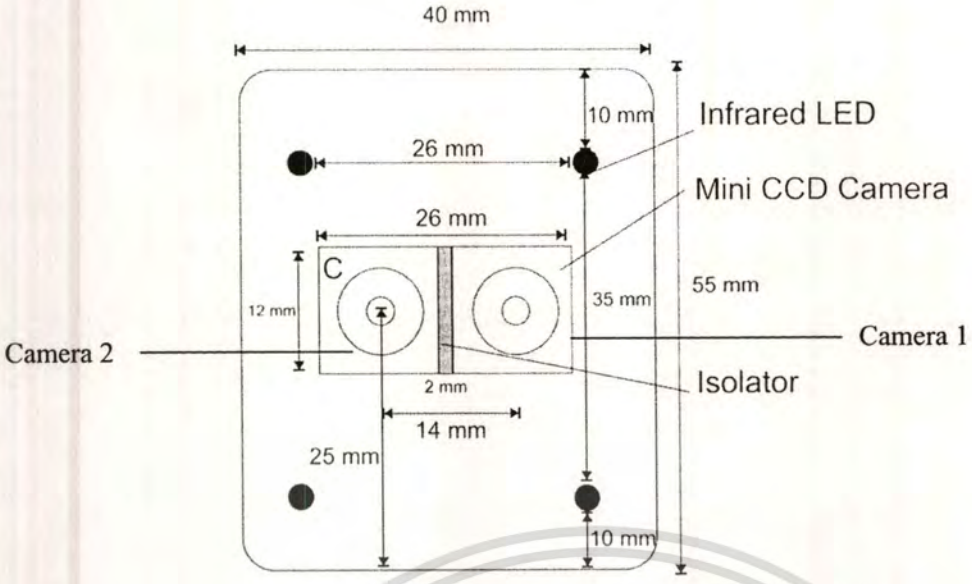


Figure 3.2 Diagram of dual cameras system installation

One of the major problems in multiple vision systems is the well known "matching" problem, which is the determination of conjugate points between stereo pair images. In order to achieve rigorously matching result on stereo pair images, the two cameras were installed parallel in horizontal direction as illustrated by Figure 3.2. The distance between two center of lens was 14mm. The right camera and left camera were defined as camera 1 and camera 2, respectively. A 2mm plastic isolator was placed between the cameras to prevent electric noise. A consumer-grade welding goggle (Senko Co., Ltd, Gangwon-do, Korea) was modified to construct a low cost eye movements tracking goggle as shown in Figure 3.3. The dual cameras system was attached on the front side of the goggle. The box of dual cameras system could be attached on the left or right side of the goggle by utilizing special two sides straps.

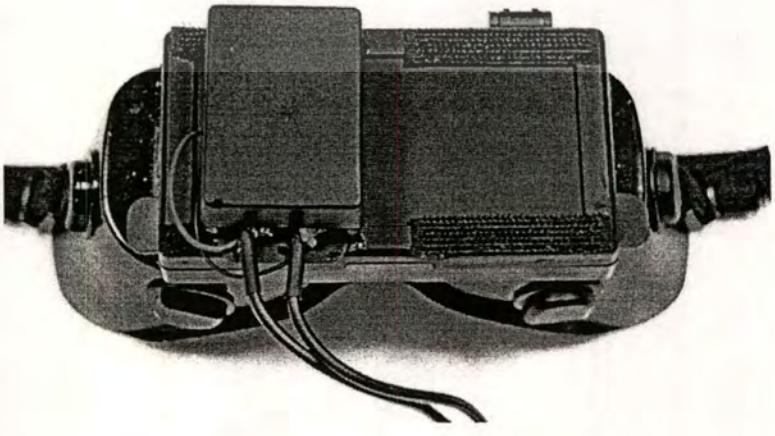


Figure 3.3 The eye movements tracking goggle

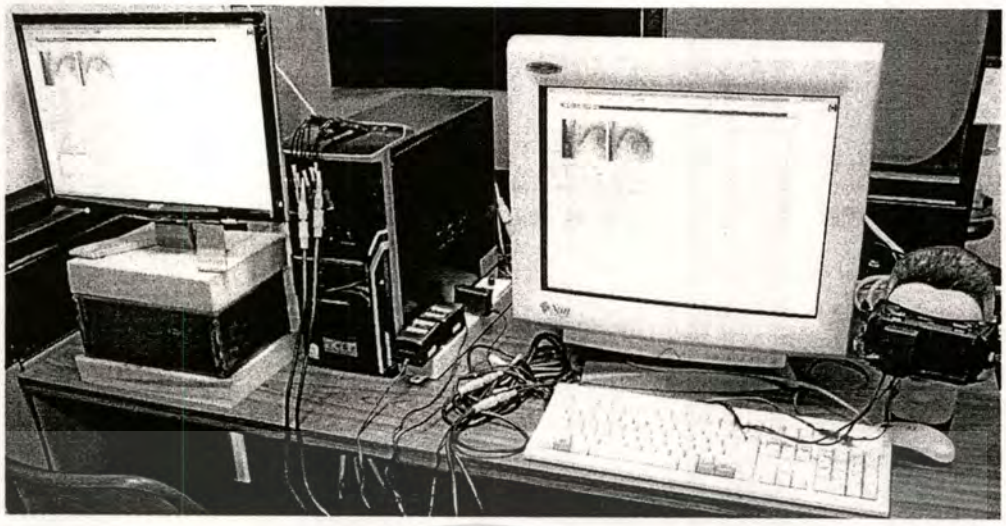


Figure 3.4 Real-time 3D eye movements tracking and visualization system

Toshiaki et al. [39] reported that in several eye movements measurement process, such as measurement of *caloric nystagmus*, the patient wore a light-occlusion goggle to prevent light stimulation that could affect the accuracy of measurement. On the other hand, there was a need to measure voluntary eye movements, such as measurement of saccade or smooth pursuit, by giving visual stimulation to the patient [40]. Hence, we considered to design a goggle that support both light-occlusion and open-field configurations by attaching a removable front side cover.

Figure 3.4 shows the real-time 3D eye movements tracking and visualization system. The system consists of three parts: (i) the goggle, (ii) personal computer, and (iii) the 3D eye movements tracking and visualization software. The goggle was used to capture real-time images sequence of the eye with a resolution of 320 x 240 in PAL video system. This image resolution was considered optimum for analysis and real-time processing using our hardware and software capability. The images were then passed to EasyCap® Video Adapter through video composite port. EasyCap® Video Adapter converted the analog images into digital images and sent to the personal computer through USB port. The specifications of personal computer used in this system were Intel® 2 Quad CPU Q6600 @ 2.4GHz, RAM of 2GB, and Windows XP® Service Pack 3 operating system. A 25" CRT SUN microsystem® and a 21" wide screen LCD ACER® monitors were used as the main monitor for the operator and the second monitor for the patient, respectively. The second monitor was used to display visual stimulus in experiment routine. The captured images were then converted to gray-scale images and processed in real-time by the 3D eye movements tracking and visualization software as shown in Figure 3.5. The software was created in C++ language using Borland C Builder 6.0, OpenCV Library [41], and OpenGL Library [42].

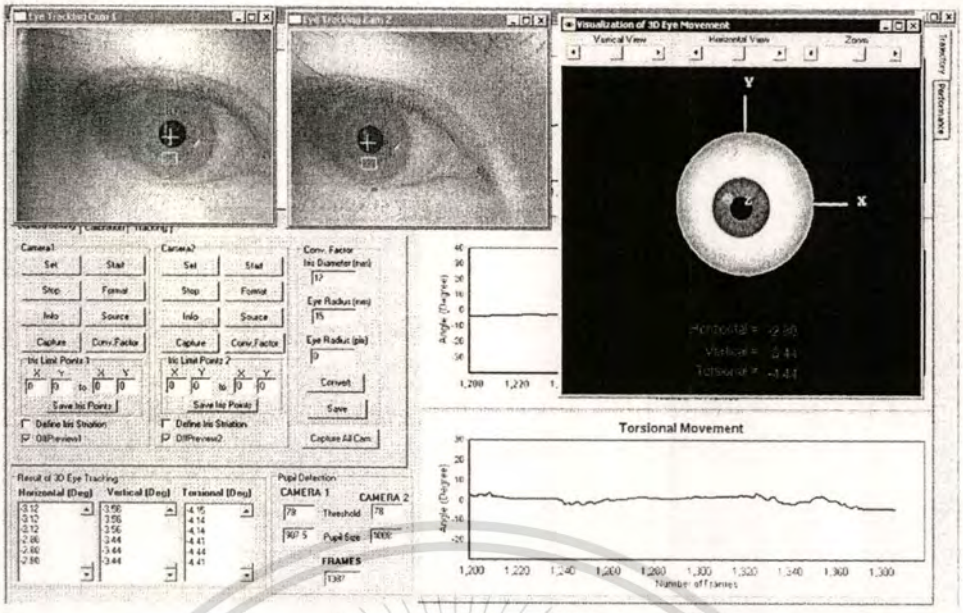


Figure 3.5 The user interface of 3D eye movements tracking and visualization software

Borland C Builder 6.0 was chosen to create menu-driven software that could be used easily by operator. OpenCV is a free and open-source computer vision library originally launched in 1999 by Intel, Inc. This library provides real-time support for video capture and image processing algorithms. OpenGL is a standard specification defining a cross-language, cross-platform API for writing applications that produce 2D and 3D computer graphics. OpenGL was developed by Silicon Graphics Inc. in 1992 and is widely used in virtual reality and visualization.

Table 3.2 shows the total cost spent to build an inexpensive eye movements tracking and visualization system. The system was accomplished at cost around US$789.

Table 3.2 Price list of eye movements tracking and visualization system's components

Components	Price (US$)
Two mini CCD cameras and three power supplies	136
Welding Goggle	6
Four IR LEDs, four resistors, small electronic cables	2
Plastic box and straps	4
Two EasyCap® Video Adapter	34
Two Extension Cable for Video Adapter	7
Standard Personal Computer + Extra Monitor [43]	600
TOTAL	**789**

3.2 Software Algorithm

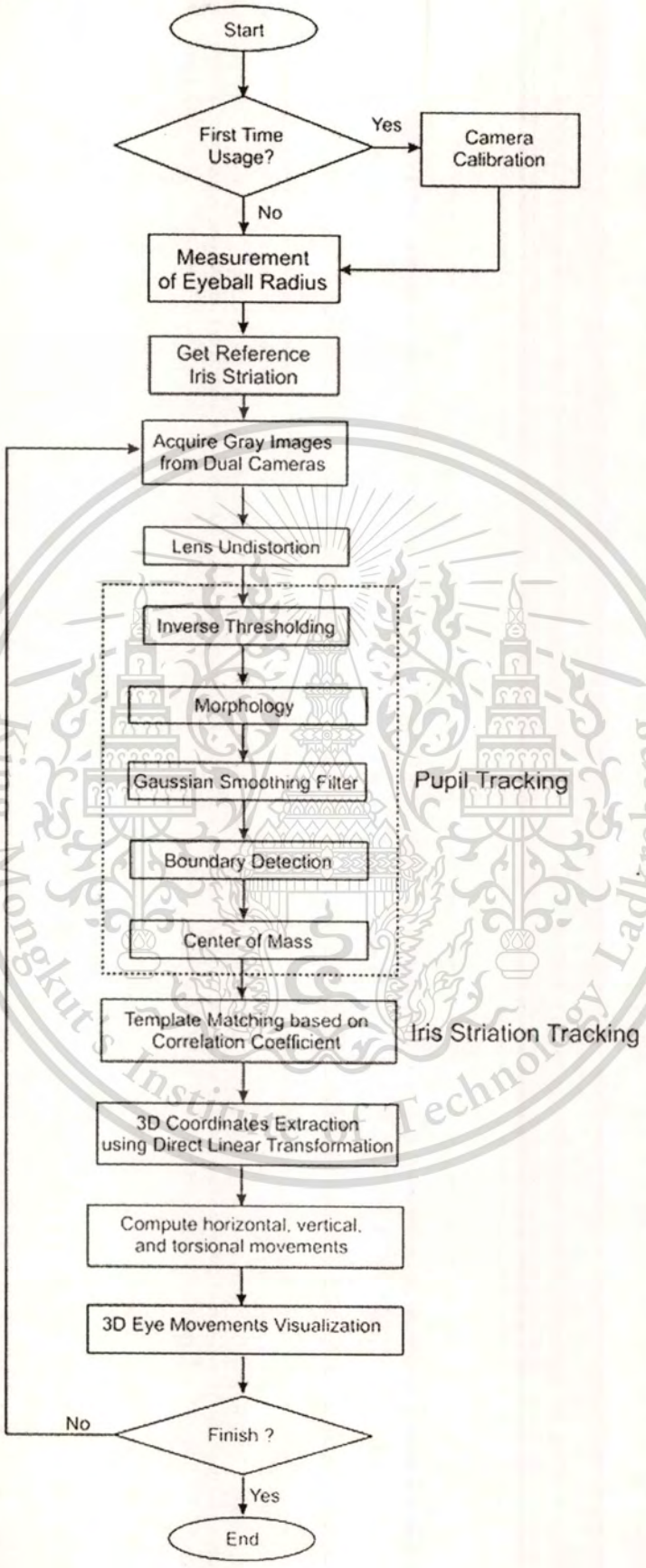


Figure 3.6 Flowchart of 3D eye movements tracking and visualization

Figure 3.6 shows the flowchart of 3D eye movements tracking and visualization software algorithm. The proposed algorithm can be explained as follows:

1. Initially, we calibrated the dual cameras system by capturing several chessboard pattern images to extract camera parameters and lens distortion coefficients.

2. Eyeball radius of the patient was determined by measuring the distance of outer part (*lateral canthi*) and inner part (*medial canthi*) of eyelid aperture (*palpebral aperture*) using a professional caliper. The iris diameter is also obtained to be used in mm to pixels conversion process for angular measurement of 3D eye movements.

3. Iris pattern used in iris striation tracking was obtained by selecting particular area of the iris. The selection was performed by capturing the eye image under infrared illumination and manually cropping an area of iris as a tracking reference.

4. After initializing various variables, color images were acquired using the dual cameras system. We then converted color images to gray-level images in order to significantly decrease computational cost, while all the details were still preserved.

5. Lens undistortion was utilized to eliminate distortion from lens aberration. The distortion parameters obtained in previous camera calibration step is used in this process. This process was necessarily included in the algorithm to obtain accurate 3D coordinates extraction result.

6. Pupil tracking was conducted by first converting a gray image into a binary image using inverse thresholding. Black area of pupil were recognized as an important part of this step. The Morphology and Gaussian smoothing filter were used to remove noise. Boundary detection using border following algorithm is used to detect the edge of the pupil. Finally, Center of Mass algorithm was applied to detect 2D coordinates (x, y) of the center of the pupil.

7. Iris striation tracking was conducted by implementing template matching algorithm. Tracking reference was taken from the iris striation selection step. In order to increase tracking accuracy, a region of interest (ROI) based on pupil center was selected. The tracking algorithm was then applied inside the defined ROI, resulting 2D coordinates (x, y) of iris striation center.

8. 3D coordinates extraction using Direct Linear Transformation (DLT) was conducted to obtain 3D coordinates (x, y, z) of pupil and iris reference centers from two captured images taken from dual cameras system. The computed 3D coordinates were used in the next step.

9. By assuming the eyeball as a perfect sphere, the angular measurement of 3D eye movements could be obtained using general trigonometric functions. Eyeball radius was measured manually. 3D coordinates of pupil center and iris reference center were calculated using DLT algorithm. Horizontal, vertical, and torsional eye movements were computed by inserting eyeball radius, 3D coordinates of pupil and iris reference centers into trigonometric equations for spherical coordinate system.

10. The tracking results were presented in real-time trajectories and 3D visualization. First, the 3D eyeball model was created and initialized. Next, real-time eye movements data was applied to the model. OpenGL simultaneously rendered the new position of 3D eyeball model based on supplied real-time data.

3.2.1 Digital Image Representation and Coordinates Conventions

Gonzalez et al. [44] explained that an image may be defined as two-dimensional function, $f(x, y)$, where x and y are spatial (plane) coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the intensity of the image at that point. The term gray-level is used often to refer the intensity of monochrome images. Color images are formed by a combination of individual 2D images. For example, in the color system, a color image consists of three (red, green, and blue) individual component images. For this reason, many of the techniques developed for monochrome images can be extended to color images by processing the component images individually.

An image may be continuous with respect to x and y coordinates, and also in amplitude. Converting such an image to digital form requires that the coordinates, as well as the amplitude, be digitized. Digitizing the coordinate values is called *sampling*, digitizing the amplitude values is called *quantization*. Thus, when x, y, and the amplitude values of f are all finite, discrete quantities, we call the image a *digital image*.

The result of sampling and quantization is a matrix of real numbers. Assume that an image $f(x, y)$ is sampled so that the resulting image has N rows and M columns, then the image size is M x N. The values of the coordinates (x, y) are discrete quantities. For notational clarity and convenience, we use integer values of these discrete coordinates.

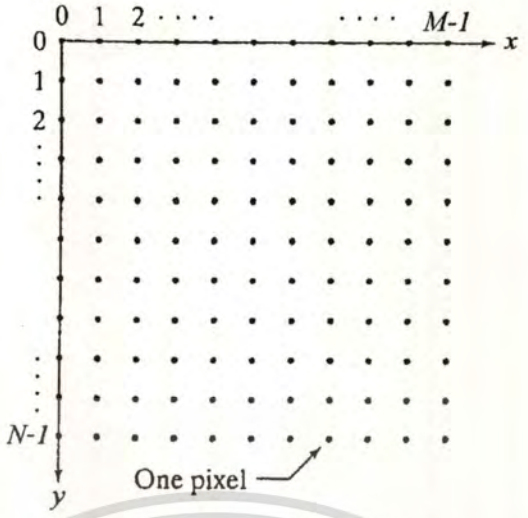


Figure 3.7 Image coordinate conventions of digital images

Figure 3.7 illustrates the coordinate conventions of digital image used in this thesis. The origin (0, 0) of coordinate system is defined to be at the top-most left pixel. The next coordinate values along the first row of the image are $(x, y) = (1, 0)$. Note that x range from 0 to $M - 1$, and y from 0 to $N - 1$, in integer increments. The unit used in digital image coordinates conventions and real physical measurement are *pixels* and *millimetres* (mm), respectively.

3.2.2 Pupil Tracking

3.2.2.1 Inverse Thresholding

Each gray-level image shown in Figure 3.8 (a) and Figure 3.8 (b) had different intensity level due to different view angle of the cameras. In order both images to have equal pixel intensity values, intensity adjustment would be the first thing to perform. The intensity adjustment was performed by mapping the intensity values in each image to new values. *Inverse thresholding* is such an intensity adjustment method that transforms an input image f to an output (segmented) binary image g as follows:

$$\begin{aligned} g(i,j) &= 1 \quad for \quad f(i,j) \le T \\ &= 0 \quad for \quad f(i,j) > T \end{aligned} \tag{3.1}$$

where T is threshold value, $g(i, j) = 1$ (white) for image element of object and $g(i, j) = 0$ (black) for image element of the background. In our case, the pupil image, the darkest area on eye image, was mapped to 1 and the others were mapped to 0. The threshold value was determined by adjustment, which was $T = 71$ in 0 to 255 gray-level range. The thresholding results were binary images with segmented white colour pupil on a black background as shown in Figure 3.8 (c) and Figure 3.8 (d).

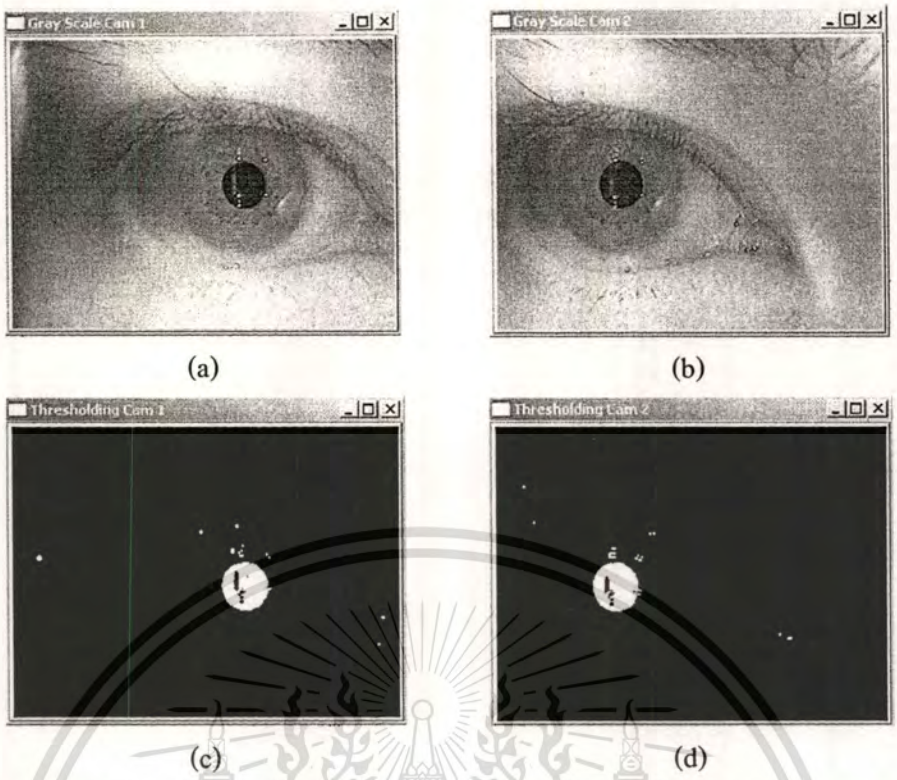


Figure 3.8 Gray-level eye images acquired from (a) camera 1; (b) camera 2 and inverse thresholding result on images acquired from (c) camera 1; (d) camera 2

However, small noise still remained in the binary images resulted by thresholding. Hence, Morphology and Gaussian smoothing filter were applied to remove the noise.

3.2.2.2 Morphology

As explained in [45], morphological operations are used predominantly for noise filtering, image simplification, skeletonizing, object marking, and many more. Mathematical morphology involves some basic concepts of set theory, such as inclusion ($\subset$ or $\supset$), intersection ($\cap$), union ($\cup$), the empty set ($\varnothing$), and set complement (^{C}). The initial assumption states that real images can be modeled using *point sets* of any dimension (e.g., *N*-dimensional Euclidean space); the Euclidean 2D space $\mathcal{E}^2$ and its system of subsets is a natural domain for planar shape description. A binary image can be treated as a 2D point set. Points belonging to objects in the image represent a set *X*. These points are pixels with value equal to one. Points of the complement set X^c correspond to the background with pixel values equal to zero. The origin, marked as a diagonal cross, has coordinates (0, 0), and coordinates of any point are interpreted as (x, y) in the common way used in mathematics. Figure 3.9 (a) shows an example of such a set points belonging to the object are denoted by small black squares.

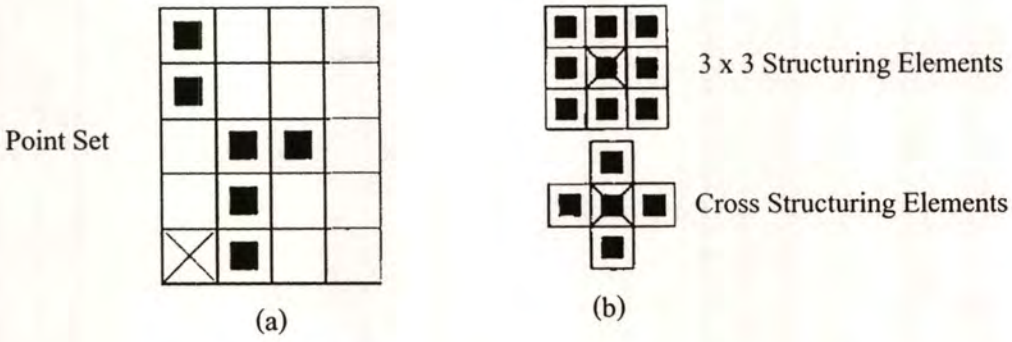


Figure 3.9 (a) A point set example and (b) some typical structuring elements

Any point x from a discrete image $X = \{(1,0), (1,1), (1,2), (2,2), (0,3), (0,4)\}$ can be treated as a vector with respect to the origin (0,0). It is assumed that only black pixels are considered, and the others are treated as a background.

A morphological transformation Ψ is given by the relation of the image (point set X) with another small point set B called a *structuring element*. Some typical structuring elements are shown in Figure 3.8 (b). B is expressed with respect to a local origin O (called the representative point). Applying morphological transformation $\Psi(X)$ to the image X means that the structuring element B is moved systematically across the entire image. Assume that B is positioned at some point in the image; the pixel in the image corresponding to the representative point O of the structuring element is called the *current pixel*. The result of the relation (which can be either zero or one) between the image X and the structuring element B in the current position is stored in the output image in the current image pixel position.

The primary morphological operations are dilation and erosion. The morphological transformation dilation $\oplus$ combines two sets using vector addition (or *Minkowski* set addition, e.g., $(a, b) + (c, d) = (a + c, b + d)$). The dilation $X \oplus B$ is the point set of all possible vector additions of pairs of elements, one from each of the sets X and B

$$X \oplus B = \left\{ p \in \varepsilon^2 : p = x + b, x \in X, b \in B \right\} \tag{3.2}$$

Dilation operation will expand the black area of X. Figure 3.10 illustrates an example of dilation:

$X = \{(1,0), (1,1), (1,2), (2,2), (0,3), (0,4)\}$

$B = \{(0,0), (1,0)\}$

$X \oplus B = \{(1,0), (1,1), (1,2), (2,2), (0,3), (0,4), (2,0), (2,1), (2,2), (3,2), (1,3), (1,4)\}$

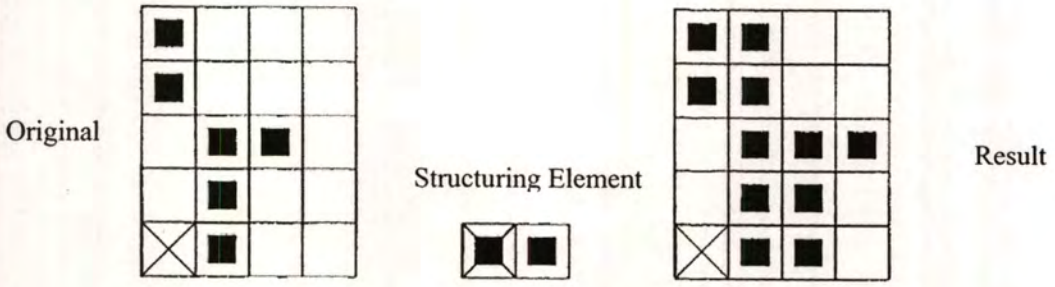


Figure 3.10 Dilation operation

Erosion Θ combines two sets using vector subtraction of set elements and is the opposite operation of dilation. The erosion of $X \Theta B$ can be explained as:

$$X \Theta B = \left\{ p \in \varepsilon^2 : p = x + b \in X \text{ for every } b \in B \right\} \tag{3.3}$$

The formula says that every point p from the image is tested, the result of the erosion is given by those point p for which all possible $x + b$ are in X. Figure 3.11 shows an example of the point set X eroded by the structuring element B. Erosion operation will contract the black area of X:

X = {(1,0), (1,1), (1,2), (0,3), (1,3), (2,3), (3,3), (1,4)}

B = {(0,0), (1,0)}

$X \Theta B$ = {(0,3), (1,3), (2,3)}

In our experiment, we used 3 x 3 structuring element to do erosion operation.. Figure 3.12 (a) and Figure 3.12 (b) show the binary images of pupil which still have remaining noise. The erosion operation yielded free-noise binary images as shown in Figure 3.12 (c) and Figure 3.12 (d).

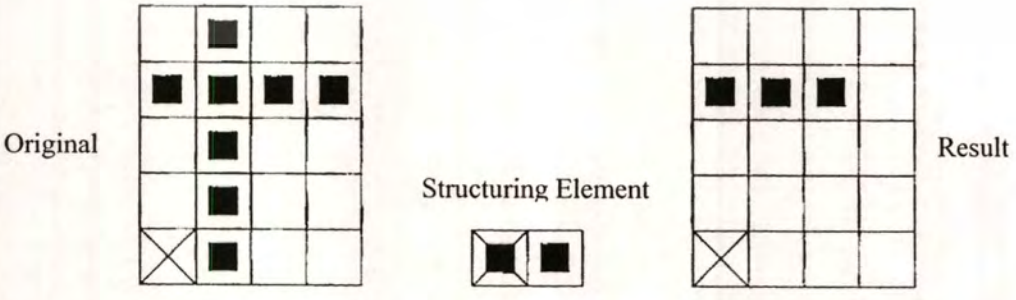


Figure 3.11 Erosion operation

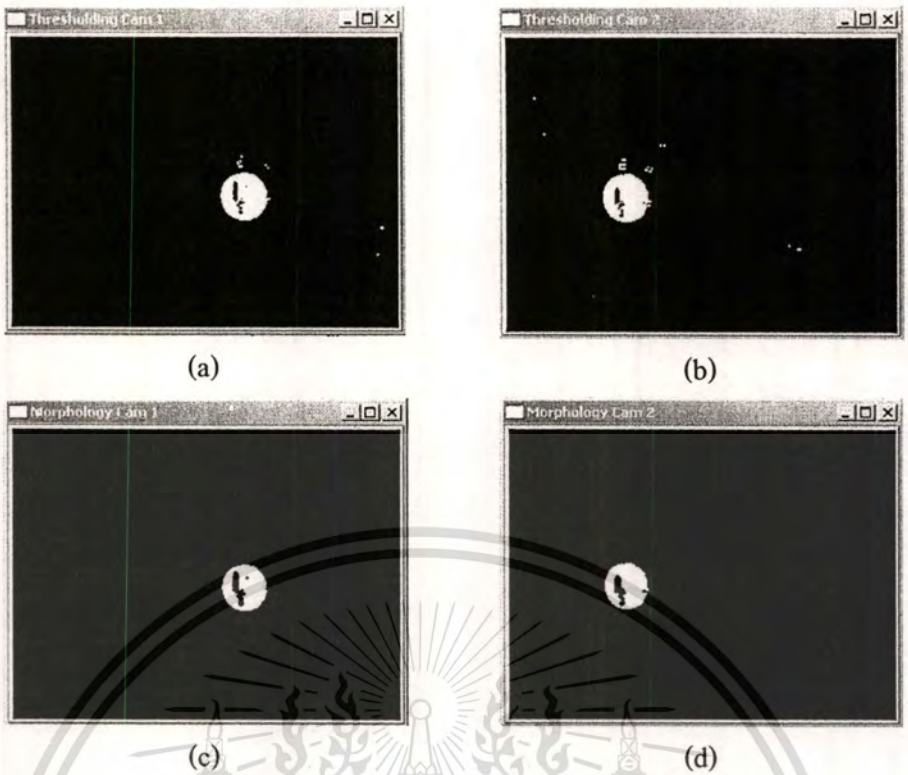


Figure 3.12 Binary eye images acquired from (a) camera 1; (b) camera 2 and morphology result on images acquired from (c) camera 1; (d) camera 2

3.2.2.3 Gaussian Smoothing Filter

Smoothing filter is used for blurring and noise reduction. Blurring is used in pre-processing steps, such as removal of small details from an image prior to (large) object extraction, and bridging small gaps in lines or curves. Gaussian smoothing filter is commonly used due to its ability to simultaneously optimize two criteria of smoothing filter: first, the filter should be smooth and roughly band limited in the frequency domain to reduce the possible number of frequencies at which function changes can take place and second, the constraint of spatial localization requires the response of a filter to be from nearby points in the image [46].

According to Ref. [47], General Gaussian distribution $G(x)$ in 1D is given by:

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} \tag{3.4}$$

where x is 1D data and σ is a standard deviation. We have also assumed that the distribution has a mean of zero (*i.e.* it is centered about the line $x = 0$). The distribution is illustrated in Figure 3.13 below:

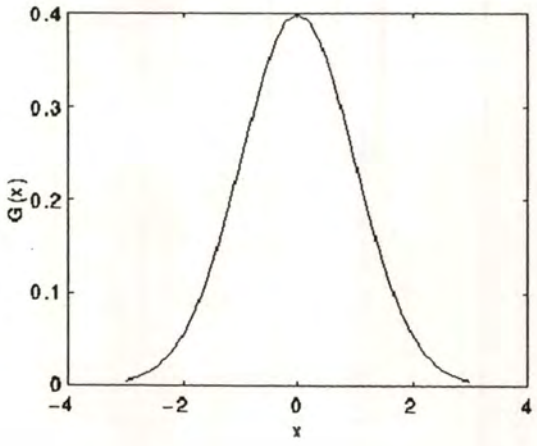


Figure 3.13 Graphical plot of 1D Gaussian distribution (mean = 0; σ =1)

In theory, the Gaussian distribution is nonzero everywhere, which would require an infinitely large convolution kernel. However, in practice it is effectively zero more than about three standard deviations from the mean. Hence, the convolution kernel can be truncated at ± 3σ [47]. A 2D isotropic (circularly symmetric) Gaussian smoothing filter is formed by:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-\left(x^2+y^2\right)}{2\sigma^2}} \tag{3.5}$$

where x and y are the 2D images coordinates and σ is a standard deviation of the associated probability distribution. The 2D Gaussian also can be written as a pair of 1D Gaussian equations as follows:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-\left(x^2+y^2\right)}{2\sigma^2}} = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-y^2}{2\sigma^2}} = G(x).G(y) \tag{3.6}$$

The distribution of 2D Gaussian with mean (0, 0) and σ =1 is depicted in Figure 3.14 below:

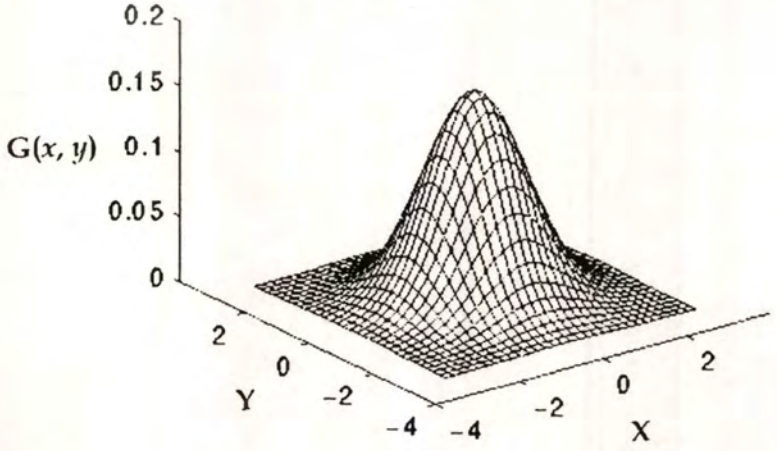


Figure 3.14 Graphical plot of 2D Gaussian distribution (mean (0, 0); σ =1)

OpenCV allows developer to define the size of Gaussian kernel used to convolve the image. As explained in [48], OpenCV performs optimum computational speed with several common kernels, 3-by-3, 5-by-5, and 7-by-7. The σ value will be defined as follows:

$$\sigma = \left(\frac{n}{2} - 1\right) * 0.30 + 0.80 \tag{3.7}$$

where σ is the standard deviation, n is the size of kernel (n = 3, 5, 7). In the other side, OpenCV also allows the developers to define the value of σ instead of the kernels size. The kernels size will be automatically determined from the value of σ. Once a suitable kernel has been calculated, then the Gaussian smoothing can be performed using standard convolution methods. The convolution can in fact be performed fairly quickly since the equation for the 2D isotropic Gaussian is separable into x and y components as explained by Eq. (3.6). Thus the 2D convolution can be performed by first convolving with a 1D Gaussian in the x direction, and then convolving with another 1D Gaussian in the y direction. In our experiment, we chose 7-by-7 kernel since it showed optimum smoothing result compared to the other OpenCV's standard kernel size as shown in Figure 3.15 below.

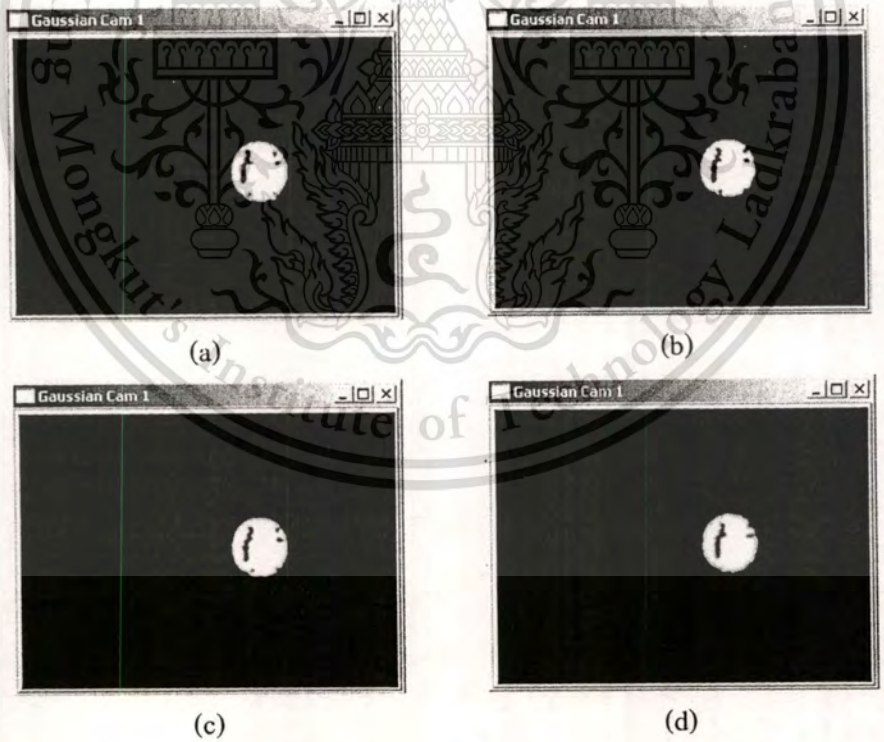


Figure 3.15 Several results of Gaussian smoothing filter implementation on binary pupil image: (a) σ = 1 ; (b) 3-by-3 kernel; (c) 5-by-5 kernel; (d) 7-by-7 kernel

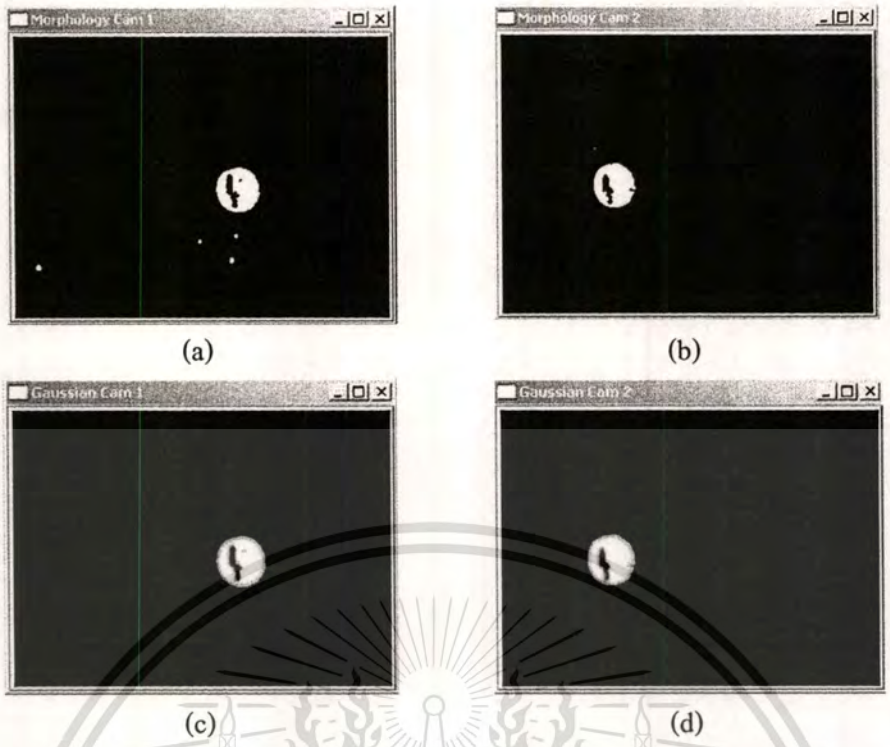


Figure 3.16 Morphology result on images acquired from (a) camera 1; (b) camera 2 and Gaussian smoothing filter result on images acquired from (c) camera 1; (d) camera 2

Figure 3.16 (a) and Figure 3.16 (b) show morphological image processing results of camera 1 and camera 2, respectively. The images were then processed with Gaussian smoothing filter results, as shown in Figure 3.16 (c) and Figure 3.16 (d). The results show that Gaussian smoothing filter connected small gaps occurred in the pupil border by blurring it, as depicted in Figure 3.16(d). The Gaussian smoothing filter will be helpful to obtain better boundary detection result on the next step.

3.2.2.4 Boundary Detection

After the pupil image has been segmented and filtered from noise, the resulting aggregate of segmented pixels usually is represented in a form suitable for further computer processing. In general, representing a region involves two choices: (1) the region is represented in terms of its external characteristic (its boundary), or (2) the region is represented in terms of its internal characteristics (the pixel comprising the region) [49]. An external representation is chosen when the primary focus is on shape characteristics. An internal representation is selected when the primary focus is on regional properties, such as colour and texture. Border following is one of the fundamental techniques in the processing of digitized binary images. It derives a sequence of the coordinates or the chain codes from the border between a region and its background. The border following technique has been studied deeply, because it has a large variety of applications, including picture recognition, picture analysis, and image data compression [50-52].

In our experiment, we used border following algorithm explained in [49] to extract the outer boundary of a binary pupil region. The border following algorithm will extract the boundary in a clockwise direction. Consequently, the output of the border following algorithm is an ordered sequence of points. It was assumed that (1) we are working with binary images in which object and background points are labeled 1 and 0, respectively, and (2) that images are padded with a border of 0s to eliminate the possibility of an object merging with the image border. Figure 3.17 shows illustration of several steps of border following algorithm. Given a binary region R or its boundary as illustrated in Figure 3.17 (a), an algorithm for following the border of R consists of the following steps:

1. Let the starting point, b_0, be the uppermost, leftmost point in the image that is labeled 1. Denote by c_0 the west neighbour of b_0 (Figure 3.17 (b)). Clearly, c_0 is always a background point. Examine the 8-neighbours of b_0, starting at c_0 and proceeding in a clockwise direction. Let b_1 denote the first neighbour encountered whose value is 1, and let c_1 be the (background) point immediately preceding b_1 in the sequence. Store the locations of b_0 and b_1 for use in Step 5.

2. Let $b = b_1$ and $c = c_1$ (Figure 3.17 (c)).

3. Let the 8-neighbours of b, starting at c and proceeding in a clockwise direction, be denoted by $n_1, n_2, \ldots\ldots, n_8$. Find the first n_k labeled 1 (Figure 3.17 (d)).

4. Let $b = n_k$ and $c = n_{k-1}$.

5. Repeat steps 3 and 4 until $b = b_0$ and the next boundary point found is b_1. The sequence of b points found when the algorithm stops constitutes the set of ordered boundary points.

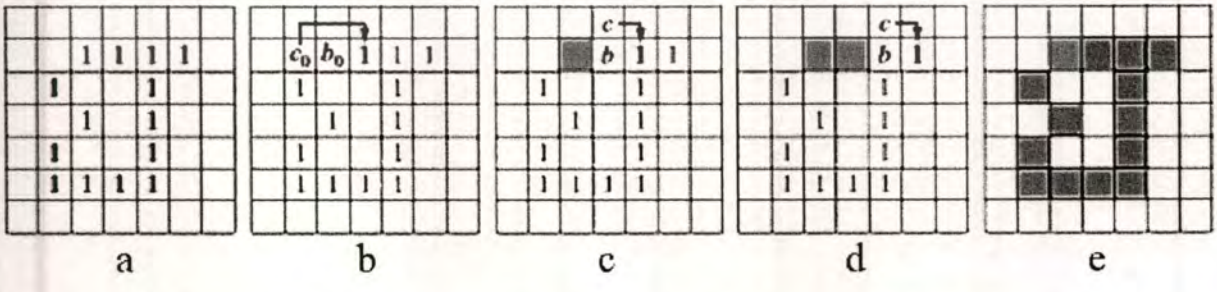


Figure 3.17 Illustration of the steps in border following algorithm. The point to be processed next is pointed by the arrow (labeled in black), the points yet to be processed are gray, and the points found by the algorithm are labeled as gray squares

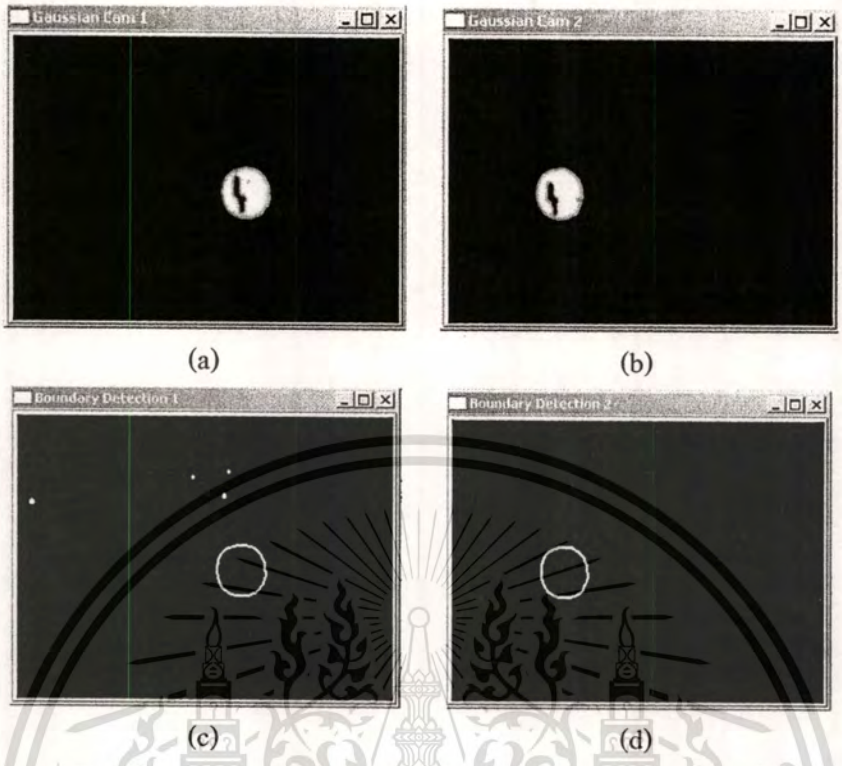


Figure 3.18 Smoothed binary pupil on images acquired from (a) camera1; (b) camera 2 and boundary detection result on images acquired from (c) camera 1; (d) camera 2

It is easily verified that continuing with this procedure will yield the correct boundary shown in Figure 3.17(e), whose point are clockwise-ordered sequence. Figure 3.18 shows the smoothed binary pupil images and the boundary detection algorithm results of camera 1 and camera 2.

3.2.2.5 Center of Mass

A quantitative measure of the shape of a set numeric values at some distance from a reference point or axis is represented by moments. In statistics, the distribution of random variables is characterized by moments [53]. In image processing research area, moments are utilized to compute center of mass of an object which is required to detect 2D position of object in spatial domain. Boundary information resulted from previous process can be used to describe a region based on the chosen representation. In our thesis, we used moments to describe the extracted pupil boundary by considering the pupil boundary as a two-dimensional density distribution function. The moments values were then used to derive center of mass of the pupil.

The 2D Cartesian moment m_{pq} of order $p + q$ with density distribution function $f(x, y)$ is

$$m_{pq} = \int_{-\infty}^{\infty} x^p y^q f(x,y)\, dx\, dy \tag{3.8}$$

According to [53], the 2D moment for a M x N discretized image $g(x, y)$ is

$$m_{pq} = \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} x^p y^q g(x,y) \tag{3.9}$$

In a discrete image, a moment is a gross characteristic of the boundary computed by integrating or summing over all pixel of the boundary [54]. A zero order moment (m_{00}) represents the total mass of the given distribution function or image. When computed a boundary of object in discrete space, the zero order moment represents the total object area.

The zero order moment (m_{00}) is

$$m_{00} = \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} g(x,y) \tag{3.10}$$

The two first order moments (m_{10}, m_{01}) are used to locate center of mass (COM) of the object. The coordinates of the COM, ($\bar{x}$, $\bar{y}$), is the intersection of the lines $x = \bar{x}$ and $y = \bar{y}$, parallel to the x and y axes respectively, about which the first order moments are zero. Alternatively, $x = \bar{x}$ and $y = \bar{y}$ represent lines where all the mass may be concentrated without change to the first order moments about the x and y axes respectively.

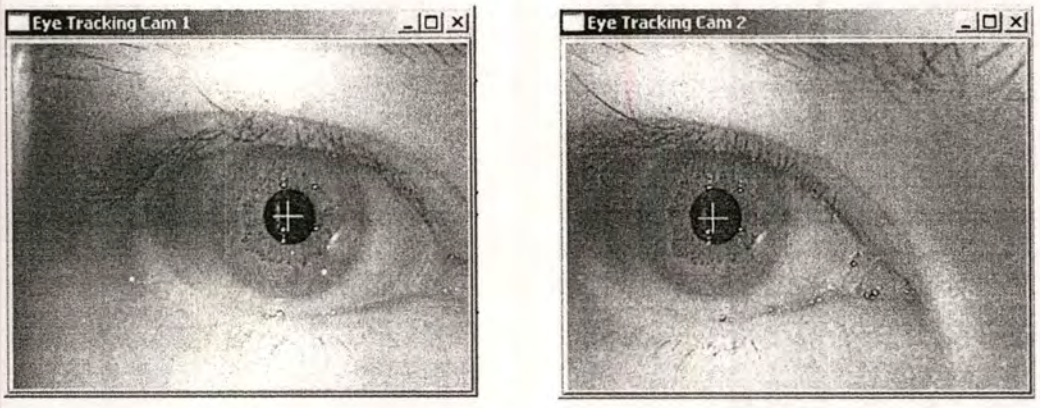
The first order moments (m_{10}, m_{01}) are defined as:

$$m_{10} = \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} x\ g(x,y) \tag{3.11}$$

$$m_{01} = \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} y\ g(x,y) \tag{3.12}$$

In terms of moments values, the coordinates of the COM are defined as:

$$\bar{x} = \frac{m_{10}}{m_{00}} \qquad \bar{y} = \frac{m_{01}}{m_{00}} \tag{3.13}$$



(a) (b)

Figure 3.19 Results of pupil tracking on images acquired from (a) camera 1; (b) camera 2

Figure 3.19 (a) and Figure 3.19 (b) show the results of pupil tracking algorithm on camera 1 and camera 2, respectively. The preceding preprocessed gray-scale images were copied and saved in the memory to display the pupil tracking results. The moments were extracted from the boundary of segmented pupil images. The coordinates of the COM were then used to define a unique 2D position that is used as a reference point to describe the positions of the pupil on each captured images. A white cross was then drawn on each image to acknowledge the 2D position of the pupil.

3.2.3 Iris Striation Tracking

The iris striation tracking is performed as a part of torsional eye movement measurement. In order to achieve better tracking performance and avoid false detection, we define the searching area based on the coordinates of COM obtained using pupil tracking algorithm. The searching area is known as Region of Interest (ROI).

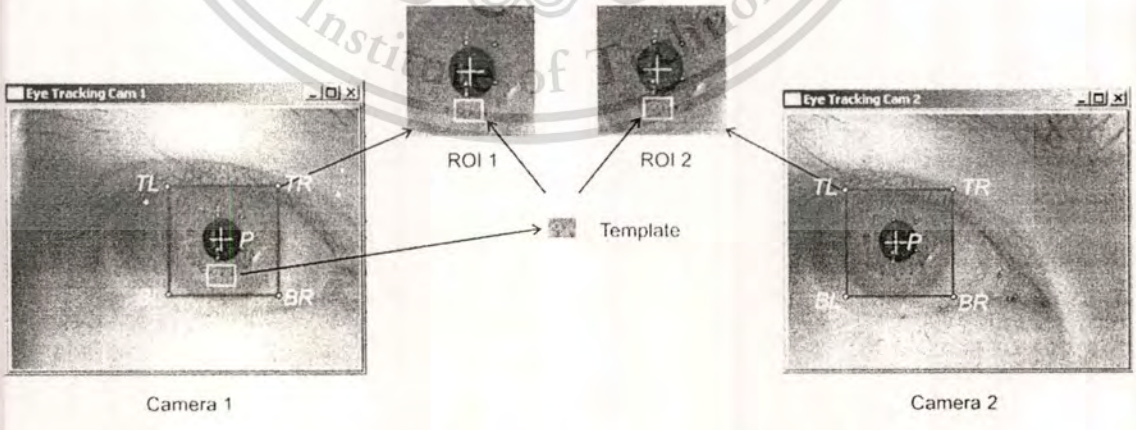


Figure 3.20 Iris striation tracking inside the Region of Interest

Figure 3.20 shows the iris striation tracking inside the ROI. The procedure can be explained as follows:

1. The iris template is manually selected from gray-scale eye image using Microsoft® Paint by operator. The cropping process is done before the experiment start. In order to obtain precise iris striation position on both cameras, the template is obtained from the best eye image of camera 1 or camera2. The size of the template is 19 x 14 pixels.

2. The preceding pupil tracking algorithm has resulted the position of pupil center. In order to achieve better performance and avoid false detection, the searching area known as ROI is automatically defined by the algorithm. The size of ROI is 100 x 100 pixels. Given the pupil center coordinates $P(x_p, y_p)$, the position of the ROI is

   Top-left point $TL(x_{tl}, y_{tl})$:

   $$(x_{tl}, y_{tl}) = (x_p - 50, y_p - 50) \tag{3.14}$$

   Top-right point $TR(x_{tr}, y_{tr})$:

   $$(x_{tr}, y_{tr}) = (x_p + 50, y_p - 50) \tag{3.15}$$

   Bottom-left point $BL(x_{bl}, y_{bl})$:

   $$(x_{bl}, y_{bl}) = (x_p - 50, y_p + 50) \tag{3.16}$$

   Bottom-right point $BR(x_{br}, y_{br})$:

   $$(x_{br}, y_{br}) = (x_p + 50, y_p + 50) \tag{3.17}$$

3. Template matching algorithm with normalized correlation coefficient is applied inside the ROI to track the iris striation on both images captured from camera 1 and camera 2. Template matching with normalized correlation coefficient was utilized to detect iris striation as this convolution-based algorithm is more intuituve and easier to be implemented on gray-level image [55].

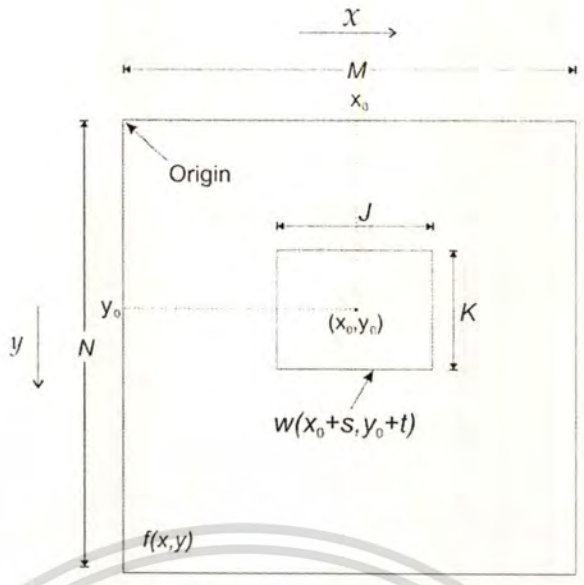


Figure 3.21 Arrangement for obtaining the correlation of f and w at point (x_0, y_0)

The basic concept of template matching algorithm with normalized correlation coefficient is shown in Figure 3.21 [55]. The picture also shows the 2D Euclidean coordinate system reference for the template matching algorithm where the x and y axes are defined on horizontal and vertical direction, respectively. The template matching is conducted by sliding a template $w(x, y)$ of size J x K within an image $f(x, y)$ of M x N, where we assume that $J \leq M$ and $K \leq N$. The similarity measurement is determined by observing the normalized correlation coefficient $\rho(x, y)$ below:

$$\rho(x,y) = \frac{\sum_t \sum_s \left[f(s,t) - \overline{f}(s,t) \right] \left[w(x+s, y+t) - \overline{w} \right]}{\sqrt{\sum_t \sum_s \left[f(s,t) - \overline{f}(s,t) \right]^2 \sum_t \sum_s \left[w(x+s, y+t) - \overline{w} \right]^2}}, \quad -1 \leq \rho(x,y) \leq 1 \qquad (3.18)$$

where x = 0, 1, 2, 3, ..., M–1 ; y = 0, 1, 2, 3, ..., N–1; s = 0, 1, 2, 3, ..., J–1; and t = 0, 1, 2, 3, ..., K–1. $\overline{w}$ is the mean value of the pixels in w (computed only once), $\overline{f}$ is the mean value of f in the region coincident with the current location of w, and the summations are taken over the coordinates common to both f and w. The normalized correlation coefficient $\rho(x, y)$ is scaled in the range of -1 to 1. A perfect match is $\rho(x, y)$ equals to 1 and a perfect mismatch is $\rho(x, y)$ equals to -1. The normalization of correlation coefficient is useful to reduce the effects of lighting differences between the template w and the image f, as explained in [56]. We assume that the origin of f is at its top-most left and the origin of w is at its center. For *one* value of (x, y), say, (x_0, y_0) inside f, application of Eq. (3.18) yields *one* value of ρ. As x and y are varied, template w moves around the image area, giving the function $\rho(x, y)$. In practical implementation, the maximum value of ρ indicates the position where w best matches f.

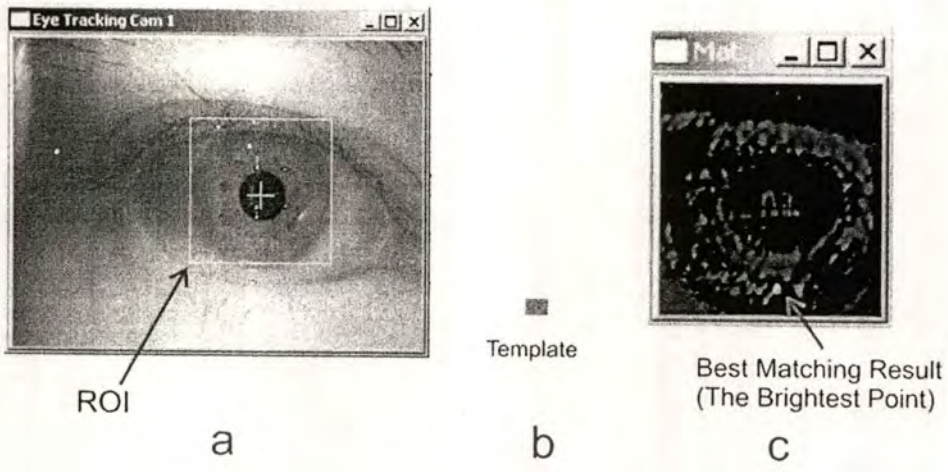


Figure 3.22 Iris striation tracking: (a) eye image with pupil center coordinates and ROI; (b) iris template; (c) best matching result is shown by the brightest point

Figure 3.22 shows practical implementation of template matching using normalized correlation coefficient on camera 1. Figure 3.22(a) shows an eye image with defined ROI. Figure 3.22(b) is the iris template. The image area inside the ROI and the iris template are $f(x, y)$ and $w(x, y)$, respectively. The normalized correlation coefficient $\rho(x, y)$ of the ROI is shown as an image in Figure 3.22(c). Note that Figure 3.22(c) has been enlarged to show the detail of image. The brightest point of $\rho(x, y)$ is in the position where the best match between f and w was found. The 2D coordinates of the center of iris template is used in torsional eye movement measurement.

3.2.4 Camera Modeling and Perspective Transformation

Most two-dimensional images are views of three-dimensional scenes from the physical perspective of a camera imaging the scene. It is often desirable to modify an observed image so as to stimulate an alternative point of view. This can be accomplished by use of a *perspective transformation* [57]. In this thesis, we use a pinhole camera model to get a handle on basic geometry of projecting rays. A *pinhole* is an imaginary wall with a tiny hole in the center that blocks all rays except those passing through the tiny aperture in the center [58]. Figure 3.23 shows the concept of pinhole camera model that projects points of light in 3D object space ρ^3 to points of light in 2D image plane ρ^2. In this imaging model, the *world coordinate system* and *camera coordinate system* are denoted by capital (X, Y, Z) and small letters (x, y, z), respectively. It is assumed that the origin of camera coordinate system coincides with the origin of world coordinate system at coordinates $C(0, 0, 0)$.

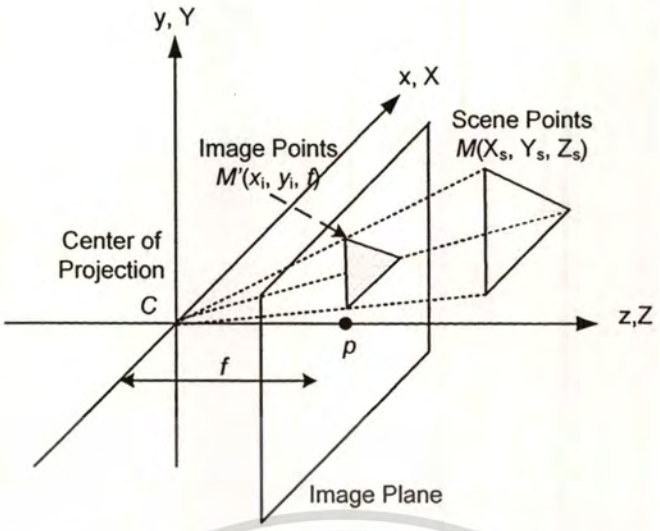


Figure 3.23 Perspective transformation on pinhole camera model
if world coordinate system coincides with camera coordinate system

Let $M(X_S, Y_S, Z_S)$ be the coordinates of object point in the scene and $M'(x_i, y_i, f)$ be the coordinates of projected point on the image plane. C is the center of projection. The optical axis of the camera is assumed passing through the center of image plane. The center point $p(x_0, y_0)$ of image plane is called principal point. The camera focal length is denoted as f. Using similar triangles, it is easy to establish that

$$x_i = f\frac{X_S}{Z_S},\ y_i = f\frac{Y_S}{Z_S} \tag{3.19}$$

The projected point $M'(x_i, y_i, f)$ is related nonlinearly to the object point $M(X_S, Y_S, Z_S)$. This relationship can be simplified by utilization of homogeneous coordinates, as introduced to the image processing community by Roberts [59]. Given vector v as a vector containing the object point coordinates

$$v = \begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix} \tag{3.20}$$

The homogeneous vector $\bar{v}$ corresponding to v is

$$\bar{v} = \begin{bmatrix} X_S \\ Y_S \\ Z_S \\ 1 \end{bmatrix} \tag{3.21}$$

Consider the following perspective transformation matrix:

$$P = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{3.22}$$

Forming the vector product

$$\bar{w} = P\bar{v} \tag{3.23}$$

Yields

$$\bar{w} = \begin{bmatrix} fX_S \\ fY_S \\ Z_S \end{bmatrix} \tag{3.24}$$

The corresponding image plane coordinates are obtained by normalized the third vector element z_s into 1 which obtains

$$w = \begin{bmatrix} f\dfrac{X_S}{Z_S} \\ f\dfrac{Y_S}{Z_S} \end{bmatrix} \tag{3.25}$$

It can be observed that the first two elements of w correspond to the imaging relationship explained by Eq. (3.19).

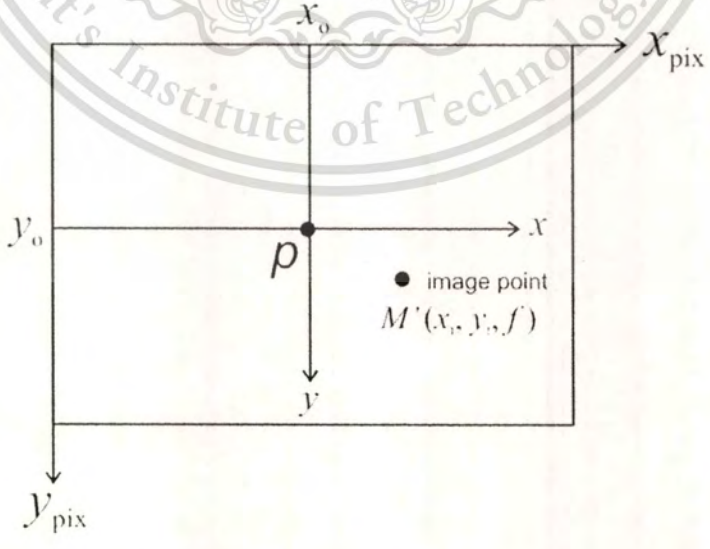


Figure 3.24 The principal point illustration on the real image plane

The next step is converting the coordinates of projected point $M'(x_i, y_i, f)$ to pixels unit. In the real situation, the center of CCD sensor is usually not on the optical axis. Therefore, a model of possible displacement (away from the optical axis) of the center of coordinates on the projection screen must be defined by involving the principal point parameter. Given that the origin of the real image sensor is located on the top-most left corner, the principal point $p(x_0, y_0)$ is defined at the center of the image plane as illustrated in Figure 3.24. Conversion to pixels unit also requires computation of adjustment scale (*aspect ratio*) of each axis of image coordinate system. Individual pixel in real CCD sensor is rectangular rather than square. Therefore, we use two different aspect ratio values for x and y axes. Skew parameter is involved in physical to pixels conversion. Skew parameter measures skewness of rows and columns in CCD sensor. The skew parameter will be zero for the most normal camera. However, in some conditions, skew parameter is not zero due to x and y axes of image coordinate system are not perpendicular each other. This condition can be seen on image enlargement taking by pinhole camera where the axis of the magnifying lens is not perpendicular to the enlarged image plane [60].

Considering several parameters above, point coordinates in pixels unit (x_{pix}, y_{pix}) can be formulated as follows:

$$\begin{aligned} x_{pix} &= k_x x_i + s y_i + x_0 \\ &= k_x f \frac{X_S}{Z_S} + sf \frac{Y_S}{Z_S} + x_0 \end{aligned} \tag{3.26}$$

$$\begin{aligned} y_{pix} &= k_y y_i + y_0 \\ &= k_y f \frac{Y_S}{Z_S} + y_0 \end{aligned} \tag{3.27}$$

where k_x is aspect ratio in x-axis, k_y is aspect ratio in y-axis, and $S$is skew parameter. If the projected point coordinates (x_{pix}, y_{pix}) are written as matrix $\tilde{u}$ in homogeneous coordinates

$$\tilde{u} = \begin{bmatrix} u & v & w \end{bmatrix}^T, \text{ where } x_{pix} = \frac{u}{w} \text{ and } y_{pix} = \frac{v}{w} \tag{3.28}$$

then we can rewrite perspective transformation equation as follows:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = G.\bar{v} = \begin{bmatrix} \alpha_x & s & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_S \\ Y_S \\ Z_S \\ 1 \end{bmatrix} \tag{3.29}$$

where $\alpha_x = k_x f$ and $\alpha_y = k_y f$ is the pixels unit values of focal length f in the x direction and y direction, respectively. Matrix G can be separated into two matrices as follows:

$$G = \begin{bmatrix} \alpha_x & s & x_0 & 0 \\ 0 & \alpha_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} = K[I_3 | O_3] \tag{3.30}$$

Matrix K is 3 x 3 an upper-triangle matrix called *camera calibration matrix* that contains *intrinsic parameter* of camera. Intrinsic parameter of camera consists of three physical parameters: focal length, skew parameter, and principal point.

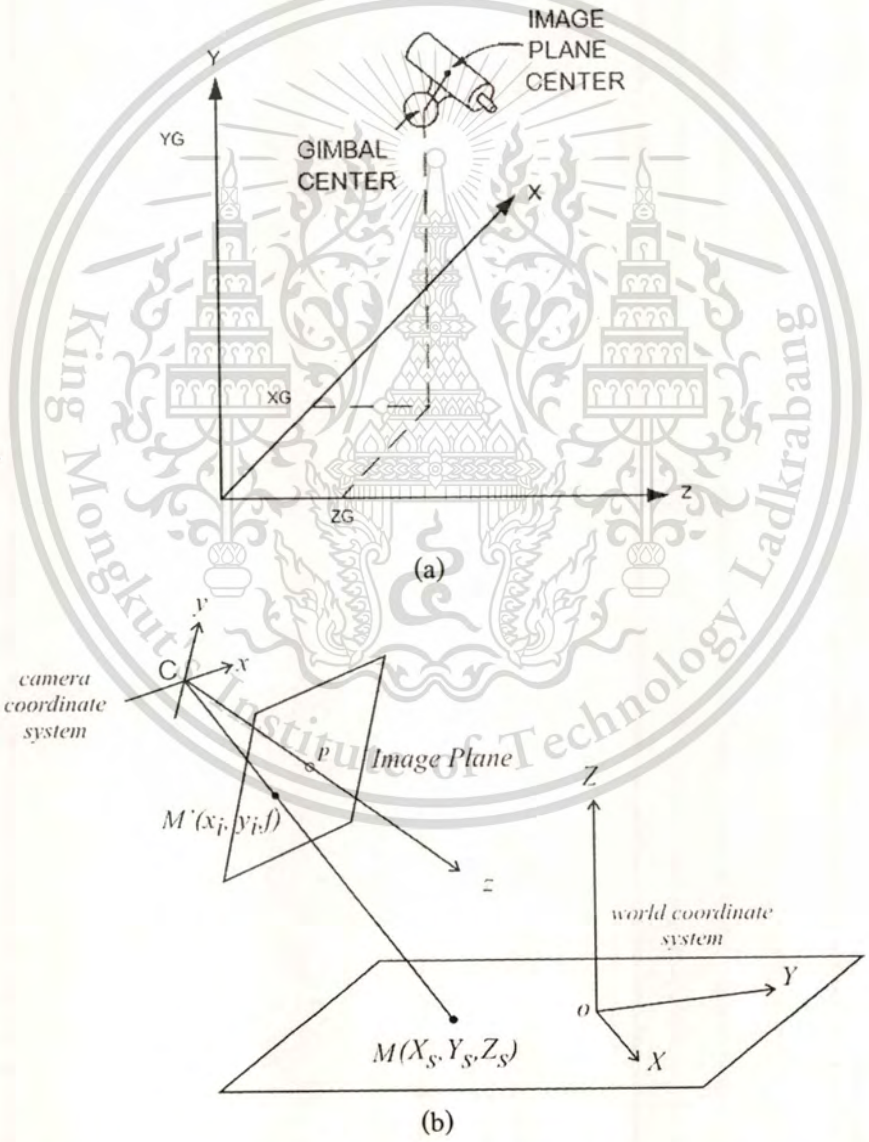


Figure 3.25 (a) Real industrial pinhole camera located in world coordinate system [57]; (b) Perspective transformation on pinhole camera model if world coordinate system does not coincide with camera coordinate system

Pinhole camera model described in Figure 3.23 based on assumption that the center of world coordinate system coincides with the center of the camera coordinate system in order to facilitate the analysis of basic pinhole camera model. In Figure 3.25, a new assumption is discussed to handle physical camera used in practical imaging geometry [61]. Figure 3.25 (a) shows the real industrial pinhole camera located at the world coordinate system. The camera is physically installed in 2D gimbal system so that the camera can spin horizontally (*pan*) with horizontal angle β and spin vertically (*tilt*) with vertical angle α. The gimbal system center located at (X_G, Y_G, Z_G) is assumed as the center of camera coordinate system related to the world coordinate system. It should be noted that the image plane center and the gimbal system center are offset by a vector with coordinate (X_O, Y_O, Z_O). Therefore, *2D rotational offset*, *gimbal-to-world coordinate translation offsets*, and *camera-to-gimbal translation offsets* must be considered in this new assumption.

Figure 3.5 (b) shows the pinhole camera model based on new assumption that the camera coordinate system does not coincide directly with the world coordinate system. If the camera were to be located at the center of the world coordinate system, not panned nor tilted with respect to the reference axes, and if the camera image plane was not offset with respect to the gimbal, the homogeneous image model would be as derived in Eq. (3.23)

$$\overline{w} = P\overline{v} \tag{3.31}$$

where $\overline{v}$ is the homogeneous vector of the world coordinates of an object point, $\overline{w}$ is the homogeneous vector of the image plane coordinates, and P is the perspective transformation matrix defined by Eq. (3.22). In this new assumption, we have to modify Eq. (3.31) in order to compensate the rotational and translational offsets.

The offset of the camera to location (X_G, Y_G, Z_G) can be accommodated by the translation operation defined as follows

$$\overline{w} = PT_G\overline{v} \tag{3.32}$$

where

$$T_G = \begin{bmatrix} 1 & 0 & 0 & -X_G \\ 0 & 1 & 0 & -Y_G \\ 0 & 0 & 1 & -Z_G \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.33}$$

Pan and tilt are modeled by a rotation transformation as follows

$$\overline{w} = PRT_G\overline{v} \tag{3.34}$$

where $R = R_\beta R_\alpha$

$$R_\beta = \begin{bmatrix} \cos\beta & 0 & -\sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.35}$$

and

$$R_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.36}$$

Finally, the camera-to-gimbal offset with offset vector (X_O, Y_O, Z_O) is modeled as

$$\overline{w} = PT_C RT_G\overline{v} \tag{3.37}$$

where

$$T_C = \begin{bmatrix} 1 & 0 & 0 & -X_o \\ 0 & 1 & 0 & -Y_o \\ 0 & 0 & 1 & -Z_o \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.38}$$

As arrangement in Eq. (3.28) and (3.29), Eq. (3.37) can be rewritten in homogeneous coordinate formula as follows

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = K\left[I_3 \mid O_3\right]\begin{bmatrix} R & -T \\ O_3^T & 1 \end{bmatrix}\begin{bmatrix} X_S \\ Y_S \\ Z_S \\ 1 \end{bmatrix} \tag{3.39}$$

where R is composite rotation matrix (size 3 x 3) consists of total rotation with respect to world coordinate system, T is translation vector (size 3 x 1) consists of total translation with respect to world coordinate system. K is camera calibration matrix (size 3 x 3) explained in Eq. (3.30).

It can be said that matrix

$$\begin{bmatrix} R & -T \\ O_3^T & 1 \end{bmatrix} \tag{3.40}$$

is a matrix that incorporates external calibration parameters (*extrinsic parameters*) of camera movement. If we express the coordinates of object point $M(X_S, Y_S, Z_S)$ as X_w, than vector $\bar{v}=[X_s \quad Y_s \quad Z_s \quad 1]^T$ can be expressed as vector $\widetilde{X}_w$ (size 2 x 1) as follows

$$\widetilde{X}_W = \begin{bmatrix} X_W \\ 1 \end{bmatrix} \tag{3.41}$$

Matrix $\tilde{u}=[u \quad v \quad w]^T$ in Eq. (3.39) can be modeled as

$$\tilde{u} = KR(X_W - t) \tag{3.42}$$

The model described in Eq. (3.42) implies that matrix $\tilde{u}$ can also be written as follows

$$\tilde{u}=[KR \mid -KRt]\begin{bmatrix} X_W \\ 1 \end{bmatrix} = M\begin{bmatrix} X_W \\ 1 \end{bmatrix} = M\widetilde{X}_W \tag{3.43}$$

M is 3 x 4 matrix which is called *projection matrix* (or *camera matrix*) and $\widetilde{X}_w$ denotes the 3D object point in homogeneous coordinates. The 3D object point is expressed up to scaling factor α in homogeneous coordinates. Hence, all αM are equivalent for $\alpha \neq 0$. If the real 2D coordinates of object in pixel unit is represented by vector

$$r = [x_{pix} \quad y_{pix}]^T = [p \quad q]^T \tag{3.44}$$

Then we can observe that a set of n noncoplanar points lies in 3D object space with coordinates X_n and its corresponding 2D n points with coordinates r_n yield one equation:

$$\alpha_n r_n = M\begin{bmatrix} X_n \\ 1 \end{bmatrix} \tag{3.45}$$

It can be seen that the camera can perform a linear perspective transformation from 3D object space ρ^3 to 2D image plane ρ^2 by utilizing perspective transformation model and homogeneous coordinate. The perspective transformation and homogeneous coordinate made the expression of nonlinear relationship between 3D object point $M(X_S, Y_S, Z_S)$ and its corresponding projected point $M'(x_i, y_i, f)$ simpler.

3.2.5 Camera Calibration

The camera calibration is a procedure to set numeric values in the camera calibration matrix K (Eq. (3.30)) and projection matrix M (Eq. (3.43)). Initially, the calibration procedure can be done by gaining projection matrix M from the coordinates of points with known scene positions. Then, the extrinsic and intrinsic parameters are computed from matrix M [62].

To obtain matrix M, observe that each known scene point in 3D object space $X=[X \quad Y \quad Z]^T$ and its corresponding 2D image point $r=[p \quad q]^T$ give one equation (Eq. (3.45)). We seek the numerical values m_{ij} in the 3 x 4 projection matrix M. Expanding Eq. (3.45), we can obtain

$$\begin{bmatrix} \alpha p \\ \alpha q \\ \alpha \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3.46}$$

$$\begin{bmatrix} \alpha p \\ \alpha q \\ \alpha \end{bmatrix} = \begin{bmatrix} m_{11}X + m_{12}Y + m_{13}Z + m_{14} \\ m_{21}X + m_{22}Y + m_{23}Z + m_{24} \\ m_{31}X + m_{32}Y + m_{33}Z + m_{34} \end{bmatrix} \tag{3.47}$$

Substituting α with $(m_{31}X + m_{32}Y + m_{33}Z + m_{34})$ at the first and second elements of vector $[\alpha p \quad \alpha q \quad \alpha]$ yields

$$\begin{aligned} p(m_{31}X + m_{32}Y + m_{33}Z + m_{34}) &= m_{11}X + m_{12}Y + m_{13}Z + m_{14} \\ q(m_{31}X + m_{32}Y + m_{33}Z + m_{34}) &= m_{21}X + m_{22}Y + m_{23}Z + m_{24} \end{aligned} \tag{3.48}$$

Thus we obtain two linear equations, each in 12 unknowns $m_{11},.....,m_{34}$, for each known corresponding 3D scene and 2D image point. If n such points are available, we can write Eq. (3.48) as a $2n$ x 12 matrix as follows

$$\begin{bmatrix} X & Y & Z & 1 & 0 & 0 & 0 & 0 & -pX & -pY & -pZ & -p \\ 0 & 0 & 0 & 0 & X & Y & Z & 1 & -qX & -qY & -qZ & -q \\ & & & & & & \vdots & & & & & \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{12} \\ \vdots \\ m_{34} \end{bmatrix} = 0 \tag{3.49}$$

$$AM = 0 \tag{3.50}$$

The matrix M actually has only 11 unknown parameters due to the unknown scaling factor, since homogeneous coordinate were used [63]. To generate a solution, at least six known corresponding scene and image points are required. Typically, more points are used for robustness and numerical stability. The over-determined Eq. (3.50) can be solved using Singular Value Decomposition [64].

In order to solve projection matrix M, performs an SVD of matrix A (size m x n) to derive

$$A_{mn} = U_{mm} D_{mn} V_{nn}^{T} \tag{3.51}$$

where matrix U (size m x m) is left orthogonal matrix, matrix D (size m x n) is diagonal matrix, and matrix V (size n x n) right orthogonal matrix. The last column of matrix V is the solution for matrix M.

Separation of extrinsic parameters (the rotation matrix R and translation matrix t) from the estimated projection matrix M can be done by recalling Eq. (3.43) that the projection matrix M can be written as follows

$$M = [KR \mid -KRt] = [A \mid b] \tag{3.52}$$

The 3 x 3 submatrix is denoted as A and the right-most column as b. To determine the translation vector t, submatrix $A = KR$ in Eq. (3.52) can be substituted to $b = -KRt$, so that we can write

$$t = -A^{-1}b \tag{3.53}$$

In order to determine the rotation matrix R, note that the camera calibration matrix K is upper triangular and the rotation matrix is orthogonal. The matrix factorization method called QR-Decomposition [65] can be used to decompose A in Eq. (3.52) into product of such matrices, and hence can recover K and R.

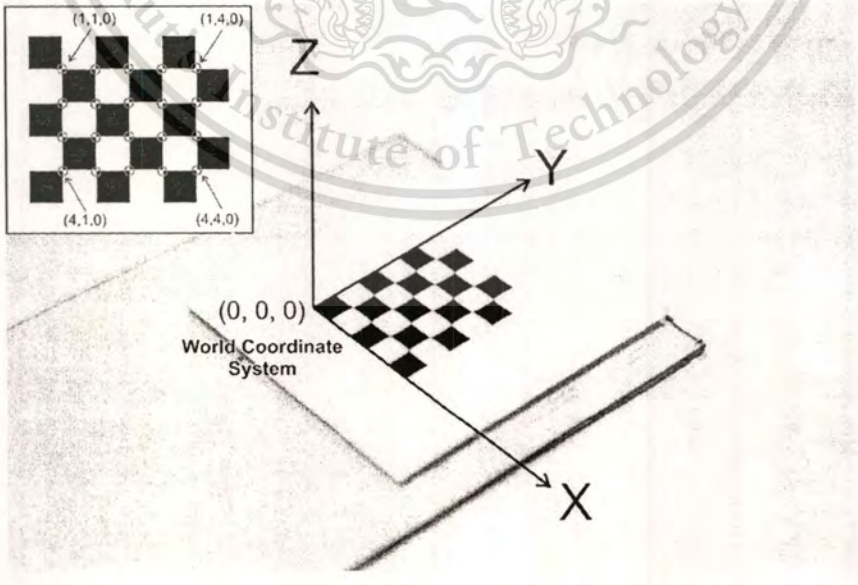


Figure 3.26 The world coordinate system of 5 x 6 tiny chessboard pattern used in camera calibration; (inset) coordinates of points on chessboard pattern.

Figure 3.26 shows a chessboard pattern used for camera calibration as used by Zhang in [66]. Our chessboard pattern consists of 5 x 6 black and white squares. The tiny square size 3 mm x 3 mm was made to accommodate close-range camera calibration. The chessboard pattern was then attached on a thin rectangular glass to allow the operator held the pattern facing the cameras. The distance between the chessboard pattern and the dual cameras system was 45 mm. The origin of the world coordinate system was set at the top-most left chessboard corner. Figure 3.26 (inset) shows coordinates (X, Y, Z) of several chessboard points acting as known scene point in 3D object space where Z = 0.

As explained in [58], the flat chessboard pattern was used in OpenCV based camera calibration because it could be constructed more precisely compared to 3D calibration object. Furthermore, flat chessboard pattern can provide multiple planar view which is useful to derive extrinsic and intrinsic parameters rather than one view of a specially constructed 3D object. Alternating black and white colors were applied on the chessboard to ensure that there is no bias toward one side or the other in measurement. Figure 3.27 shows eight distinct positions of chessboard pattern captured by camera 1 and camera 2 simultaneously.

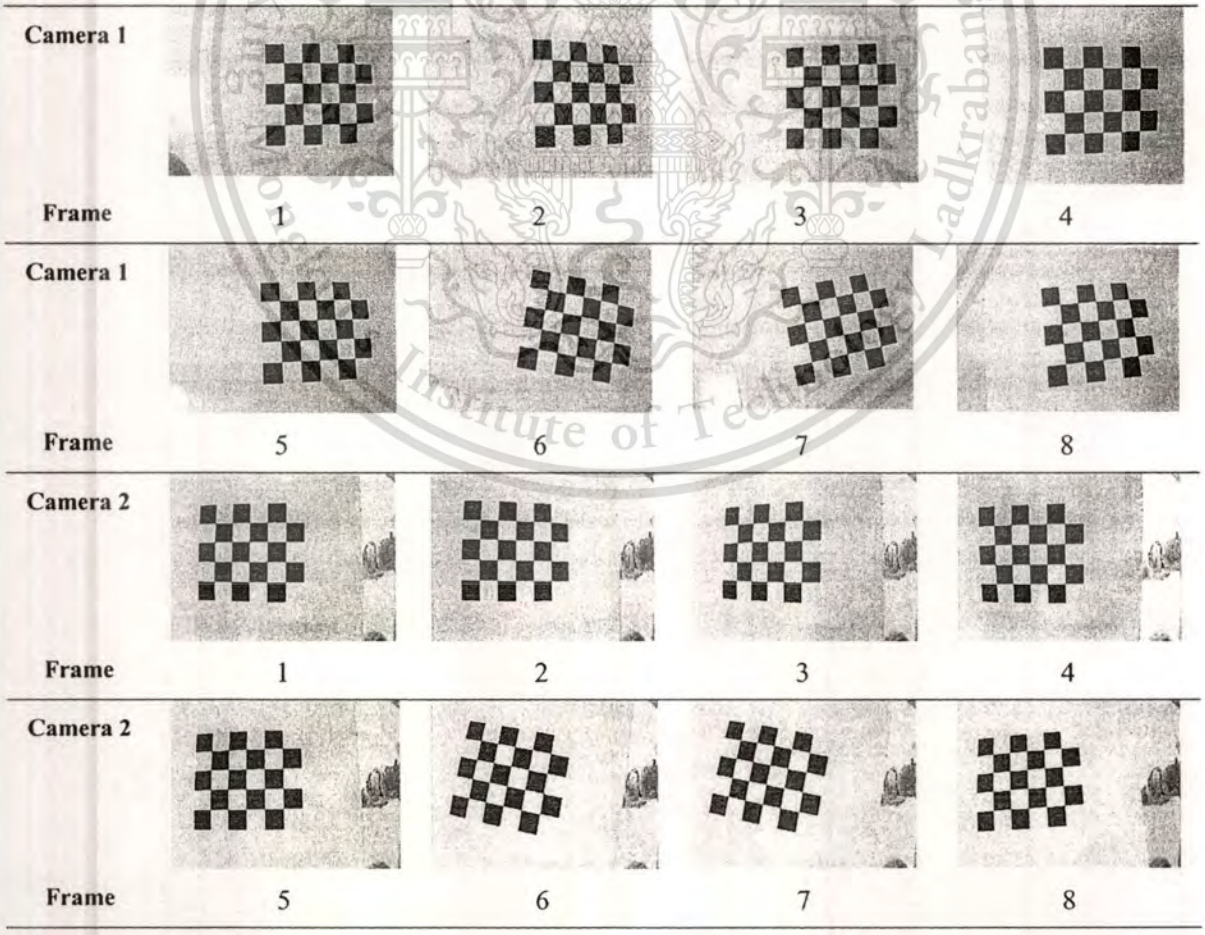


Figure 3.27 Eight distinct positions of chessboard pattern captured by camera 1 and camera 2

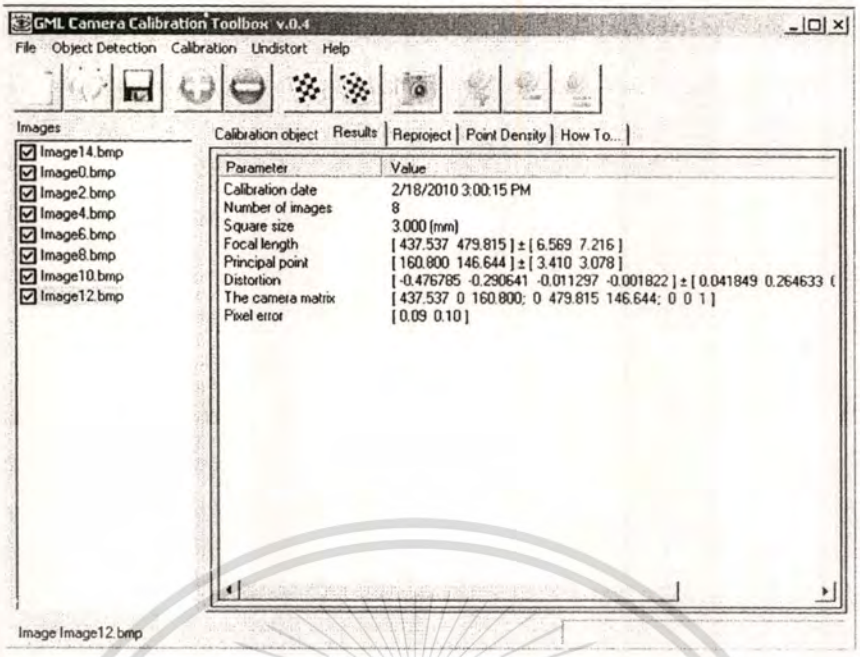


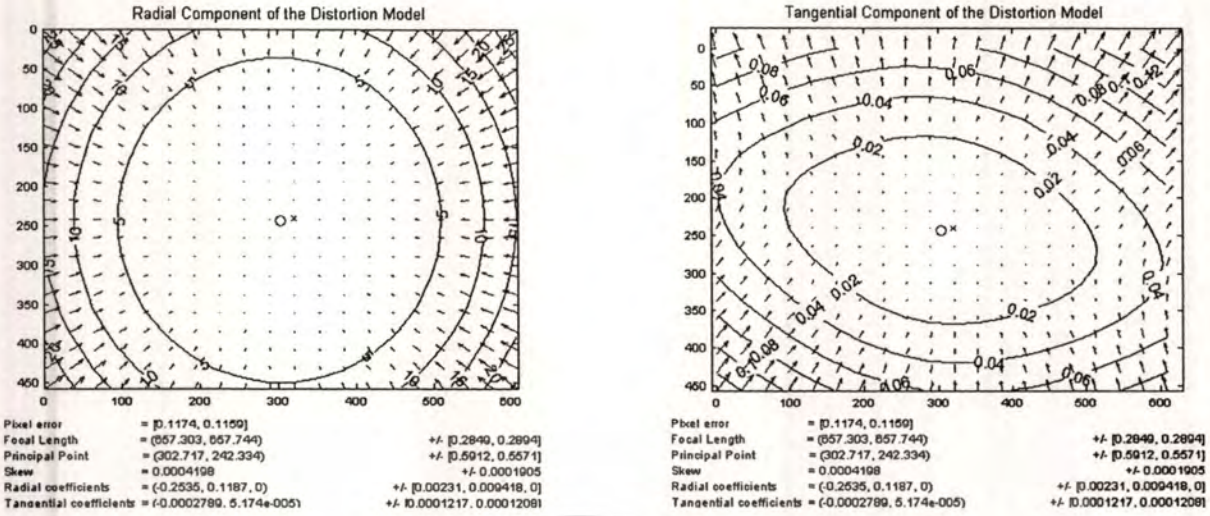
Figure 3.28 GML Camera Calibration Toolbox

The captured chessboard images were then input to the GML Camera Calibration Toolbox [67]. GML Camera Calibration Toolbox is a free, open source, and OpenCV based camera calibration software running on Windows XP Operating System. The toolbox merely used to conduct faster camera calibration process in practical experiment routine. Using the toolbox, extrinsic and intrinsic parameters of each camera can be extracted automatically. Figure 3.28 shows GML Camera Calibration Toolbox extracting intrinsic parameters of camera 1 after detecting all chessboard corner coordinates.

3.2.6 Lens Undistortion

So far, we have assumed that the lens performs ideal perspective transformation, as a pinhole camera does. However, in practical usage, no real lens is perfect. This is mainly for reasons of manufacturing; it is much easier to make a "spherical" lens than to make a more mathematically ideal "parabolic" lens. It is also difficult to mechanically align the lens and the image sensor exactly. In this thesis, we consider two major lens aberrations and how to model the distortion coefficients. The approach to model the lens distortion is mostly taken from a reference proposed by Brown [68].

Radial distortion arises as the shape of the lens and *tangential distortion* arise from the assembly process of the camera as whole. Figure 3.29(a) and Figure 3.29(b), which are taken from Bouguet's page of Matlab Camera Calibration Toolbox [69], illustrate radial and tangential distortions, respectively.

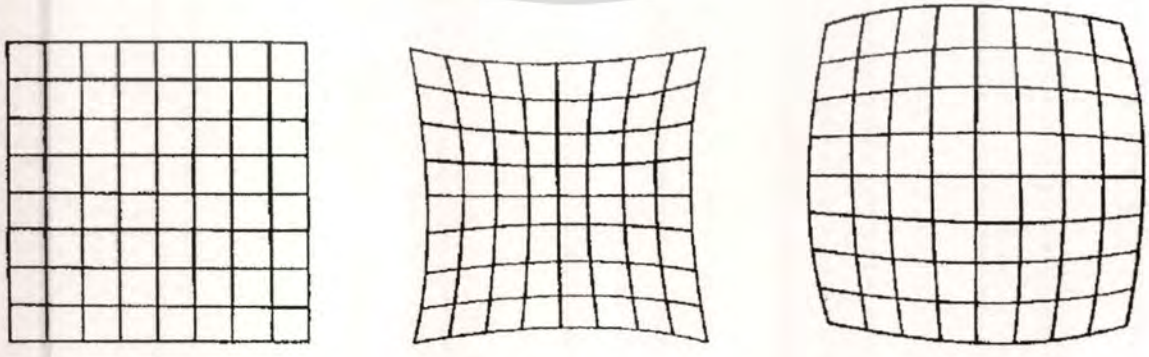


(a) (b)

Figure 3.29 Common lens aberration occurred on industrial camera:

(a) radial distortion; (b) tangential distortion [69]

Radial distortion can be noticed by observing the pixel near the edge of image. In Figure 3.29 (a), the arrows show where the points on an external rectangular grid are displaced in a radially distorted image. This distortion occurs because rays farther from the center of the lens are bent more (barrel) or less (pincushion) than that those closer from as illustrated in Figure 3.30. In barrel distortion (Figure 3.30 (c)), magnification decreases with distance from the center of image. On the contrary, in pincushion distortion (Figure 3.30 (b)), magnification increases with distance from the center of image. The distortion is 0 at the center of the image. In practice, this distortion is small and can be characterized by the first few terms of a Taylor series expansion. For common camera the first two terms k_1 and k_2 are usually used. For highly-distorted camera, such as fish eye lens, the third radial distortion term k_3 is included.



(a) (b) (c)

Figure 3.30 Radial distortion on 8 x 8 grids: (a) undistorted; (b) pincushion; (c) barrel

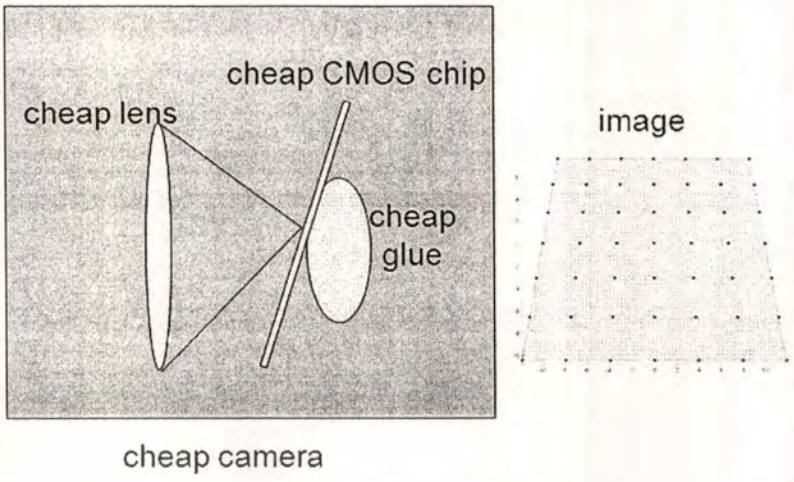


Figure 3.31 Tangential distortion on cheap industrial camera [70]

The second common distortion is tangential distortion. In Figure 3.29(b), the arrows show where points on an external rectangular grid are displaced in a tangentially distorted image. This distortion is due to manufacturing defects resulting from the lens not being exactly parallel to the imaging plane. This can happen when the imager is glued to the back of the camera, as shown in Figure 3.31 [70]. The tangential distortion can be modeled by including two tangential distortion parameters p_1 and p_2. To obtain radial and tangential distortions parameters, observe that the undistorted (perfect) projected 2D point location (x_{pix}, y_{pix}) of 3D point coordinates (X_S, Y_S, Z_S) has been defined in Eq. (3.26) and Eq. (3.27) as follows

$$\begin{aligned} x_{pix} &= k_x x_i + s y_i + x_0 \\ &= k_x f \frac{X_S}{Z_S} + sf \frac{Y_S}{Z_S} + x_0 \end{aligned} \tag{3.54}$$

$$\begin{aligned} y_{pix} &= k_y y_i + y_0 \\ &= k_y f \frac{Y_S}{Z_S} + y_0 \end{aligned} \tag{3.55}$$

where (x_0, y_0) is principal point, f is focal length, k_x is aspect ratio in x-axis, k_y is aspect ratio in y-axis, and s is skew parameter. Let (x_d, y_d) be the distorted location of (x_{pix}, y_{pix}), then we can define the distance r of distorted point location (x_d, y_d) with respect to principal point as follows:

$$r = \sqrt{(x_d - x_0)^2 + (y_d - y_0)^2} \tag{3.56}$$

where

$$\tilde{x} = x_d - x_0 \;\; ; \;\; \tilde{y} = y_d - y_0 \tag{3.57}$$

Correction factors of radial distortion in x and y axes can be formulated as:

$$c_x = \tilde{x}\left(k_1 r^2 + k_2 r^4 + k_3 r^6\right) \tag{3.58}$$

$$c_y = \tilde{y}\left(k_1 r^2 + k_2 r^4 + k_3 r^6\right) \tag{3.59}$$

Correction factors of tangential distortion in x and y axes can be formulated as:

$$d_x = 2p_1\tilde{x}\tilde{y} + p_2\left(r^2 + 2\tilde{x}^2\right) \tag{3.60}$$

$$d_y = p_1\left(r^2 + 2\tilde{y}^2\right) + 2p_2\tilde{x}\tilde{y} \tag{3.61}$$

The perfect point location (x_{pix}, y_{pix}) can be written as

$$x_{pix} = x_d + \left(c_x + d_x\right) \tag{3.62}$$

$$y_{pix} = y_d + \left(c_y + d_y\right) \tag{3.63}$$

Deriving both Eq. (3.62) and Eq. (3.63) yields

$$\begin{aligned} x_{pix} &= x_d + \tilde{x}\left(k_1 r^2 + k_2 r^4 + k_3 r^6\right) + 2p_1\tilde{x}\tilde{y} + p_2\left(r^2 + 2\tilde{x}^2\right) \\ y_{pix} &= y_d + \tilde{y}\left(k_1 r^2 + k_2 r^4 + k_3 r^6\right) + p_1\left(r^2 + 2\tilde{y}^2\right) + 2p_2\tilde{x}\tilde{y} \end{aligned} \tag{3.64}$$

$$\begin{aligned} x_{pix} &= x_d + \tilde{x}k_1 r^2 + \tilde{x}k_2 r^4 + \tilde{x}k_3 r^6 + 2p_1\tilde{x}\tilde{y} + p_2\left(r^2 + 2\tilde{x}^2\right) \\ y_{pix} &= y_d + \tilde{y}k_1 r^2 + \tilde{y}k_2 r^4 + \tilde{y}k_3 r^6 + p_1\left(r^2 + 2\tilde{y}^2\right) + 2p_2\tilde{x}\tilde{y} \end{aligned} \tag{3.65}$$

Rearranging the elements of equations yields

$$\begin{aligned} \left(x_{pix} - x_d\right) - \tilde{x}k_1 r^2 - \tilde{x}k_2 r^4 - \tilde{x}k_3 r^6 - 2p_1\tilde{x}\tilde{y} - p_2\left(r^2 + 2\tilde{x}^2\right) = 0 \\ \left(y_{pix} - y_d\right) - \tilde{y}k_1 r^2 - \tilde{y}k_2 r^4 - \tilde{y}k_3 r^6 - p_1\left(r^2 + 2\tilde{y}^2\right) - 2p_2\tilde{x}\tilde{y} = 0 \end{aligned} \tag{3.66}$$

Eq. (3.66) can be written in matrix form $AX = 0$ where X is 6 x 1 vector consists of distortion parameters. Lens distortion parameters can be estimated by observing a set of known undistorted (perfect) projected 2D points of 3D object scene and its corresponding distorted (real) 2D points using a chessboard pattern and GML Calibration Toolbox. If n such undistorted and distorted points are available, matrix A (size $2n$ x 6) and vector X can be written as follows:

$$\begin{bmatrix} (x_{pix}-x_d) & -\tilde{x}r^2 & -\tilde{x}r^4 & -\tilde{x}r^6 & -2\tilde{x}\tilde{y} & -\left(r^2+2\tilde{x}^2\right) \\ (y_{pix}-y_d) & -\tilde{y}r^2 & -\tilde{y}r^4 & -\tilde{y}r^6 & -\left(r^2+2\tilde{y}^2\right) & -2\tilde{x}\tilde{y} \\ & & & \vdots & & \end{bmatrix} \begin{bmatrix} 1 \\ k_1 \\ k_2 \\ k_3 \\ p_1 \\ p_2 \end{bmatrix} = 0 \quad (3.67)$$

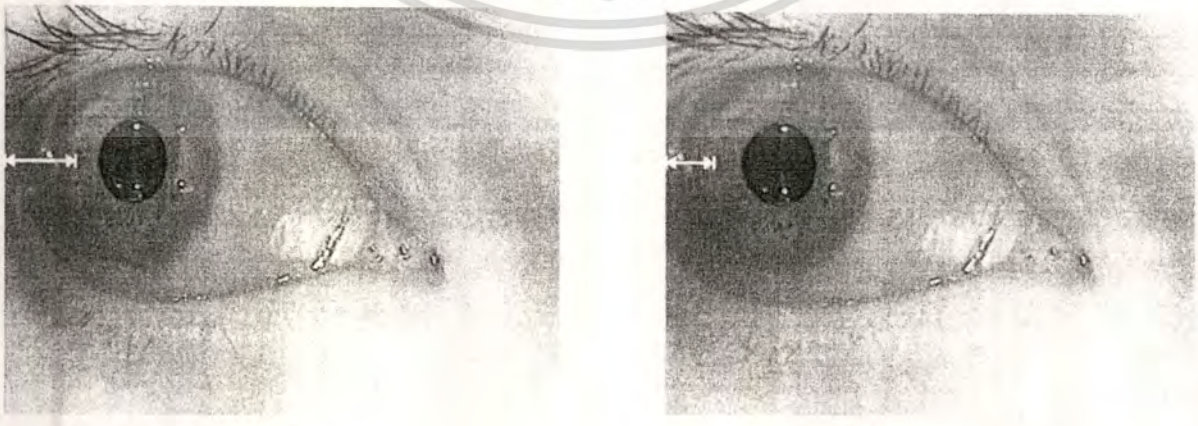
$$AX = 0 \quad (3.68)$$

In order to solve matrix X, performs an SVD [64] of matrix A (size m x n) to derive

$$A_{mn} = U_{mm} D_{mn} V_{nn}^{T} \quad (3.69)$$

where matrix U (size m x m) is left orthogonal matrix, matrix D (size m x n) is diagonal matrix, and matrix V (size n x n) right orthogonal matrix. The last column of matrix V is the solution for vector X. After radial and tangential distortions parameters are obtained, the lens undistortion is done by computing the correct location of the point (x_{pix}, y_{pix}) using Eq. (3.62) and Eq. (3.63).

Figure 3.32(a) and Figure 3.32(b) show the image captured from distorted lens and the result of lens undistortion algorithm, respectively. Both images were captured from camera 2. The most noticeable change is a particular iris striation location measured from the left edge of the image as shown by the white arrow. The distance between the iris striation and the left edge of the image was simply shorter after lens undistortion algorithm (Figure 3.32(b)). The iris striation position was improved by adding correction values with respect to principal point as explained in Eq. (3.62) and Eq. (3.63). The lens distortion algorithm is very useful to improve accuracy of 3D coordinates extraction algorithm explained in the next section.



(a) (b)

Figure 3.32 Captured image from camera 2: (a) distorted lens and (b) undistorted lens

3.2.7 3D Coordinates Extraction

Computing 3D information of a scene from optical measurement is very important for a wide variety of measurement. Eye movements tracking involves gauging small part of human eye that can lead into inaccuracy problem as result of geometric distortion due to point projection from 3D object to image sensor. Several approaches have been conducted to achieve precise measurement of small object by implementing a widely-used 3D coordinates extraction algorithm, Direct Linear Transformation (DLT), as explained in [71-74]. DLT algorithm is originally proposed by Abdel-Aziz and Karara [75]. In this thesis, DLT algorithm was utilized to obtain accurate eye movements tracking result. Furthermore, the accuracy of DLT algorithm can be enhanced by involving lens undistortion algorithm explained in the previous section.

The DLT algorithm requires two or more cameras to transform image coordinates into object space coordinates. It constructs relationship between digitized coordinates from the two or more camera views and the corresponding coordinates in 3D object space. In this thesis, the 3D coordinates of pupil center and iris striation center were simultaneously extracted from their corresponding 2D coordinates on two pair of images. An illustration of 3D coordinates extraction of pupil center is shown in Figure 3.33. Similar approach is used to extract 3D coordinates of iris striation center.

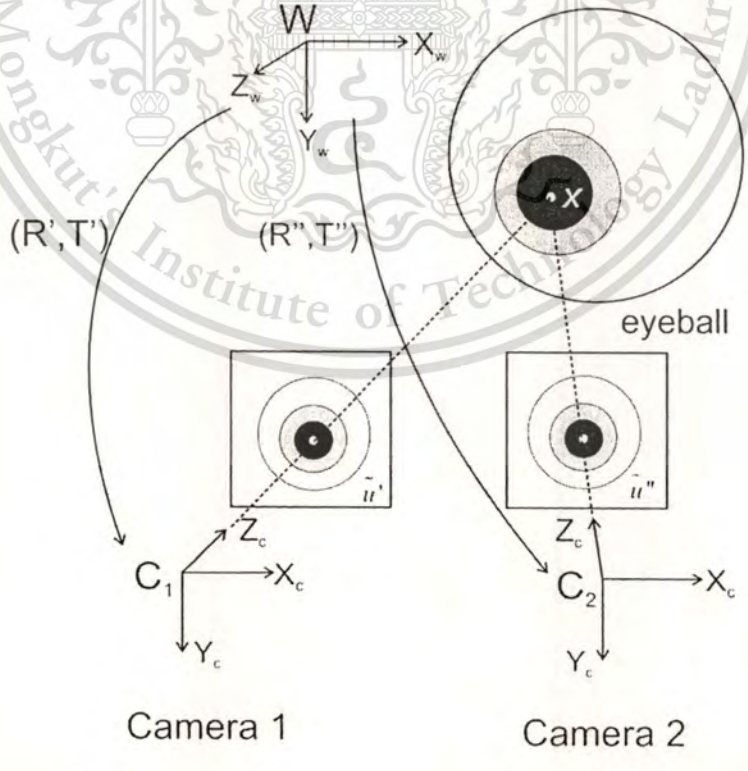


Figure 3.33 Illustration of 3D coordinates extraction of pupil center.
Similar approach is used to obtain 3D coordinates of iris striation center

In Figure 3.33, the origins of camera coordinate system C_1 and C_2 do not coincide with the origin of world coordinate system W . Therefore, the second assumption (Eq. (3.39)) is used to obtain intrinsic and extrinsic parameter of camera 1 and camera 2. Matrix R' and vector T' denote rotational matrix and translational vector of camera 1, respectively. Matrix R'' and vector T'' denote rotational matrix and translational vector of camera 2, respectively. $\tilde{u}'$ and $\tilde{u}''$ are coordinates of pupil center on images captured from camera 1 and camera 2, respectively. X is the extracted 3D coordinates of pupil center.

In the previous section, we have discussed about estimating projection matrix M from a camera (Eq. (3.49)). Let the projection matrix for camera 1 and camera 2 are denoted by M' and M'', respectively. The projection matrix M' for camera 1 is divided into three row vectors $m'^{T}_1, m'^{T}_2, m'^{T}_3$ while the projection matrix M'' for camera 2 is divided into three row vectors $m''^{T}_1, m''^{T}_2, m''^{T}_3$. Recalling Eq. (3.43) for camera 1, the perspective transformation equation is formed as follows

$$\tilde{u}' = \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} = \begin{bmatrix} m'_{11} & m'_{12} & m'_{13} & m'_{14} \\ m'_{21} & m'_{22} & m'_{23} & m'_{24} \\ m'_{31} & m'_{32} & m'_{33} & m'_{34} \end{bmatrix} \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = M'X = \begin{bmatrix} m'^{T}_1 \\ m'^{T}_2 \\ m'^{T}_3 \end{bmatrix} X \tag{3.70}$$

Similarly for camera 2, the perspective transformation equation is defined as follows

$$\tilde{u}'' = \begin{bmatrix} u'' \\ v'' \\ w'' \end{bmatrix} = \begin{bmatrix} m''_{11} & m''_{12} & m''_{13} & m''_{14} \\ m''_{21} & m''_{22} & m''_{23} & m''_{24} \\ m''_{31} & m''_{32} & m''_{33} & m''_{34} \end{bmatrix} \begin{bmatrix} X'' \\ Y'' \\ Z'' \\ 1 \end{bmatrix} = M''X = \begin{bmatrix} m''^{T}_1 \\ m''^{T}_2 \\ m''^{T}_3 \end{bmatrix} X \tag{3.71}$$

where $\tilde{u}'$ and $\tilde{u}''$ denote the homogeneous image coordinates of point on camera 1 and camera 2, respectively. X denotes the corresponding homogeneous 3D coordinates of point on scene. In order to eliminate the unknown scale factor, consider the ratio between three rows in the projection matrix M' and M''

$$\begin{aligned} u':v':w' &= m'^{T}_1 X : m'^{T}_2 X : m'^{T}_3 X \\ u'':v'':w'' &= m''^{T}_1 X : m''^{T}_2 X : m''^{T}_3 X \end{aligned} \tag{3.72}$$

Thus three equations hold for both camera 1 and camera 2

$$\begin{aligned} u'm'^{T}_2 X &= v'm'^{T}_1 X & u''m''^{T}_2 X &= v''m''^{T}_1 X \\ u'm'^{T}_3 X &= w'm'^{T}_1 X & u''m''^{T}_3 X &= w''m''^{T}_1 X \\ v'm'^{T}_3 X &= w'm'^{T}_2 X & v''m''^{T}_3 X &= w''m''^{T}_2 X \end{aligned} \tag{3.73}$$

Eq. (3.73) can be written in a matrix form as follows

$$\begin{bmatrix} u'm'^{T}_{2} - v'm'^{T}_{1} \\ u'm'^{T}_{3} - w'm'^{T}_{1} \\ v'm'^{T}_{3} - w'm'^{T}_{2} \end{bmatrix} X = 0 \qquad \begin{bmatrix} u''m''^{T}_{2} - v''m'^{T}_{1} \\ u''m''^{T}_{3} - w''m''^{T}_{1} \\ v''m''^{T}_{3} - w''m''^{T}_{2} \end{bmatrix} X = 0 \tag{3.74}$$

Multiplying the first row in camera 1 matrix by w', the first row in camera 2 matrix by w'', the second row of camera 1 matrix by $-v'$,and the second row of camera 2 matrix by $-v''$, we get

$$\begin{aligned} \left(u'w'm'^{T}_{2} - v'w'm'^{T}_{1} - u'v'm'^{T}_{3} + v'w'm'^{T}_{1}\right)X &= \left(u'w'm'^{T}_{2} - u'v'm'^{T}_{3}\right)X = 0 \\ \left(u''w''m''^{T}_{2} - v''w''m''^{T}_{1} - u''v''m''^{T}_{3} + v''w''m''^{T}_{1}\right)X &= \left(u''w''m''^{T}_{2} - u''v''m''^{T}_{3}\right)X = 0 \end{aligned} \tag{3.75}$$

Extracting the equation corresponding to the third row of the matrix in Eq. (3.74), we get

$$\left(-w'm'^{T}_{2} + v'm'^{T}_{3}\right)X = 0 \qquad \left(-w''m''^{T}_{2} + v''m''^{T}_{3}\right)X = 0 \tag{3.76}$$

By observing Eq. (3.75) and Eq. (3.76), it is known that the first row and the third row in Eq. (3.74) are linearly dependent

$$\begin{aligned} \left(u'w'm'^{T}_{2} - u'v'm'^{T}_{3}\right)X &= -u'\left(-w'm'^{T}_{2} + v'm'^{T}_{3}\right)X = 0 \\ \left(u''w''m''^{T}_{2} - u''v''m''^{T}_{3}\right)X &= -u''\left(-w''m''^{T}_{2} + v''m''^{T}_{3}\right)X = 0 \end{aligned} \tag{3.77}$$

Since only two rows are linearly independent, we use the second and the third rows

$$\begin{aligned} \left(u'm'^{T}_{3} - w'm'^{T}_{1}\right)X = 0 \qquad \left(u''m''^{T}_{3} - w''m''^{T}_{1}\right)X = 0 \\ \left(v'm'^{T}_{3} - w'm'^{T}_{2}\right)X = 0 \qquad \left(v''m''^{T}_{3} - w''m''^{T}_{2}\right)X = 0 \end{aligned} \tag{3.78}$$

Eq. (3.78) can be written in matrix form

$$\begin{bmatrix} u'm'^{T}_{3} - w'm'^{T}_{1} \\ v'm'^{T}_{3} - w'm'^{T}_{2} \\ u''m''^{T}_{3} - w''m''^{T}_{1} \\ v''m''^{T}_{3} - w''m''^{T}_{2} \end{bmatrix} X = AX = 0 \tag{3.79}$$

The matrix A has size 4 x 4 and X is a 4 x 1 vector. We are interested in a nontrivial solution $(X \neq 0)$ of Eq. (3.79), and therefore consider the case $\det(A) = 0$. SVD [64] can be performed to derive

$$A_{mn} = U_{mm} D_{mn} V_{nn}^{T} \tag{3.80}$$

where matrix U (size m x m) is left orthogonal matrix, matrix D (size m x n) is diagonal matrix, and matrix V (size n x n) right orthogonal matrix. The last column of matrix V is the solution for vector X.

3.2.8 Angular Measurement of 3D Eye Movements

3.2.8.1 Convention of Measurement

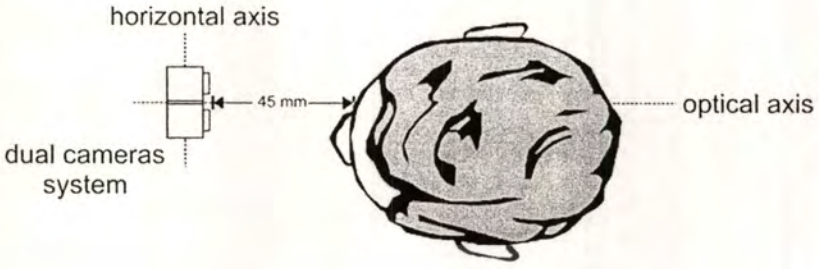


Figure 3.34 Configuration of dual cameras system with respect to eye

Eye movements in 3D scene consist of translation as well as rotations. Translation of the eye, which can be due to a translation of the eye in the orbit [76], as well as to head movement, will not be dealt in this thesis, and the terms *eye movements* or *eye positions* will refer only to the rotational component. Therefore, several assumptions are defined:

1. The eye exhibits ideal *ball* and *socket* behavior so that all eye movements are pure rotations around the center of the eye, with no translation of this center.
2. The optical axis of the eye (given approximately by the axis through the center of the pupil and the center of rotation of the eye) coincides with the visual axis, i.e., the line of sight.
3. The dual cameras system is placed directly in the front of the eye approximately 45 mm measured from the center of pupil to the front side of cameras. The placement is configured such that the horizontal axis of the cameras is perpendicular with the optical axis of the eye as illustrated in Figure 3.34.

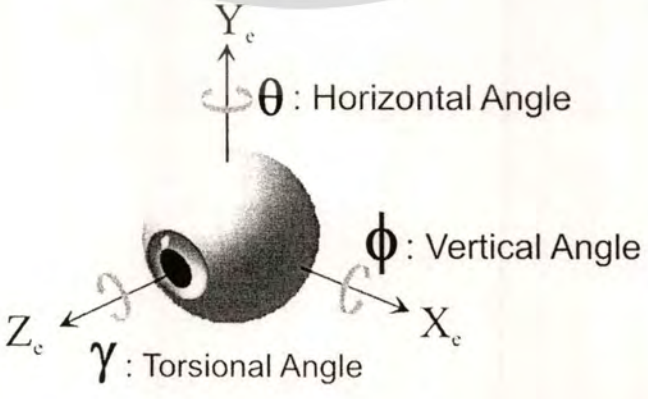


Figure 3.35 Coordinate system convention of 3D eye movements

The coordinate convention for 3D eye movements follows the right-handed coordinate system [8] as shown in Figure 3.35. The eye coordinate system (X_e, Y_e, Z_e) is defined such that Z_e-axis is perpendicular to the image plane and passes through the center of the pupil, Y_e-axis is parallel to interaural axis of human body, and X_e-axis parallel to the earth vertical axis. As the dual cameras system is installed in the wearable goggle, the center of the head-fixed coordinate system coincides with the center of the eye coordinate system. The rotation angles are termed as θ (horizontal rotation / pan), ϕ (vertical rotation / tilt), and γ (torsional rotation / yaw).

Before vertical, horizontal, and torsional angular eye positions are determined, physical parameters of the eye must be measured. The parameters consist of eye diameter and iris diameter. Both parameters are used to compute millimetres-to-pixels conversion factor α. A professional vernier caliper (Mitutoyo, Kawaski, Japan) with accuracy ±0.02mm was used to measure the physical parameters. Two pieces of band-aid were attached to the carbide-tipped of the caliper to prevent harming eye. Eyeball diameter of the patient was determined by carefully measure the distance of outer part (*lateral canthi*) and inner part (*medial canthi*) of eyelid aperture (*palpebral aperture*), as shown in Figure 3.36 (a). Iris diameter was measured by placing the caliper on the front of eye, as shown in Figure 3.36 (b).

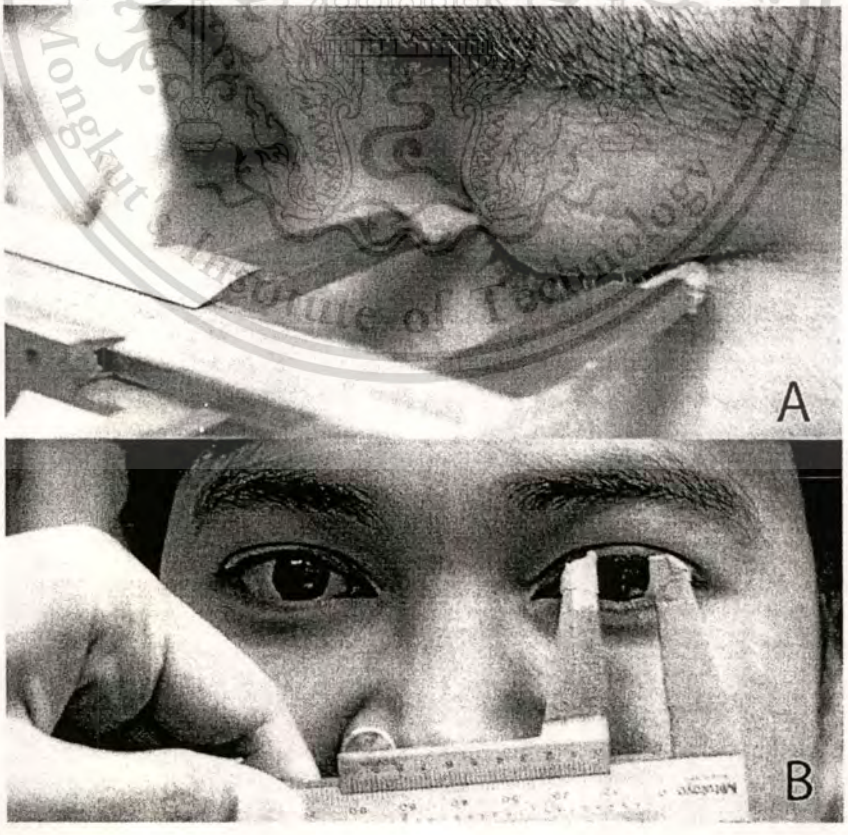
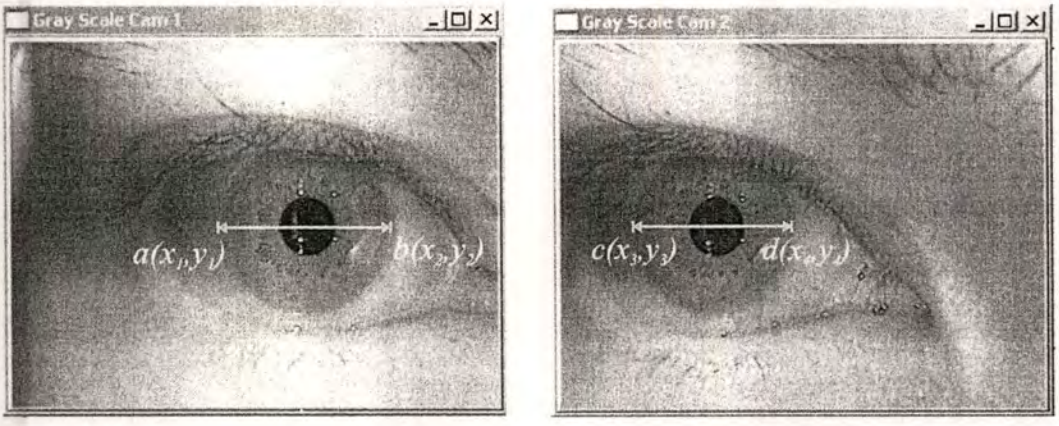


Figure 3.36 Measurement of physical parameters: (a) eye diameter; (b) iris diameter



(a) (b)

Figure 3.37 Measurement of iris diameter in pixels unit
using captured images from (a) camera 1 and (b) camera 2

In order to obtain the radius of the eyeball, we can simply divide the diameter of eyeball by two:

$$R_{eye} = \frac{D_{eye}}{2} \tag{3.81}$$

where R_{eye} and D_{eye} denote the radius of the eyeball and the diameter of the eyeball in millimetres unit, respectively.

Next, the iris diameter in pixels unit was measured. The measurement procedure is illustrated in Figure 3.37. Initially, single eye image from camera 1 was captured and two diametrically opposed points on the iris edge, $a(x_1, y_1)$ and $b(x_2, y_2)$, were marked. In order to make sure that the iris diameter was correctly measured, the line pointer was configured such that $y_1 = y_2$. Similar procedure was applied for image captured from camera 2, resulting $c(x_3, y_3)$ and $d(x_4, y_4)$ where $y_3 = y_4$.

Next, 3D coordinates of the left point $A(X_1, Y_1, Z_1)$ was obtained from $a(x_1, y_1)$ and $c(x_3, y_3)$ using DLT algorithm. Similar computation was applied to the right points. The 3D coordinates of the right point $B(X_2, Y_2, Z_2)$ was obtained from $b(x_2, y_2)$ and $d(x_4, y_4)$. Since the iris was modeled as a plane section as explained in [77], the iris diameter in pixels unit d_{iris} can be obtained by measuring the absolut value of the horizontal distance as follows

$$d_{iris} = |X_1 - X_2| \tag{3.82}$$

The millimetres-to-pixels conversion factor α can be defined as follows

$$\alpha = \frac{d_{iris}}{D_{iris}} \tag{3.83}$$

where d_{iris} and D_{iris} denote the iris diameter in pixels unit and the iris diameter in millimetres unit, respectively. Finally, the radius of the eyeball in pixels unit r_{eye} can be determined using this formula:

$$r_{eye} = \alpha R_{eye} \tag{3.84}$$

where α is millimetres-to-pixels conversion factor, R_{eye} is the radius of the eyeball in millimetres unit.

3.2.8.2 Determining Vertical and Horizontal Angular Positions

The DLT algorithm yields 3D coordinates of pupil center $P'(X, Y, Z)$. The reference position of pupil center is determined when the eye is seeing at the line-of-sight so that the optical axis coincides with Z_e-axis. This alignment provides initial coordinates $P(X_0, Y_0, Z_0)$ which is used to measure the angular position of the eye. N denotes the center of the eye coordinate system as illustrated in Figure 3.38. Based on general right-handed trigonometric functions, the vertical angular position (ϕ) measured in radians unit is determined as follows

$$\phi = -\arcsin\left(\frac{\Delta Y}{r_{eye}}\right) \tag{3.85}$$

where $\Delta Y = |Y - Y_0|$ and r_{eye} is the radius of the eyeball obtained from Eq. (3. 84). The vertical angular position in degrees unit (ϕ_{deg}) can be determined as follows

$$\phi_{deg} = \phi * 57.2957795 \tag{3.86}$$

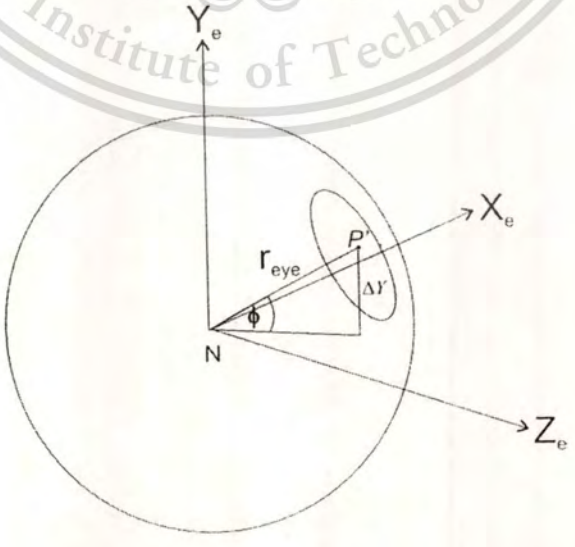


Figure 3.38 Measurement of vertical angular position (ϕ)

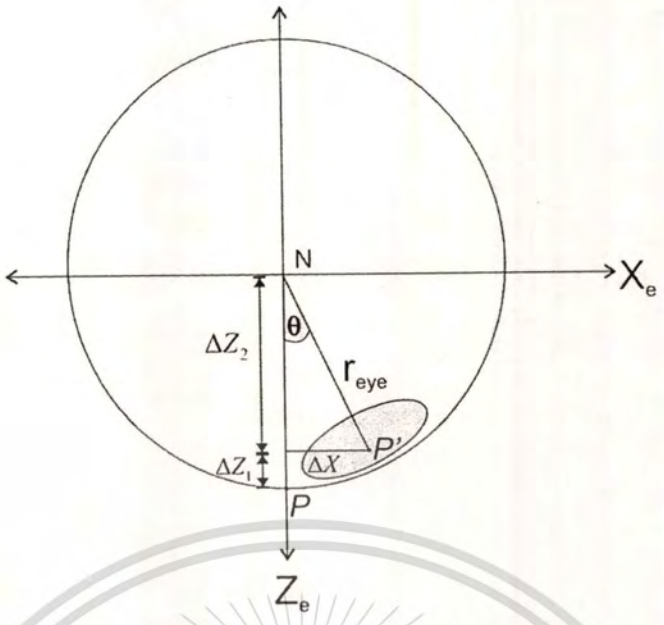


Figure 3.39 Measurement of horizontal angular position (θ)

The measurement of horizontal angular position (θ) is illustrated in Figure 3.39. Since the Z position of the pupil center is measured with respect to the center of world coordinate, we shall determine the Z position of pupil center with respect to the eye coordinate system. Thus, the radius of the eyeball r_{eye} can be written in term of Z position as follows

$$r_{eye} = \Delta Z_1 + \Delta Z_2 \tag{3.87}$$

where $\Delta Z_1 = |Z - Z_0|$ and ΔZ_2 is the distance of current pupil center position P' from the center of eye measured on Z_e-axis. If the pupil center is located at reference position, then

$$r_{eye} = \Delta Z_2 \tag{3.88}$$

Based on general right-handed trigonometric functions, the horizontal angular position (θ) measured in radians unit is determined as:

$$\theta = \arctan\left(\frac{\Delta X}{\Delta Z_2}\right) \tag{3.89}$$

where $\Delta X = |X - X_0|$ and $\Delta Z_2 = r_{eye} - \Delta Z_1$. The vertical angular position in degrees unit (θ_{deg}) can be determined as follows

$$\theta_{deg} = \theta * 57.2957795 \tag{3.90}$$

3.2.8.3 Determining Torsional Angular Position

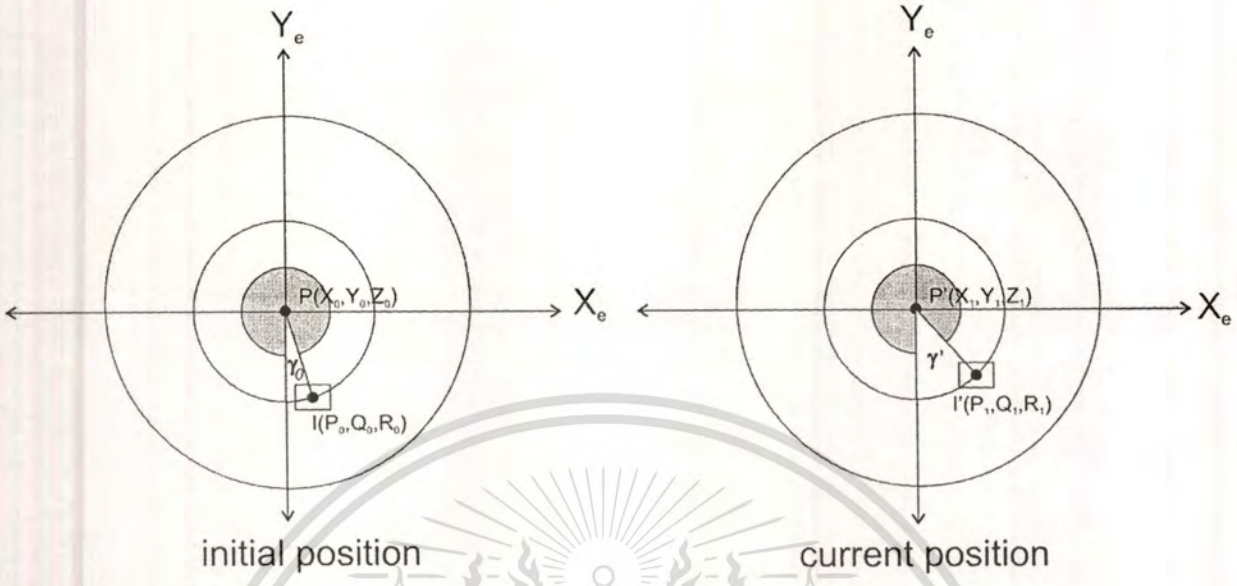


Figure 3.40 Measurement of torsional angular position (γ)

As explained in the previous section, the iris is modeled as a plain section. In order to measure the torsional angular position accurately, the 3D coordinates resulted from DLT algorithm are utilized. However, only two components of 3D coordinates will be used to measure the torsional position, X and Y components.

Figure 3.40 illustrates the method to measure torsional angular position (γ). $P(X_0, Y_0, Z_0)$ and $P'(X_1, Y_1, Z_1)$ denote the initial and current 3D coordinates of the pupil center, respectively. $I(P_0, Q_0, R_0)$ and $I'(P_1, Q_1, R_1)$ denote the initial and current 3D coordinates of the iris striation, respectively. γ_0 and γ' denote the initial and current torsional angles, respectively. Measuring the torsional movement γ in radians unit is done by subtracting γ_0 from γ'

$$\begin{aligned}\gamma &= \gamma' - \gamma_0 \\ &= \arctan\left(\frac{P_1 - X_1}{Q_1 - Y_1}\right) - \arctan\left(\frac{P_0 - X_0}{Q_0 - Y_0}\right)\end{aligned} \tag{3.91}$$

The torsional angular position in degrees unit (γ_{deg}) can be determined as follows

$$\gamma_{deg} = \gamma * 57.2957795 \tag{3.92}$$

3.2.9 3D Visualization of Eye Movements

The 3D visualization was mainly developed using OpenGL library. The OpenGL library is a software interface to graphics hardware. GL stands for Graphics Library. It allows us to create interactive programs that produce color images of moving three-dimensional object. Using OpenGL library, we can control computer-graphics technology to produce either realistic images or ones that depart from reality in imaginative ways.

The 3D eyeball was created to provide an easier tool for medical expert to analyze voluntary and spontaneous eye movements. The basic structure of the 3D eyeball is simple where its tasks are the initialization of certain states that control how OpenGL renders and the specification of objects to be rendered. Rendering is the process by which a computer creates images from models. These models, or objects, are constructed from geometric primitives. An example of a core OpenGL program for rendering a 3D eyeball is shown in Figure 3.41. The program consists of four parts: "Drawing axes", "Rotation", "Texture mapping" and "Drawing sphere".

```
/* Drawing axes */
float XUP[3] = {1,0,0},
      YUP[3] = {0,1,0},
      ZUP[3] = {0,0,1},
      ORG[3] = {0,0,0};

glLineWidth (2.0);
  glBegin (GL_LINES);
    glVertex3fv(ORG); glVertex3fv(XUP);
    glVertex3fv(ORG); glVertex3fv(YUP);
    glVertex3fv(ORG); glVertex3fv(ZUP);
  glEnd ();

glRasterPos3f (1.1,0.0, 0.0);
glPrint("X");
glRasterPos3f (0.0, 1.1, 0.0);
glPrint("Y");
glRasterPos3f (0.0, 0.0, 1.1);
glPrint("Z");


/* Rotation */
glRotatef(theta, 0.0, 1.0, 0.0);
glRotatef(phi, 1.0, 0.0, 0.0);
glRotatef(gamma, 0.0, 0.0, 1.0);

/* Texture mapping */
glBindTexture(GL_TEXTURE_2D, textureName);

/* Drawing sphere */
q = gluNewQuadric();
gluSphere(q,0.4f,32,16);
gluDeleteQuadric(q);
```

Figure 3.41 Example of OpenGL program for rendering a 3D eyeball

In the "Drawing axes" part, the 3D points coordinates of the origin (ORG), the end of positive *X*-axis (XUP), the end of positive *Y*-axis (YUP), and the end of positive *Z*-axis (ZUP) are defined. The first line is OpenGL command glLineWidth() to specify the width of the axes lines. The next eight commands, which are bracketed by glBegin() and glEnd(), define the object *X*, *Y*, and *Z* axes to be drawn. The end points of the line are defined by the glVertex3fv() command, taking parameter from previously defined axes coordinates. The next six commands render the "X", "Y", and "Z" characters. The glRasterPos3f() command is used to define the 3D position in OpenGL window coordinate system, which is called raster position. The glPrint() command is used to write a character on OpenGL window. In the "Rotation" part, the glRotate() command changes the object's orientation. In OpenGL commands, 3D vertices are treated internally as 3D homogeneous vertices comprising four coordinates. Vertex transformations such as rotations can be represented by applying an appropriate 4 x 4 matrix to the coordinates representing the vertex in homogeneous coordinate. The glRotatef(TYPE angle, TYPE x, TYPE y, TYPE z) command multiplies the current matrix by a matrix that rotates an object in a counterclockwise direction about the ray from the origin through the point (*x*, *y*, *z*). The angle parameter specifies the angle of rotation in degrees. The command glRotatef(theta, 0.0, 1.0, 0.0) will rotate the 3D eyeball horizontally in *theta* degrees about the *Y*-axis. The command glRotatef(phi, 1.0, 0.0, 0.0) will rotate the 3D eyeball vertically in *phi* degrees about the *X*-axis. The command glRotatef(gamma, 0.0, 0.0, 1.0) will rotate the 3D eyeball torsionally in *gamma* degrees about the *Z*-axis. The values of *theta*, *phi*, and *gamma* are taken from real-time results of 3D eye movements tracking.

In the "Texture mapping" part, the glBindTexture(GL_TEXTURE_2D, textureName) command binds the texture specified by "textureName" to the GL_TEXTURE_2D target. In our 3D eyeball, the texture consists of the pupil, iris, and the outer part of the eye (*sclera*). In the "Drawing sphere" part, the gluNewQuadric() command initializes quadric object to be rendered as a sphere by gluSphere() command. Quadric object is used to render and model the surfaces of spheres and cylinders and to tessellate disks (circles) and partial disks (arcs). The gluSphere() command draws a sphere of the given radius centered around the origin. The first argument "q" denote the quadric object, the second argument "0.4f" specifies the radius of sphere, the third argument "32" specifies the number of subdivisions around the *Z*-axis (similar to lines of longitude), and the fourth argument "16" specifies the number of subdivisions along the *Z*-axis (similar to lines of latitude). The gluDeleteQuadric() command is used to delete the quadric object after rendered in order to avoid memory leak.

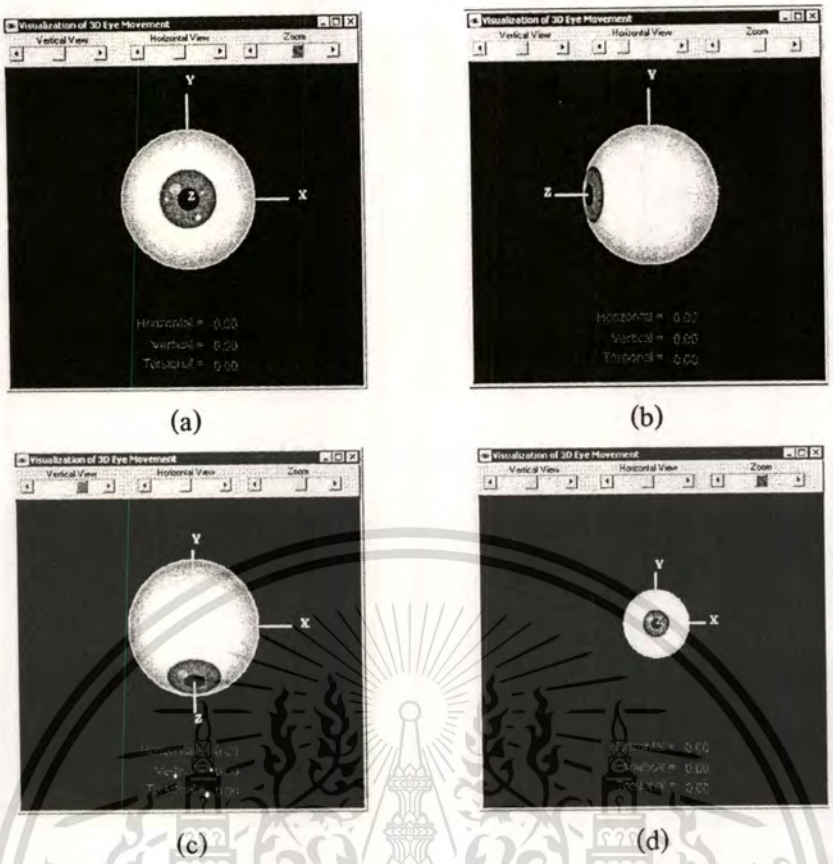
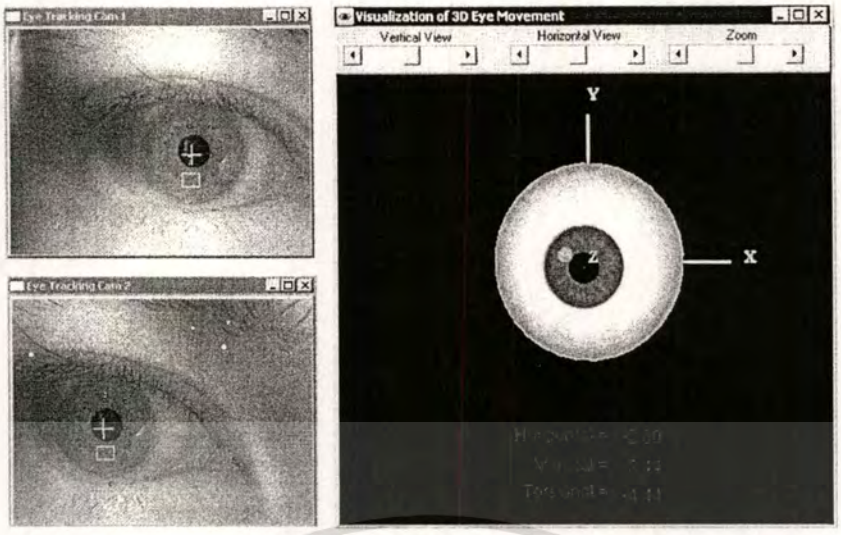
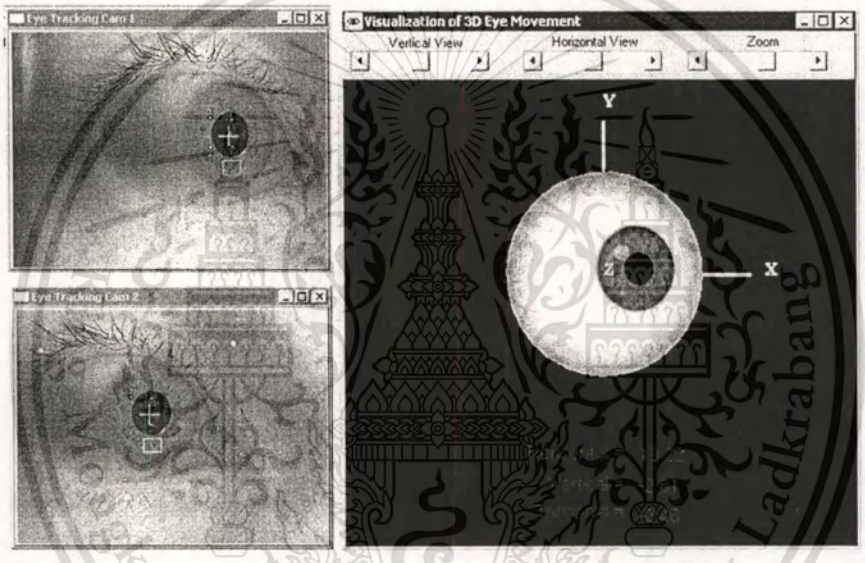


Figure 3.42 Several point-of-views of the 3D eyeball:
(a) default position; (b) rotated horizontally; (c) rotated vertically; (c) zoomed out

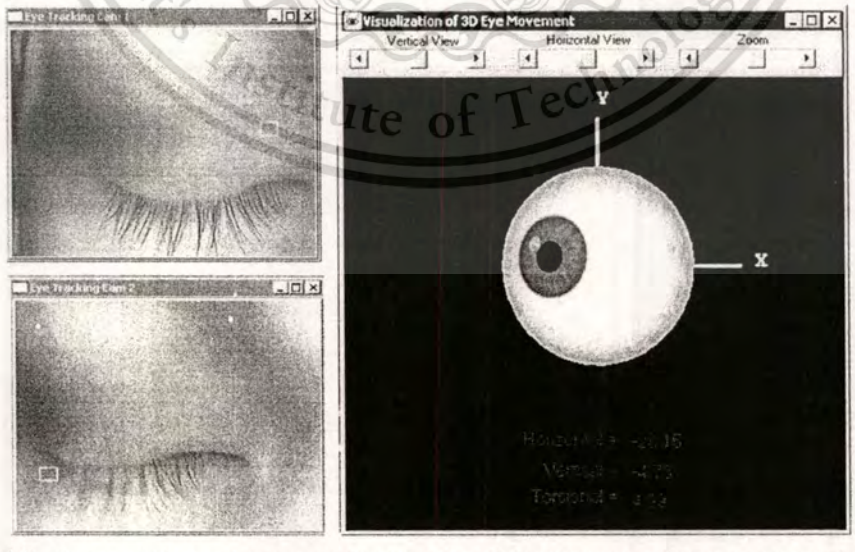
The 3D eyeball can be viewed interactively from several point-of-views by changing its orientation and size. The glRotatef(TYPE angle, TYPE x, TYPE y, TYPE z) command is utilized to change the orientation based on rotation. The glScalef(TYPE x, TYPE y, TYPE z) command produces a general scaling along the *X*, *Y*, and *Z* axes. The three arguments of glScalef() command indicate the desired scale factors along each of the three axes. Figure 3.42 shows several point-of-views of the 3D eyeball. Figure 3.42 (a) shows the default 3D eyeball position when it is launched. The horizontal and vertical rotation-based point-of-views are shown in Figure 3.42 (b) and Figure 3.42 (c), respectively. The scaling of the 3D eyeball about *Z*-axis produces zoom in and zoom out effects (Figure 3.42 (d)). The zoom in effect enlarges the size of the eyeball while the zoom out effect shrinks the size of the eyeball. Figure 3.43 shows several representative snapshots of the real eye and the OpenGL 3D eyeball. Figure 3.43 (a) shows a snapshot when the eye gazed around optical axis (θ = -2.80; ϕ = 3.44; γ =-4.44). Figure 3.43 (b) shows a snapshot when the eye gazed to the horizontal eccentric position (θ = 19.22; ϕ = -2.81; γ =-2.46). The tracking and visualization failed when the patient closed his eye as the algorithm was unable to detect the pupil and iris striation. This condition is shown in Figure 3.43 (c).



(a) Frame 87: the eye gazed around the optical axis



(b) Frame 96: the eye gazed to eccentric position



(c) Frame 143: tracking and visualization failed because the eye was occluded

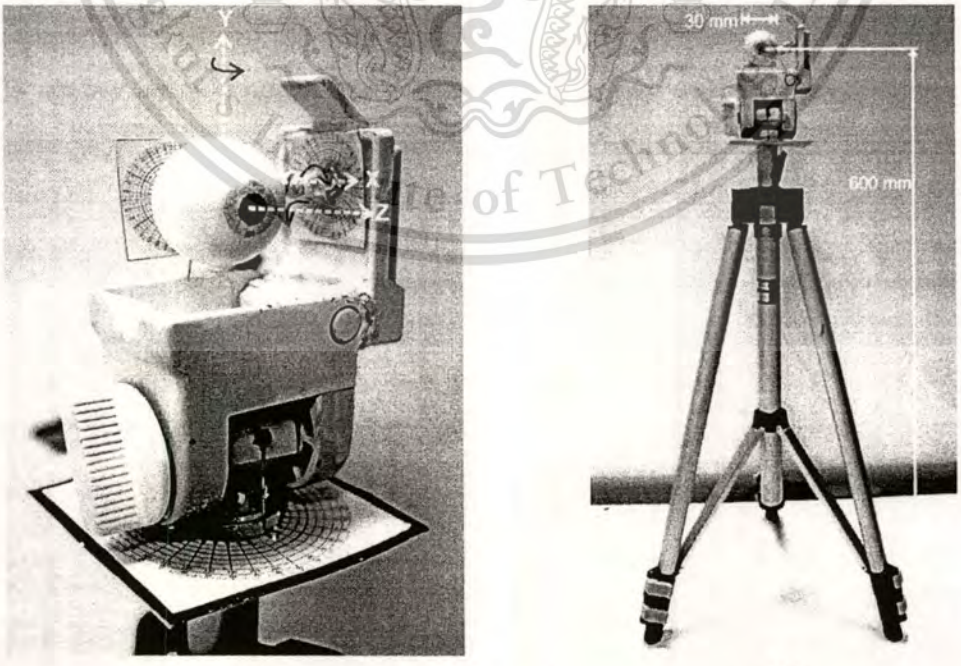
Figure 3.43 Representative snapshots from real-time tracking and visualization

3.3 Experimental Methods

3.3.1 System Validation

Validation of the eye movements tracking and visualization system was performed by initially performing camera calibration to obtain camera and distortion parameters as explained in the previous section. A tiny 5 x 6 chessboard pattern was positioned approximately 45 mm facing the dual cameras system. Eight different positions were then simultaneously captured as shown in Figure 3.27. In order to practically extract intrinsic, extrinsic, and lens distortion parameters, GML Camera Calibration Toolbox was used.

We constructed a calibration device consisting of an artificial eyeball fixed to a gimbal system as used in [5]. The gimbal system was made by modifying a camera tripod. Horizontal and torsional axes were added to transform the tripod into a Fick-gimbal system, enabling complete 3D eye movements. Three printed protractors were added in horizontal, vertical, and torsional axes. The resolution of the protractors was 0.25°. The sizes of the protractors were 65 x 65 mm, 30 x 30 mm, 30 x 30 mm for horizontal, vertical, and torsional axes respectively. Three needles were installed pointing to the protractors to ensure the precision of gimbal system in measurement process. Figure 3.44 shows the constructed calibration device. Positive value is shown by the direction of the arrows in Figure 3.44 (a). The distance between centre of artificial eyeball and the ground was 600 mm as shown in Figure 3.44 (b).



(a) (b)

Figure 3.44 Artificial eye mounted on Fick-gimbal system

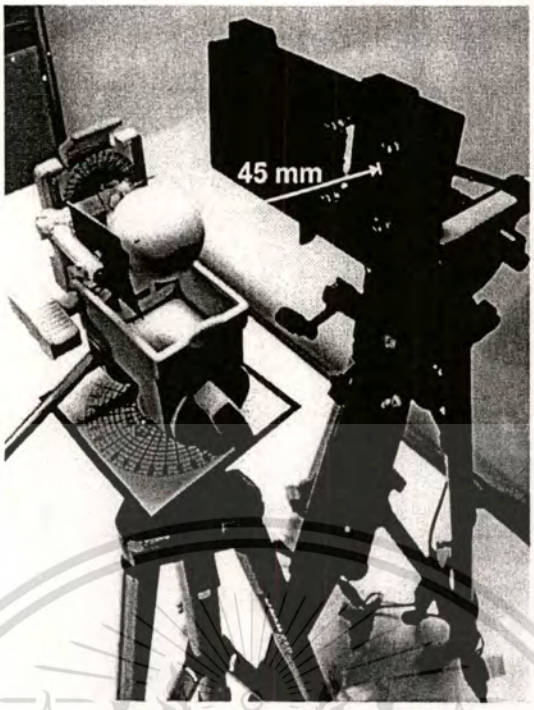


Figure 3.45 The alignment of calibration device and dual cameras system

The artificial eyeball consisted of a white plastic sphere having dimensions similar to that of human eyeball (radius 15 mm). A 4 mm black circle was drawn in the centre of the sphere acting as a "pupil". A blue circle with 12 mm diameter was drawn around the pupil acting as an "iris". Three short straight lines simulating a human iris striation were scribed at the iris circle. The artificial eyeball was installed such that the centre of sphere coincided with the center of gimbal system. Rotation of the artificial eyeball was possible about horizontal, vertical and torsional of the gimbal with resolution 0.25°. The calibration device and the dual cameras system were then positioned as shown in Figure 3.45. They were aligned such that the line of sight of the artificial eye was orthogonal to the horizontal axis of the cameras, and the pupil was approximately centered in the image plane when the artificial eye was in the reference position. The distance between the dual cameras system and the artificial eye was approximately 45 mm. The gimbal system was adjusted such that horizontal, vertical, and torsional movements of the gimbal corresponded to horizontal, vertical, and torsional movements of the artificial eye in the image plane.

Five settings of torsional angular position were made with the artificial eye rotated to angles of horizontal and vertical excursion ranging between ±20° at steps of 10°, resulting 125 unique positions. Each unique position consisted of Horizontal-Vertical-Torsional (HVT) values. At each unique position, 5 images were acquired using dual cameras system to ensure the accuracy of measurement.

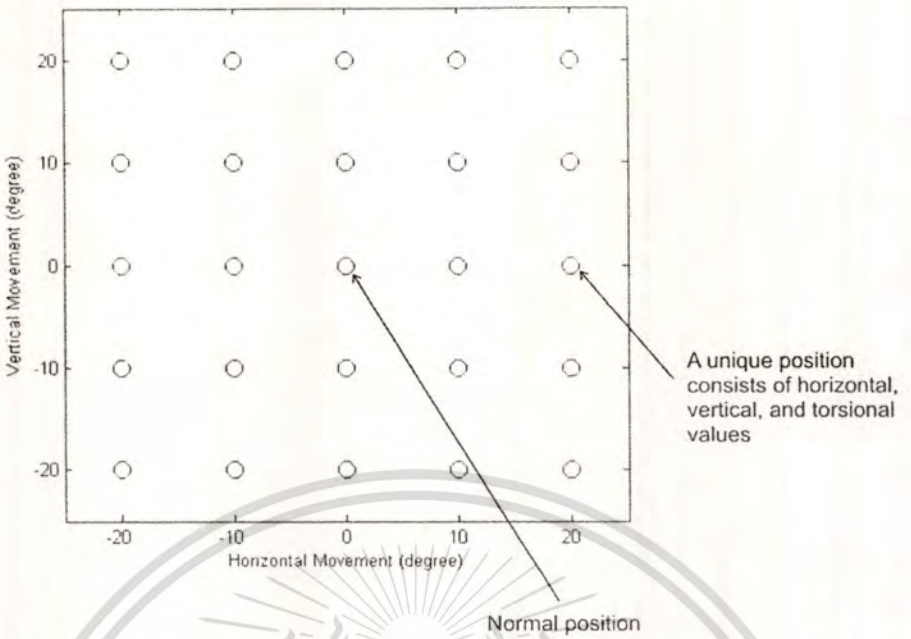


Figure 3.46 Calibration target of artificial eyeball

The mean of horizontal, vertical, and torsional values were then computed from 5 images and defined as the measured HVT value for each unique position. A normal position is defined when the artificial eyeball was aligned at the line-of-sight, while an eccentric position was defined when the artificial eyeball shifted away from the line-of-sight. An illustration of calibration target using artificial eyeball is illustrated in Figure 3.46.

The effect of lens undistortion algorithm was investigated. In this experiment, the data was collected from both distorted and undistorted lens. The undistorted result was also compared to single camera method proposed by [5]. Furthermore, the effect of inaccuracy in measurement of physical parameters of the eye was also probed. In this experiment, ±5% and ±10% bias factors of eye radius were introduced to simulate error on determining physical parameters of the eye. The data collecting was conducted five times using normal (unbiased), overestimated (5% and 10% bias factors), and underestimated (-5% and -10% bias factors) eyeball radius.

3.3.2 Computational Time

One major aim of measuring computational time was to investigate whether the eye movements tracking and visualization system could be used in real-time measurement. The measurement was conducted on a personal computer with Intel® 2 Quad CPU Q6600 @ 2.4 GHz, RAM 2 GB, and Windows XP® Service Pack 3 operating system at 320 x 240 pixels image size. Computational time of single frame processing was measured by computing average processing time of 1500 frames.

3.3.3 Practical Implementation on Human

3.3.3.1 Voluntary Eye Movements

The eye movements tracking and visualization system was implemented to detect voluntary eye movements in real-time operation. The aim of this experiment was to investigate system implementation on several patients with various ages, sexes, and physical parameters of the eye (iris pattern, eye radius, and iris diameter). Voluntary eye movement is a movement of the eye that can be consciously controlled by human, including fixating eye gaze to specific target (*fixation*) or tracking a moving target (*smooth pursuit*). Generally, voluntary eye movements consist of vertical and horizontal movements. Visual stimulus was used to stimulate and standardize the voluntary eye movements.

The 21" wide screen ACER® monitor was used to display visual stimulus in experiment routine. The resolution of the monitor was 1440 x 900 pixels with 60Hz refresh rate. In order to allow the patient gazed on the target, the front side cover of the goggle was disclosed. In this experiment, the measurement was performed on the right eye. Initially, the physical parameters of the patient's eye, such as eye radius, iris diameter, and iris striation, was collected by the operator. Then, the physical parameters data was entered into the system. Finally, several parameters in the system was configured to allow the system tracked the pupil and the iris striation properly. Figure 3.47 shows adjustment of the system to track the patient's right eye.

The patient was positioned approximately 550mm at the front of the monitor. The center of the screen was located straight ahead to the right eye. The vertical and horizontal amplitudes of the stimulus were 97mm and 200mm, respectively. The amplitudes were adjusted in order to allow the patient followed the target at maximum excursion 10° in vertical direction and 20° in horizontal direction as shown in Figure 3.48.

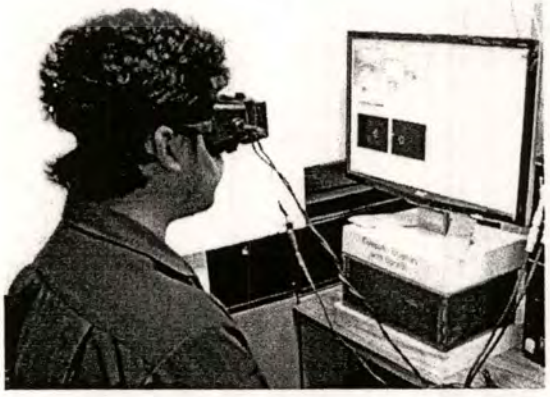
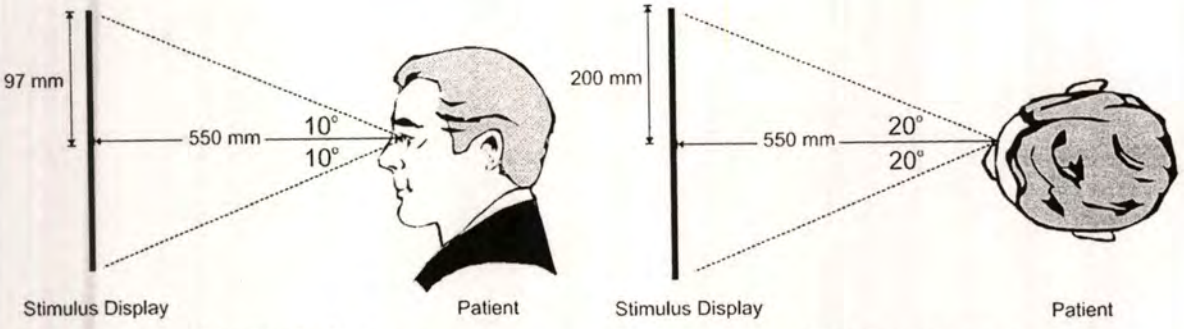


Figure 3.47 Adjustment of the system before experiment



(a) Vertical gaze (b) Horizontal gaze

Figure 3.48 The configuration of voluntary eye movements measurement experiment

The visual stimulus was developed using OpenGL library. It consisted of a white point, which is actually a small rectangle with size 25 x 25 pixels, acting as a moving target in black background. The usage of small white rectangular target and black background for visual stimulus was mainly adapted from previous research [78, 79]. The initial position of the target was in the center of the monitor which was set as the reference position for the measurement.

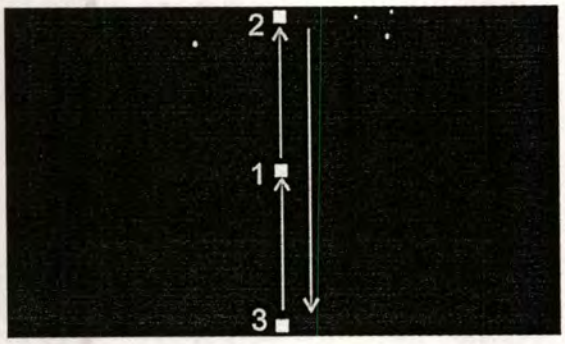
Four tests were performed in this experiment. Each test was performed about 30 seconds. A 1 minute break-time was given between two tests to avoid dizziness on the patient. The tests are:

- **Vertical smooth pursuit test:** patient was asked to follow a white target that moved vertically on sinusoidal wave pattern with amplitude 10° and period 2s (Figure 3.49 (a)). The sequence of the target position was 1-2-3-1-2-3-1……1.

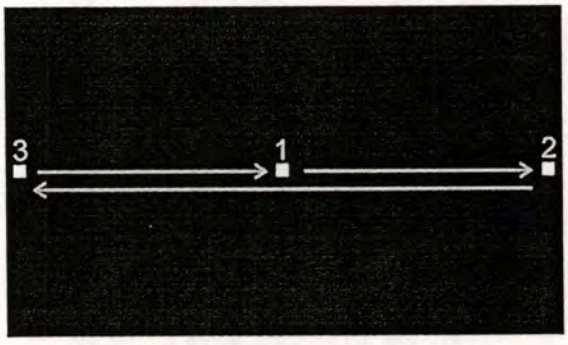
- **Horizontal smooth pursuit test:** patient was asked to follow a white target that moved horizontally on sinusoidal wave pattern with amplitude 20° and period 2s (Figure 3.49 (b)). The sequence of the target position was 1-2-3-1-2-3-1……1.

- **Vertical fixation test:** patient was asked to fixate the gaze on a white target for 1s that appeared alternately on the top and bottom of the monitor (Figure 3.49 (c)). The movement of the target was generated by square wave pattern with amplitude 10° and period 2s. The sequence of the target position was 0-1-2-1-2-1-2……0.

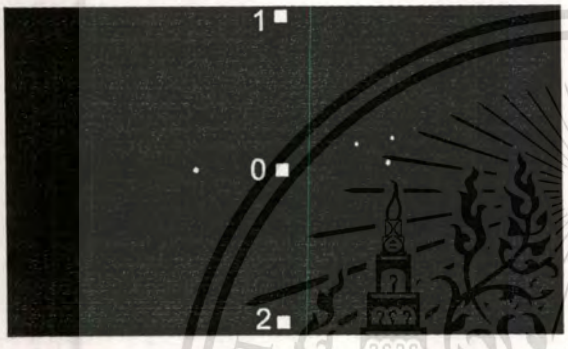
- **Horizontal fixation test:** patient was asked to fixate the gaze on a white target for 1s that appeared alternately on the right and left of the monitor (Figure 3.49 (d)). The movement of the target was generated by square wave pattern with amplitude 20° and period 2s. The sequence of the target position was 0-1-2-1-2-1-2……0.



(a) Vertical smooth pursuit test



(b) Horizontal smooth pursuit test



(c) Vertical fixation test

(d) Horizontal fixation test

Figure 3.49 Illustration of target positions in each voluntary eye movements test

3.3.3.2 Involuntary Eye Movements

The eye movements tracking and visualization system was used to observe involuntary eye movements. Involuntary eye movements is a movement of the eye that can not be controlled consciously by human, including Vestibulo-Ocular Reflex (VOR) and several eye movements disorders, as explained in chapter 2. The aim of this experiment was to investigate system implementation on a patient that suffered from eye movements disorder. In this case, the system was used to detect symptom of Benign Paroxysmal Positional Vertigo (BPPV). The experiment was held at Outpatient Department (OPD) Ear Nose Throat, Srinakarinviroj Hospital, Pathumthani, Thailand.

The inner ear consists of three semicircular canals that contribute to a sense of balance. Over time, calcium carbonate particles in the inner ear can break off and accumulate behind a membrane in the posterior canal. This particle can be thought of as "ear rocks" (*otoconia*). Moving the head in certain directions may cause the ear rocks to tug on hairlike sensors, triggering a type of dizziness called BPPV. In most cases when it is found that a patient has BPPV, a noninvasive positioning procedure is used to clear the ear rocks out of the canal and deposit them back into the part of the inner ear where they belong. Figure 3.50 illustrates the location of the ear rocks inside the inner ear.

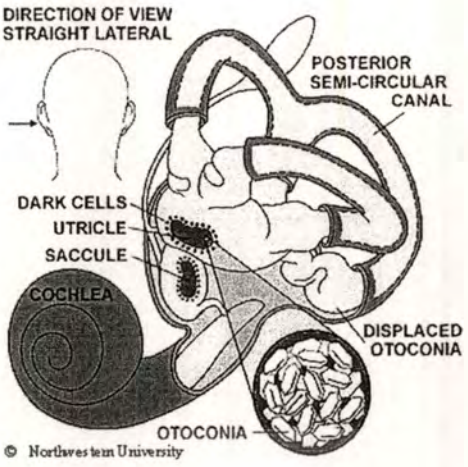
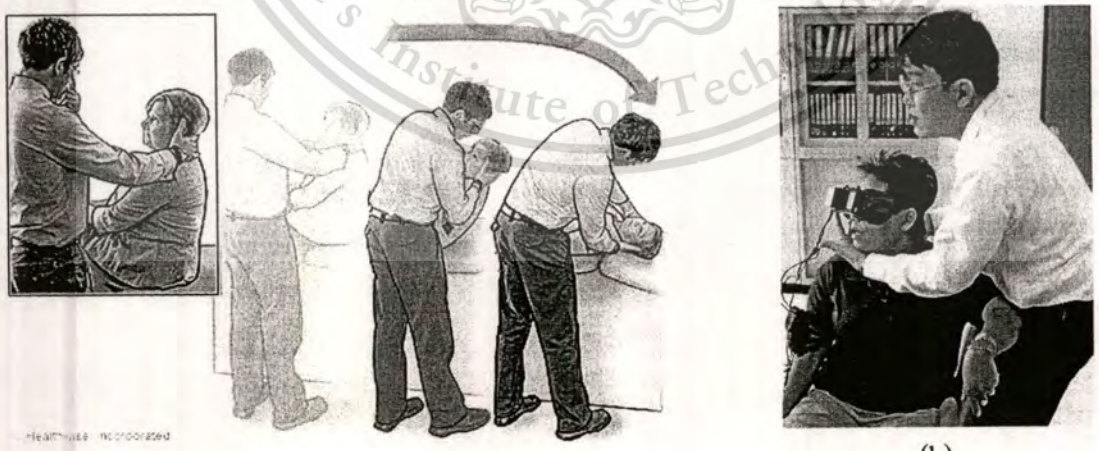


Figure 3.50 Location of the ear rocks inside the inner ear [80]

The Dix-Hallpike test was performed to investigate BPPV symptoms as shown in Figure 3.51(a) [81, 82]. During this test, the patient was brought from sitting to a supine position, with the head turned 45° to one side and extended about 20° backward. Once supine, the eyes were typically observed for about 30s. If no nystagmus ensued, the patient was brought back to sitting. There was a delay of about 30s again, and then the other side was tested.

Figure 3.51(b) shows the Dix-Hallpike test conducted at Srinakarinvirooj Hospital with the eye movements tracking and visualization system used to record and detect the nystagmus. The test was performed on complete darkness by installing the front side cover on the goggle because it had been reported in [83] that the ocular torsional response in reduced light was larger than the torsional response in light.



(a) (b)

Figure 3.51 Dix-Hallpike test to investigate BPPV symptoms:

(a) Procedure of Dix-Hallpike test [82];

(b) The eye movements tracking and visualization system was used during Dix-Hallpike test

Chapter 4

Experimental Results

In this chapter, system validation and implementation results are presented. Initially, a camera calibration procedure was performed to obtain intrinsic, extrinsic, and lens distortion parameters of the camera. Next, artificial eyeball and gimbal system were used in *in vitro* system validation. Computational time was also measured to investigate real-time performance. Finally, the system was implemented on human to detect voluntary and involuntary eye movements.

4.1 System Validation

The camera calibration results were shown in Table 4.1 and Table 4.2 as follows:

Table 4.1 Calibration result of camera 1

Parameter	Value (pixels)
Focal Length	$\alpha_x = 437.537$; $\alpha_y = 479.815$
Principal Point	$x_0 = 160.800$; $y_0 = 146.644$
Distortion coefficients	$k_1 = -0.476785$; $k_2 = -0.290641$; $p_1 = -0.011297$; $p_2 = -0.001822$
Rotation Matrix	$R_1 = \begin{bmatrix} 0.99 & -0.09 & 0.10 \\ -0.11 & -0.96 & 0.25 \\ 0.07 & -0.26 & -0.96 \end{bmatrix}$
Translation Vector	$T_1 = [3.94 \quad 9.49 \quad -46.31]^T$

Table 4.2 Calibration result of camera 2

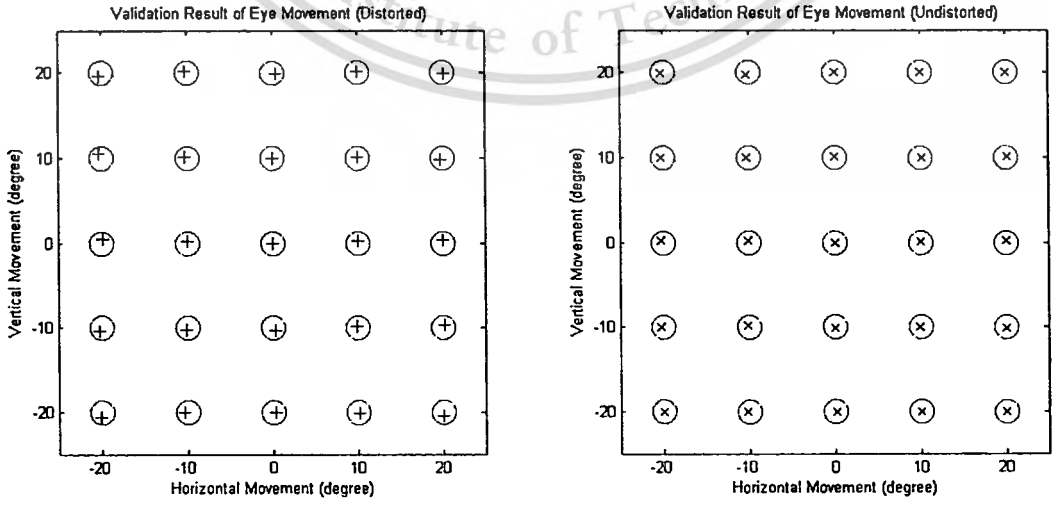
Parameter	Value (pixels)
Focal Length	$\alpha_x = 428.224$; $\alpha_y = 466.388$
Principal Point	$x_0 = 175.369$; $y_0 = 149.405$
Distortion coefficients	$k_1 = -0.126796$; $k_2 = -0.721747$; $p_1 = -0.003700$; $p_2 = -0.032267$
Rotation Matrix	$R_2 = \begin{bmatrix} 0.99 & -0.07 & 0.08 \\ -0.08 & -0.97 & 0.22 \\ 0.06 & -0.23 & -0.97 \end{bmatrix}$
Translation Vector	$T_2 = [-10.62 \quad 10.77 \quad -46.74]^T$

The intrinsic and extrinsic parameters of the camera were measured in pixels unit. Focal length in x and y directions are denoted as α_x and α_y, respectively. Principal point is denoted by coordinates (x_0, y_0). Radial distortion coefficients are k_1 and k_2, while tangential distortion coefficients are p_1 and p_2. The extrinsic parameters consist of rotation and translation matrices which are denoted as matrix R and matrix T, respectively.

4.1.1 Effect of Lens Undistortion

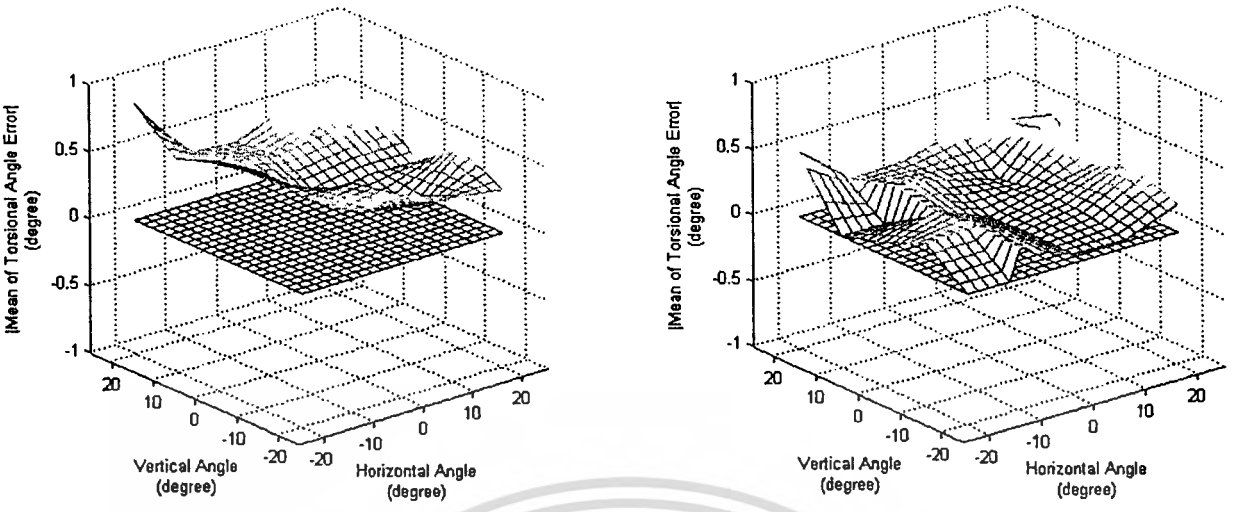
The effect of lens undistortion algorithm was verified using artificial eyeball and gimbal system. The artificial eyeball was rotated in horizontal and vertical angles with excursion ranging ±20° at steps of 10°. At each point, five different angular torsional positions were made, resulting 125 unique positions consisted of horizontal-vertical-torsional (HVT) values. At each unique position, 5 images were acquired using dual cameras system to ensure the accuracy of measurement. The average of horizontal, vertical, and torsional values were then computed from 5 images and defined as the measured HVT value for each unique position.

Figure 4.1 illustrates the results of calibration target validation from distorted and undistorted lens. Ideal calibration angle is denoted by "O". Actual angles determined from software processing are denoted by "+" and "×" for distorted and undistorted lens, respectively. A close proximity between ideal calibration angle and actual angle was proven by low degree of error. Furthermore, the lens undistortion algorithm was able to improve horizontal and vertical tracking accuracy, particularly at the eccentric positions as shown at (-20°, 20°), (-20°, -20°), (-20°, 10°), (-20°, -10°), (20°, 10°), (20°, -10°), and (20°, -20°).



(a) Distorted Lens (b) Undistorted Lens

Figure 4.1 Results of calibration target validation tested using artificial eyeball



(a) Distorted Lens (b) Undistorted Lens

Figure 4.2 Torsional error in horizontal and vertical excursions

The effect of lens undistortion algorithm for torsional measurement was also investigated as shown in Figure 4.2. Each torsion value for the graphical plots was obtained by computing absolute mean of torsional angular error from five different torsional settings at each calibration target position. The ideal torsional angular error was plotted as a plane graphical plot at zero angle while the actual error was plotted above it. Figure 4.2 (a) shows absolute means of torsional error of distorted lens (max. value = 0.87°, min. value = 0.17°). Figure 4.2 (b) shows absolute means of torsional error of undistorted lens (max. value = 0.49°, min. value = 0.06°). The experimental data demonstrates that the lens undistortion algorithm improved the accuracy of torsional measurement at each calibration target position significantly. The error in torsional angular position was primarily caused by incorrect placement of the iris template due to lens distortion.

Table 4.3 Absolute mean of error of distorted and undistorted lens (in degrees)

Method	Horizontal	Vertical	Torsional
Distorted Lens	0.32	0.38	0.47
Undistorted Lens	0.15	0.14	0.20

Table 4.4 Absolute standard deviation of error of distorted and undistorted lens (in degrees)

Method	Horizontal	Vertical	Torsional
Distorted Lens	0.24	0.26	0.30
Undistorted Lens	0.14	0.14	0.18

Statistical representations of horizontal, vertical, and torsional values taken from 125 unique points are presented in Table 4.3 and Table 4.4. Error reduction ε after lens undistortion was computed using the following formula:

$$\varepsilon = \frac{|Mean|_{distorted} - |Mean|_{undistorted}}{|Mean|_{distorted}} \times 100\% \quad (4.1)$$

The lens undistortion algorithm was able to reduce the errors in horizontal, vertical, and torsional angular positions up to 53.13%, 63.16%, and 57.45%, respectively.

The performance of our system was also compared to single camera method proposed in [5] as similar *in vitro* validation test was also used in our research. The statistical comparison results are shown in Table 4.5 and Table 4.6. By adjusting the camera manually using four equidistant infrared reflections, the single camera method provided high accuracy in horizontal (mean = 0.20°, S.D. = 0.14 °), vertical (mean = 0.30°, S.D. = 0.16°), and torsional (mean = 0.38°, S.D. = 0.28 °) angular positions. Small translational misalignment of the camera which led to geometric distortion was found to influence the tracking's error on single camera method. In our study, we used a new method to reduce the error caused by geometric distortion and lens aberration using 3D coordinates extraction and lens undistortion algorithms. The experimental results reveal that our system exhibited better performance in horizontal (mean = 0.15°, S.D. = 0.14 °), vertical (mean = 0.14°, S.D. = 0.14 °), and torsional (mean = 0.20°, S.D. = 0.18°) angular positions compared to single camera method. Complete validation results of our method for distorted and undistorted lens are shown in Table 4.7 and Table 4.8, respectively.

Table 4.5 Absolute mean of error of single camera & dual cameras methods (in degrees)

Method	Horizontal	Vertical	Torsional
Single Camera + Marker Array	0.20	0.30	0.38
Our Method (Dual Cameras, Undistorted)	0.15	0.14	0.20

Table 4.6 Absolute standard deviation of error of single camera & dual cameras methods (in degrees)

Method	Horizontal	Vertical	Torsional
Single Camera + Marker Array	0.14	0.16	0.28
Our Method (Dual Cameras, Undistorted)	0.14	0.14	0.18

Table 4.7 Horizontal, vertical, and torsional validation result of distorted lens (in degrees)

Degrees	Torsion	-20 (right)			-10 (right)			0			10 (left)			20 (left)		
		Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.
20 (down)	20 (CCW)	-19.98	19.64	20.71	-10.80	20.14	19.59	0.36	20.17	20.39	10.74	20.08	19.70	19.80	20.90	20.90
	10 (CCW)	-20.58	20.28	10.51	-9.85	19.54	10.60	0.19	19.75	9.82	9.46	19.56	9.49	19.63	20.80	10.98
	0	-19.98	20.27	-0.37	-10.26	19.95	0.62	0.59	19.88	-0.21	10.04	19.97	-0.29	20.61	19.93	0.90
	-10 (CW)	-19.43	20.95	-10.98	-9.79	20.97	-10.46	0.14	20.02	-10.28	10.41	20.38	-10.53	19.70	20.70	-10.70
	-20 (CW)	-20.50	20.93	-19.53	-10.28	20.06	-19.65	0.64	20.30	-20.69	10.76	20.52	-20.27	20.67	20.40	-20.88
10 (down)	20 (CCW)	-20.68	9.53	20.37	-9.89	10.10	20.79	0.40	9.99	20.17	9.90	10.60	20.90	19.86	9.62	20.90
	10 (CCW)	-20.03	9.62	9.89	-10.32	9.89	10.69	-0.02	9.78	10.46	9.93	9.70	9.90	20.43	10.80	9.08
	0	-20.27	10.37	0.90	-9.98	9.61	0.66	-0.09	10.24	-0.43	9.97	10.15	-0.21	20.40	10.36	0.14
	-10 (CW)	-19.59	9.47	-9.78	-10.07	10.29	-10.00	0.71	10.59	-10.36	10.51	10.71	-9.61	20.43	10.80	-9.98
	-20 (CW)	-20.20	9.47	-20.72	-9.40	9.62	-19.92	0.71	10.75	-19.45	9.87	10.16	-20.70	19.84	10.40	-19.01
0	20 (CCW)	-20.04	-0.90	19.84	-10.20	-0.60	20.26	0.36	-0.06	19.99	10.68	-0.45	20.97	19.50	-0.70	20.90
	10 (CCW)	-20.80	-0.58	9.93	-10.47	-0.21	10.22	0.03	-0.02	9.96	9.94	-0.91	10.04	19.48	-0.51	10.90
	0	-19.98	-0.03	-0.38	-9.99	-0.13	-0.07	0.00	0.00	0.00	9.49	0.07	0.50	20.46	0.64	0.80
	-10 (CW)	-19.53	-0.41	-10.37	-10.04	-0.43	-10.12	-0.12	-0.24	-9.84	9.50	-0.01	-9.52	19.56	0.40	-9.50
	-20 (CW)	-19.71	-0.01	-20.32	-9.70	-0.20	-20.68	-0.35	-0.20	-19.90	10.02	-0.13	-20.27	19.97	0.53	-20.47
-10 (up)	20 (CCW)	-20.06	-10.18	20.41	-10.44	-10.81	19.87	0.13	-10.75	19.74	9.85	-10.31	20.90	20.70	-10.80	20.89
	10 (CCW)	-19.73	-9.80	10.07	-10.03	-10.36	10.41	-0.05	-10.21	9.93	10.45	-9.92	9.57	20.80	-10.72	10.98
	0	-19.54	-10.49	0.74	-9.93	-9.92	-0.02	-0.01	-10.19	-0.23	9.88	-10.12	0.46	19.60	-10.26	0.08
	-10 (CW)	-20.32	-9.77	-10.27	-9.80	-9.80	-9.94	-0.05	-9.65	-10.32	10.58	-10.03	-9.70	19.78	-10.72	-9.09
	-20 (CW)	-19.20	-9.44	-20.90	-10.42	-9.98	-19.76	-0.20	-9.65	-19.70	10.21	-10.34	-20.18	20.25	-10.26	-20.99
-20 (up)	20 (CCW)	-20.38	-19.49	20.90	-9.97	-20.05	20.90	0.74	-20.70	19.64	9.99	-20.47	19.01	20.80	-19.90	20.98
	10 (CCW)	-19.79	-20.49	9.98	-9.71	-20.42	10.50	0.55	-20.56	10.76	10.05	-20.32	9.49	20.58	-19.78	10.90
	0	-19.92	-20.12	0.41	-10.04	-20.15	-0.13	0.47	-19.47	-0.39	10.13	-20.10	0.70	19.89	-19.80	0.72
	-10 (CW)	-20.08	-19.80	-10.20	-10.28	-19.33	-9.68	-0.44	-19.12	-9.30	10.69	-19.90	-10.50	20.18	-19.59	-10.10
	-20 (CW)	-20.67	-19.48	-19.95	-10.14	-20.80	-19.98	-0.30	-19.94	-20.27	10.24	-20.30	-20.87	19.57	-19.10	-20.89

Table 4.8 Horizontal, vertical, and torsional validation result of undistorted lens (in degrees)

Degrees	Torsion	-20 (right)			-10 (right)			0			10 (left)			20 (left)		
		Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.
20 (down)	20 (CCW)	-19.97	20.19	19.72	-10.03	20.06	20.22	0.88	20.00	20.36	10.35	19.83	20.12	19.90	19.99	20.58
	10 (CCW)	-20.11	19.92	10.13	-9.88	20.02	10.56	0.09	20.23	10.25	10.30	20.26	10.19	20.43	20.39	9.51
	0	-20.06	20.26	0.39	-10.06	20.05	0.52	-0.13	20.13	0.05	10.18	20.21	0.04	20.08	19.76	0.16
	-10 (CW)	-19.95	19.82	-10.21	-9.98	20.05	-10.14	-0.20	19.91	-10.28	10.15	20.07	9.82	19.44	19.76	-10.78
	-20 (CW)	-20.02	20.00	-20.36	-10.14	20.42	19.96	-0.05	20.06	-20.02	10.18	20.04	19.66	20.15	20.62	-20.46
10 (down)	20 (CCW)	-20.03	10.19	19.94	-10.02	10.17	19.95	0.12	10.21	20.10	10.27	9.95	20.35	20.17	10.24	19.52
	10 (CCW)	-20.06	9.95	10.31	-9.91	10.05	10.01	0.07	10.25	10.19	10.20	9.83	10.44	20.15	9.78	10.31
	0	-20.23	10.26	-0.90	-9.91	10.05	0.30	-0.04	9.94	0.03	9.97	9.97	0.32	20.21	10.11	0.09
	-10 (CW)	-20.14	10.14	-10.31	-10.01	10.00	-9.97	-0.19	10.02	-10.06	9.65	10.14	-9.96	20.03	10.04	-9.87
	-20 (CW)	-20.01	9.99	-20.66	-10.06	9.98	-20.00	-0.09	10.10	-20.58	10.34	9.76	-20.21	20.19	10.25	-19.67
0	20 (CCW)	-20.14	-0.38	20.54	-10.02	-0.11	20.18	0.36	0.06	19.99	9.90	0.23	19.78	20.35	0.08	19.91
	10 (CCW)	-20.28	-0.16	10.11	-10.07	-0.01	10.08	0.03	-0.02	9.96	9.94	0.08	10.10	20.04	0.28	10.02
	0	-19.93	-0.26	0.15	-10.04	0.04	-0.29	0.00	0.00	0.00	9.89	0.08	-0.29	20.19	-0.10	0.22
	-10 (CW)	-19.89	-0.25	-10.26	-10.12	-0.24	-10.14	-0.12	-0.24	-9.84	10.28	-0.15	-10.22	19.87	-0.19	-9.91
	-20 (CW)	-19.95	-0.38	-19.97	-10.16	-0.17	-20.02	-0.35	-0.20	-19.90	10.34	-0.57	-19.71	20.28	-0.56	-20.07
-10 (up)	20 (CCW)	-20.29	-10.00	20.01	-10.21	-10.08	20.06	0.08	-10.05	20.06	10.45	-10.00	20.42	20.39	-10.19	19.88
	10 (CCW)	-20.28	-10.27	10.59	-10.08	-10.09	10.12	0.06	-10.10	10.01	10.10	-10.06	9.98	20.19	-9.94	10.19
	0	-20.22	-10.05	-0.33	-10.31	-9.98	0.37	0.03	-9.91	-0.19	10.09	-9.98	0.20	20.08	-10.06	-0.18
	-10 (CW)	-19.72	-10.23	-10.29	-10.34	-10.19	-10.02	-0.05	-9.94	-10.02	10.10	-10.04	-10.09	20.11	-9.87	-10.44
	-20 (CW)	-20.34	-10.34	-20.43	-10.05	-10.08	-20.01	-0.09	-10.86	-19.91	10.09	-9.98	-19.46	20.08	-10.07	-20.26
-20 (up)	20 (CCW)	-20.06	-20.07	19.95	-10.13	-20.14	20.05	0.05	-19.90	19.97	10.08	-19.95	20.23	20.29	-20.14	20.21
	10 (CCW)	-20.36	-19.86	10.06	-10.09	-20.05	10.10	0.05	-20.01	10.08	10.09	-19.94	10.48	20.04	-19.96	10.10
	0	-19.90	-20.10	0.38	-10.09	-19.92	0.02	0.09	-20.32	-0.09	10.08	-19.95	0.10	20.10	-19.79	-0.54
	-10 (CW)	-19.95	-20.42	-10.37	-9.91	-20.01	-10.11	-0.01	-19.99	-10.09	10.54	-19.79	-9.78	20.49	-20.00	-10.04
	-20 (CW)	-19.90	-20.00	-19.82	-10.05	-19.84	-20.01	-0.04	-20.20	-20.04	10.13	-19.56	-19.90	20.06	-20.03	-20.27

4.1.2 Effect of Inaccuracy in Measurement of Physical Parameters of The Eye

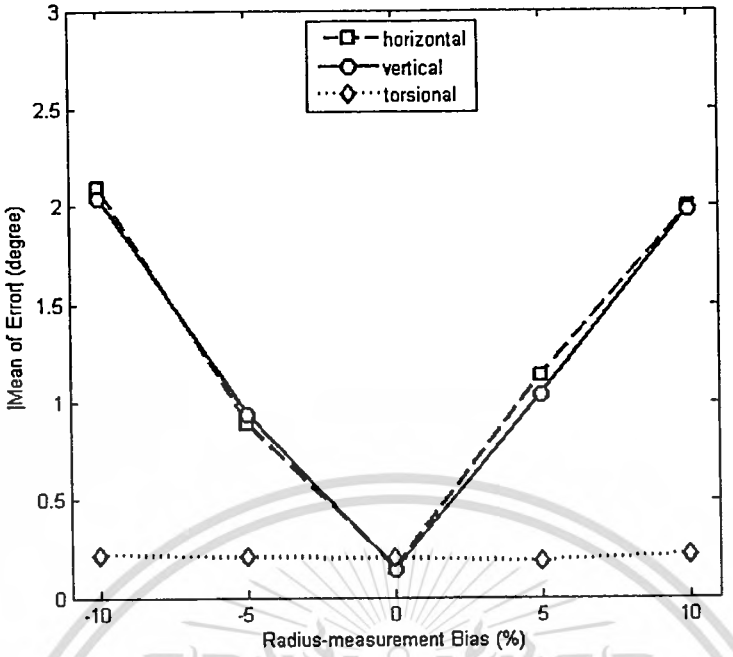
The effect of inaccuracy in measurement of physical parameters of the eye was simulated. The experiment was conducted five times using normal (unbiased), overestimated (5% and 10% bias factors), and underestimated (-5% and -10% bias factors) eyeball radius. The data was collected from 125 unique positions of calibration target. The statistical results are shown in Table 4.9 and Table 4.10. Figure 4.3 shows the error related with determining horizontal, vertical, and torsional angular eye positions as the measurement of the eyeball radius provided to the algorithm was biased up to ±10%. The 3D angular positions of the eye were found to be affected by inaccuracy of measuring eyeball radius. For example, if the eyeball radius was underestimated by a factor of 10%, error in determining eye position increased in both horizontal (mean = 2.09°, S.D. = 0.95°) and vertical (mean = 2.04°, S.D. = 0.96°) positions. In the other side, error in determining torsional eye position (mean = 0.22°, S.D. = 0.18°) was more robust and only slightly influenced by the 10% underestimated eyeball radius.

Table 4.9 Absolute mean of error for simulation of inaccuracy in measuring eyeball radius (in degrees)

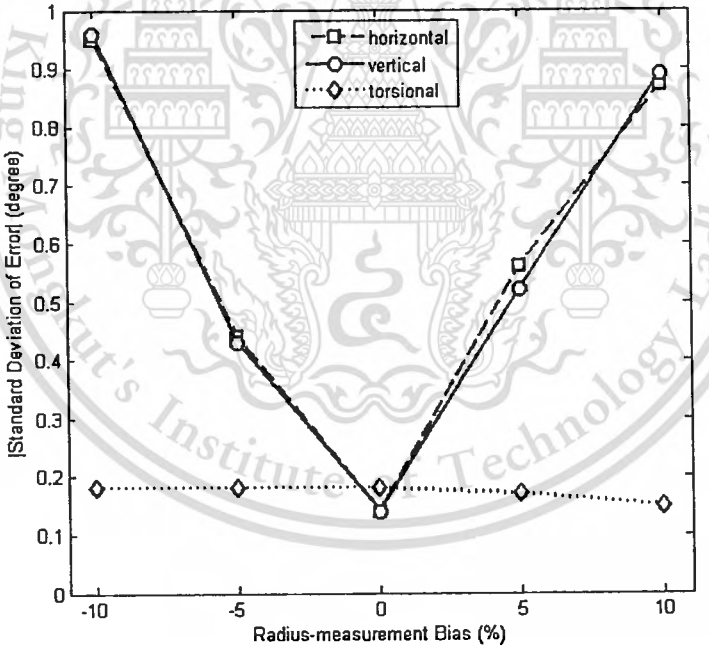
Bias (%)	Eyeball Radius (mm)	Horizontal	Vertical	Torsional
-10	13.50	2.09	2.04	0.22
-5	14.25	0.89	0.94	0.21
0	15.00	0.15	0.14	0.20
5	15.75	1.14	1.04	0.19
10	16.50	1.99	1.98	0.22

Table 4.10 Absolute standard deviation of error for simulation of inaccuracy in measuring eyeball radius (in degrees)

Bias (%)	Eyeball Radius (mm)	Horizontal	Vertical	Torsional
-10	13.50	0.95	0.96	0.18
-5	14.25	0.44	0.43	0.18
0	15.00	0.14	0.14	0.18
5	15.75	0.56	0.52	0.17
10	16.50	0.87	0.89	0.15



(a) Absolute mean of error



(b) Absolute standard deviation of error

Figure 4.3 Experimental error produced by inaccuracy in measuring eyeball radius.

Table 4.11 and Table 4.12 show 3D angular positions result of undistorted lens with 5% and 10% overestimated eyeball radius, respectively. Table 4.13 and Table 4.14 show 3D angular positions result of undistorted lens with 5% and 10% underestimated eyeball radius, respectively. The angular measurement was based on right-handed coordinate system.

Table 4.11 Horizontal, vertical, and torsional validation result of undistorted lens (in degrees, bias: 5%, eyeball radius = 15.75 mm)

Degrees	Torsion	-20 (right)			-10 (right)			0			10 (left)			20 (left)		
		Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.
20 (down)	20 (CCW)	-18.78	18.66	20.14	-7.74	18.15	20.02	0.13	18.31	19.92	8.99	18.96	20.56	18.15	18.03	20.11
	10 (CCW)	-18.65	18.37	10.09	-8.87	18.70	9.98	0.69	18.79	10.23	8.95	19.10	9.89	18.75	18.97	10.10
	0	-18.07	18.68	-0.17	-8.86	18.80	-0.16	-0.26	18.83	0.36	9.36	19.15	0.59	18.96	18.14	0.68
	-10 (CW)	-19.17	19.03	-10.06	-8.58	18.58	-10.62	-0.19	18.68	-10.25	8.81	18.93	-10.04	18.92	18.76	-10.01
	-20 (CW)	-19.03	18.96	-20.12	-8.96	18.51	-20.30	-0.36	19.20	-20.70	9.09	19.21	-19.94	18.84	18.35	-20.11
10 (down)	20 (CCW)	-18.13	8.98	20.23	-8.14	8.59	20.18	0.01	8.70	20.12	9.00	9.23	20.15	18.86	8.81	20.00
	10 (CCW)	-18.12	8.28	10.05	-8.57	8.54	9.98	-0.43	8.89	10.52	8.55	8.71	9.84	18.21	8.46	10.02
	0	-19.19	8.33	0.15	-8.64	8.57	0.31	-0.21	9.15	-0.02	9.09	9.12	0.26	18.89	8.27	-0.26
	-10 (CW)	-18.30	8.36	-10.02	-8.63	8.79	-10.25	-0.51	9.07	-9.97	8.58	8.94	-10.02	18.40	8.57	-9.94
	-20 (CW)	-18.83	8.81	-20.13	-8.49	9.19	-19.90	-0.15	9.14	-20.02	8.54	8.91	-20.04	18.49	8.44	-20.05
0	20 (CCW)	-18.02	-0.28	19.89	-8.26	0.49	20.32	0.08	0.12	20.49	9.28	0.22	20.08	18.75	0.31	20.05
	10 (CCW)	-18.78	-0.21	10.16	-8.22	0.21	10.24	0.03	0.05	9.92	9.11	0.04	10.09	18.65	0.12	10.01
	0	-18.04	-0.04	0.30	-8.58	-0.29	0.04	0.00	0.00	0.00	9.14	-0.14	-0.26	18.04	0.04	0.67
	-10 (CW)	-18.35	-0.22	-9.72	-8.88	-0.30	-10.24	-0.09	-0.44	-10.12	9.19	-0.32	-10.02	18.01	-0.46	-10.07
	-20 (CW)	-18.40	-0.21	-20.06	-8.12	-0.05	-20.19	-0.02	-0.20	-19.96	8.94	-0.02	-20.04	18.20	-0.18	-20.11
-10 (up)	20 (CCW)	-18.05	-8.27	20.13	-9.00	-9.51	19.95	0.67	-9.01	20.08	8.90	-8.92	19.70	18.20	-8.77	20.23
	10 (CCW)	-18.82	-8.52	10.40	-8.95	-8.50	9.95	0.04	-8.00	10.15	8.67	-8.99	9.97	18.08	-8.87	10.00
	0	-18.84	-9.13	0.45	-8.56	-8.53	-0.60	0.09	-9.03	0.31	9.02	-8.80	0.22	18.87	-9.04	0.12
	-10 (CW)	-18.13	-8.43	-9.62	-8.04	-8.44	-9.98	-0.12	-8.70	-10.06	8.75	-8.78	-10.10	18.50	-8.71	-10.04
	-20 (CW)	-18.52	-9.02	-20.05	-8.51	-9.22	-20.17	-0.06	-8.97	-20.11	8.81	-8.66	-20.18	18.72	-8.19	-20.08
-20 (up)	20 (CCW)	-18.74	-18.89	20.31	-8.95	-19.41	20.46	1.11	-18.96	20.18	8.97	-20.92	20.51	18.89	-18.22	20.04
	10 (CCW)	-18.00	-18.72	9.87	-8.37	-18.70	10.13	0.31	-18.94	9.97	9.09	-19.13	10.25	18.41	-18.82	10.10
	0	-18.83	-18.13	-0.39	-8.54	-18.41	0.04	0.64	-19.40	-0.19	8.86	-19.15	0.01	19.04	-18.43	0.50
	-10 (CW)	-18.83	-18.85	-9.43	-8.26	-18.35	-10.20	-0.34	-18.99	-9.42	8.96	-18.98	-10.52	18.42	-18.58	-9.96
	-20 (CW)	-18.83	-18.60	-20.30	-8.07	-18.14	-20.19	-0.33	-19.03	-20.21	8.57	-18.75	-19.54	18.61	-17.97	-20.28

Table 4.12 Horizontal, vertical, and torsional validation result of undistorted lens (in degrees, bias: 10%, eyeball radius = 16.50 mm)

Degrees	Torsion	-20 (right)			-10 (right)			0			10 (left)			20 (left)		
		Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.
20 (down)	20 (CCW)	-17.05	17.02	20.23	-7.70	17.90	20.41	0.23	18.21	20.65	7.56	17.78	20.64	17.71	17.82	19.84
	10 (CCW)	-17.45	17.47	10.30	-7.79	17.25	10.19	0.25	17.84	9.87	7.40	17.85	9.83	17.35	17.88	10.29
	0	-17.34	17.07	-0.21	-7.24	17.61	-0.39	0.10	17.18	0.45	7.46	17.52	0.14	17.53	17.60	0.07
	-10 (CW)	-18.05	17.12	-10.01	-7.70	17.40	-10.22	0.77	17.43	-10.37	7.24	18.31	-10.25	18.04	17.82	-10.01
	-20 (CW)	-17.47	17.18	-19.93	-7.51	17.31	-20.26	0.32	17.12	-19.77	7.91	17.85	-20.54	17.93	17.49	-20.51
10 (down)	20 (CCW)	-18.18	7.80	20.19	-7.38	7.39	20.29	0.40	7.43	19.77	7.16	7.05	20.06	17.53	7.77	20.25
	10 (CCW)	-17.68	7.90	10.01	-7.94	7.64	10.18	0.12	8.29	10.29	7.68	7.94	10.37	17.71	7.74	10.22
	0	-17.24	7.67	0.32	-7.25	7.54	0.19	0.12	8.29	-0.36	7.49	7.68	0.28	17.61	7.51	0.11
	-10 (CW)	-17.22	7.64	-9.80	-7.42	7.34	-10.17	0.12	7.30	-9.90	7.44	7.96	-10.04	17.81	7.66	-10.21
	-20 (CW)	-17.54	7.81	-19.90	-7.38	7.34	-20.26	0.54	7.84	-20.38	7.29	7.73	-19.88	17.96	7.30	-20.14
0	20 (CCW)	-17.30	0.39	20.03	-7.17	0.37	19.97	0.01	0.18	20.22	7.99	0.52	20.39	17.64	0.51	20.26
	10 (CCW)	-17.18	0.11	10.30	-7.46	0.26	10.16	0.15	0.30	10.36	7.90	0.31	10.33	17.86	0.14	10.30
	0	-17.46	-0.34	0.08	-7.72	0.43	0.70	0.00	0.00	0.00	7.43	0.18	-0.09	17.30	0.76	-0.13
	-10 (CW)	-17.48	-0.14	-10.22	-8.08	-0.34	-9.72	-0.40	0.24	-10.01	7.68	-0.19	-10.11	17.49	0.45	-10.15
	-20 (CW)	-17.92	-0.15	-19.89	-7.17	-0.37	-19.97	-0.08	0.13	-19.70	7.56	-0.32	-20.02	17.65	-0.47	-20.11
-10 (up)	20 (CCW)	-17.43	-7.96	20.20	-7.27	-7.22	20.19	0.65	-7.29	20.20	7.54	-7.90	20.19	17.86	-7.77	19.74
	10 (CCW)	-17.40	-7.47	10.10	-7.84	-7.82	10.27	0.39	-7.42	9.72	7.25	-7.46	10.15	17.85	-7.49	10.24
	0	-17.40	-7.55	0.18	-7.61	-7.43	0.01	-0.28	-8.07	-0.22	7.23	-8.12	0.26	17.74	-7.74	0.22
	-10 (CW)	-17.91	-7.09	-10.21	-7.49	-7.65	-10.03	-0.32	-8.05	-10.01	7.64	-7.77	-9.94	17.86	-7.22	-10.02
	-20 (CW)	-17.22	-7.77	-20.30	-7.88	-7.68	-20.20	-0.67	-7.61	-20.21	7.56	-7.73	-19.72	17.70	-7.38	-20.10
-20 (up)	20 (CCW)	-18.06	-17.42	19.97	-7.97	-17.13	20.40	0.59	-17.52	19.72	8.04	-17.88	20.20	17.34	-17.38	19.80
	10 (CCW)	-17.24	-17.70	10.23	-7.66	-17.65	10.06	-0.60	-17.42	10.48	7.56	-17.66	10.60	18.19	-17.87	10.17
	0	-17.25	-17.07	-0.05	-8.09	-17.34	0.29	-0.83	-17.33	0.23	8.00	-18.06	-0.14	17.76	-17.30	-0.21
	-10 (CW)	-18.08	-17.61	-10.43	-7.77	-17.24	-10.51	-0.09	-17.23	-10.04	7.29	-17.24	-10.22	17.37	-17.67	-10.56
	-20 (CW)	-17.88	-17.21	-20.58	-7.68	-17.85	-20.30	-0.34	-18.08	-20.29	7.17	-17.58	-20.02	17.69	-17.43	-19.85

Table 4.13 Horizontal, vertical, and torsional validation result of undistorted lens (bias: -5%, eyeball radius = 14.25 mm)

Degrees	Torsion	-20 (right)			-10 (right)			0			10 (left)			20 (left)		
		Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.
20 (down)	20 (CCW)	-21.38	21.62	20.05	-11.12	20.95	20.41	0.13	21.25	19.73	11.05	21.07	20.11	21.12	21.05	20.02
	10 (CCW)	-20.80	21.06	10.51	-11.02	21.08	10.22	0.05	21.04	10.02	11.20	20.90	10.40	21.96	21.05	9.81
	0	-21.26	21.11	-0.20	-11.01	20.80	0.02	0.05	21.25	0.16	11.10	21.02	-0.09	20.80	21.04	-0.28
	-10 (CW)	-20.55	21.58	-10.01	-11.14	21.51	-10.11	0.07	21.06	-10.18	10.88	21.11	-10.03	21.10	21.16	-9.91
	-20 (CW)	-20.96	21.40	-20.10	-11.23	21.64	-20.28	0.39	21.36	-20.37	11.42	21.27	-20.10	21.09	21.07	-20.11
10 (down)	20 (CCW)	-21.02	11.76	20.24	-10.88	11.19	20.25	0.11	11.04	19.73	11.05	11.23	20.15	21.86	11.24	19.91
	10 (CCW)	-21.08	10.83	10.10	-11.05	11.39	10.29	0.17	11.05	9.89	11.07	11.15	9.86	21.20	11.18	9.71
	0	-21.38	10.93	0.35	-11.02	11.36	-0.28	0.24	11.03	0.26	11.19	11.02	0.36	21.15	11.18	-0.18
	-10 (CW)	-20.68	11.15	-10.00	-11.07	11.26	-9.85	0.20	11.19	-10.16	10.80	11.00	-9.91	21.09	11.05	-10.03
	-20 (CW)	-20.63	11.43	-20.48	-10.91	10.72	-19.89	0.17	11.09	-20.01	11.24	11.15	-20.98	21.74	11.21	-19.95
0	20 (CCW)	-20.65	0.25	20.36	-11.29	0.15	20.67	0.32	0.20	20.05	11.30	0.04	19.91	21.22	0.21	19.95
	10 (CCW)	-20.22	0.10	10.33	-10.57	0.17	9.95	0.18	0.04	10.13	10.82	0.05	10.16	21.26	0.19	10.24
	0	-20.79	0.29	0.38	-11.02	-0.03	0.29	0.00	0.00	0.00	11.05	0.03	0.30	21.51	-0.08	-0.09
	-10 (CW)	-21.24	-0.35	-10.02	-10.96	-0.24	-10.20	-0.25	-0.35	-9.97	11.31	-0.06	-10.10	20.65	-0.22	-10.14
	-20 (CW)	-20.93	-0.30	-20.39	-11.29	-0.10	-20.98	-0.36	-0.11	-20.14	10.93	-0.06	-20.13	20.75	-0.13	-20.67
-10 (up)	20 (CCW)	-21.19	-10.88	19.94	-11.03	-11.04	20.09	0.28	-11.20	19.91	11.19	-11.21	20.16	21.57	-11.03	20.32
	10 (CCW)	-21.16	-10.61	10.59	-11.00	-10.76	10.39	0.12	-11.14	9.95	11.50	-11.04	10.38	21.46	-11.21	9.98
	0	-21.26	-11.10	-0.16	-10.71	-11.20	0.30	0.09	-11.09	0.06	11.38	-11.26	0.20	21.41	-11.10	0.07
	-10 (CW)	-20.97	-11.54	-10.39	-10.97	-10.98	-10.07	0.17	-11.17	-10.10	11.22	-11.27	-10.26	21.09	-11.13	-9.89
	-20 (CW)	-21.25	-11.50	-20.66	-10.45	-11.17	-20.44	0.13	-11.18	-19.97	10.80	-11.20	-20.21	21.51	-11.15	-20.21
-20 (up)	20 (CCW)	-21.10	-21.03	20.21	-11.02	-21.26	20.20	0.37	-21.03	20.36	11.13	-21.09	20.07	20.85	-21.10	20.29
	10 (CCW)	-21.05	-21.02	9.95	-11.06	-21.09	10.11	0.02	-21.06	10.00	11.14	-21.06	9.76	21.30	-21.19	10.02
	0	-20.72	-21.02	-0.24	-10.45	-21.08	0.16	0.22	-21.24	0.41	11.02	-21.06	0.07	20.91	-21.04	0.08
	-10 (CW)	-21.08	-21.04	-10.20	-10.87	-21.11	-10.16	0.17	-21.11	-10.32	11.15	-21.15	-9.75	20.95	-21.08	-10.04
	-20 (CW)	-21.06	-21.23	-20.52	-11.09	-21.19	-20.01	0.04	-21.04	-20.29	11.14	-21.03	-20.59	20.61	-21.06	-20.04

Table 4.14 Horizontal, vertical, and torsional validation result of undistorted lens (bias: -10%, eyeball radius = 13.50 mm)

Degrees	Torsion	-20 (right)			-10 (right)			0			10 (left)			20 (left)		
		Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.	Hor.	Vert.	Tor.
20 (down)	20 (CCW)	-23.09	23.04	20.34	-12.44	22.03	20.37	0.57	22.59	19.87	12.38	22.43	20.22	22.79	22.30	20.20
	10 (CCW)	-23.11	22.46	10.27	-12.36	22.69	10.59	0.03	22.40	10.38	12.10	23.07	10.35	22.81	22.43	10.34
	0	-22.62	22.12	-0.11	-12.44	22.04	0.34	-0.03	22.40	0.03	12.22	22.46	-0.09	22.27	22.02	0.01
	-10 (CW)	-22.05	22.14	-10.58	-13.08	22.45	-10.19	-0.60	22.67	-10.09	12.09	22.06	-9.92	22.94	23.03	-10.20
	-20 (CW)	-22.30	22.09	-20.66	-13.14	22.01	-20.05	-0.71	22.42	-20.31	12.98	22.60	-20.12	22.53	22.83	-20.05
10 (down)	20 (CCW)	-22.55	12.86	20.25	-12.26	12.60	20.20	0.52	12.82	20.21	13.11	12.76	20.10	22.33	13.01	20.20
	10 (CCW)	-22.30	13.02	10.27	-12.06	12.37	9.81	0.08	12.23	10.19	13.17	12.13	10.09	23.18	12.92	10.11
	0	-22.31	12.56	0.28	-12.86	12.69	0.11	-0.23	12.36	0.22	13.12	13.19	-0.61	22.63	12.32	-0.06
	-10 (CW)	-22.14	12.33	-10.39	-12.93	12.17	-10.65	-0.39	12.36	-10.08	13.65	12.71	-10.20	22.91	13.12	-10.30
	-20 (CW)	-22.77	12.77	-20.12	-12.48	12.43	-20.10	-0.36	12.42	-20.46	12.84	12.83	-20.00	22.23	12.27	-19.94
0	20 (CCW)	-22.33	0.19	20.09	-12.14	0.25	20.27	0.10	0.06	19.95	13.06	0.50	20.35	22.74	0.39	19.76
	10 (CCW)	-22.88	0.19	9.85	-12.03	-0.30	10.30	0.41	0.07	10.03	12.75	0.14	10.16	22.72	0.17	10.05
	0	-22.27	0.06	0.29	-12.34	-0.14	0.05	0.00	0.00	0.00	13.31	0.30	-0.42	22.76	0.04	-0.01
	-10 (CW)	-22.89	-0.04	-10.34	-12.06	-0.56	-10.32	0.22	-0.19	-10.27	12.24	0.10	-10.80	22.68	0.09	-10.09
	-20 (CW)	-22.43	-0.28	-20.18	-12.77	-0.11	-20.03	-0.03	-0.06	-20.02	12.76	-0.43	-20.10	22.52	-0.57	-20.45
-10 (up)	20 (CCW)	-22.09	-12.32	20.59	-12.51	-12.73	19.99	0.45	-12.19	20.39	12.99	-12.44	20.39	22.79	-12.92	19.73
	10 (CCW)	-22.48	-12.54	9.91	-12.20	-12.64	9.98	0.50	-13.02	9.82	12.66	-12.14	10.26	22.83	-12.69	10.31
	0	-22.80	-12.57	0.35	-11.97	-13.08	0.16	-0.22	-12.93	0.04	12.33	-12.48	0.29	22.81	-12.41	0.10
	-10 (CW)	-22.06	-12.07	-9.86	-12.04	-12.34	-10.15	-0.59	-12.86	-10.29	12.15	-12.32	-10.30	22.22	-12.56	-10.11
	-20 (CW)	-22.22	-12.08	-20.12	-12.42	-12.50	-19.77	-0.21	-12.43	-20.58	12.46	-12.03	-19.96	22.15	-12.31	-19.83
-20 (up)	20 (CCW)	-22.25	-22.59	20.49	-12.34	-22.36	20.40	0.95	-22.80	19.90	12.41	-22.60	20.71	22.62	-22.31	19.87
	10 (CCW)	-23.20	-22.01	10.40	-12.92	-22.84	10.39	0.29	-22.99	10.03	12.17	-22.97	9.91	22.33	-22.03	9.61
	0	-22.09	-22.30	-0.20	-12.47	-22.22	0.30	0.09	-22.73	0.01	12.57	-22.13	0.12	22.01	-22.18	0.45
	-10 (CW)	-22.15	-22.98	-10.09	-12.43	-22.15	-10.04	-0.16	-22.76	-10.14	12.25	-23.13	-10.16	22.03	-22.23	-10.60
	-20 (CW)	-22.96	-22.31	-19.97	-12.09	-22.14	-20.00	-0.09	-22.35	-20.08	12.08	-22.08	-19.90	22.58	-22.18	-20.38

4.2 Computational Time

The computational time was observed to verify whether the eye movements tracking and visualization system could be used in real-time measurement. Computational time of single frame processing was measured by computing average processing time from 1500 frames. The experiment was performed using personal software and hardware with specifications as follows:

1. Intel® 2 Quad CPU Q6600 @ 2.4 GHz
2. Windows XP® Service Pack 3
3. RAM 2.00 GB
4. EasyCap® Analog to Digital Converter
5. Image Resolution 320 x 240

Table 4.15 shows processing time for individual steps. The total response time to process single video frame is 40.38ms. Thus, the system can be used for real-time processing with average sampling rate 25 frame per second (fps) which is useful for clinicians in the diagnosis of several vestibular disorders.

Table 4.15 Processing time for single video frame

Processing Step	Duration (ms)
Frame Grabbing	4.52
Lens Undistortion	10.32
Pupil Tracking	8.88
Iris Striation Tracking	16.06
3D Extraction, Angular Measurement, & Visualization	0.6
Total Time	**40.38**

4.3 Practical Implementation on Human

4.3.1 Voluntary Eye Movements

In this experiment, 10 healthy participants without any history of vertigo were involved. The participants consisted of 6 males and 4 females with various ages and physical parameters of the eye. The ages of the subject were ranging from 24 to 45 years old with average value 29.7 ± 7.6 years old. The minimum and maximum measured iris diameters were 11mm and 13mm with average value 11.08 ± 0.7mm. The minimum and maximum measured eyeball diameters were 24.5mm and 32mm with average value 28.6 ± 1.9mm. The data of participants in this experiment is shown in Table 4.16.

The smooth pursuit and fixation tests were performed on the right eye in vertical and horizontal directions. The results are presented in Figure 4.4 – Figure 4.11. The visual stimulus was denoted as dashed lines while the eye movement was denoted as solid lines. Figure 4.4 – Figure 4.7 show the result of vertical and horizontal smooth pursuit tests. As expected before, the participants were able to follow the target as the target moved up-down in vertical direction (10° amplitude and period 2s) and right-left in horizontal direction (20° amplitude and period 2s). Some errors occurred due to eye blinking such as shown in Figure 4.6 (subject # 5, frame 770-780). Figure 4.8 – Figure 4.11 show the result of vertical and horizontal fixation tests. The experimental data shows that it was difficult for the participants to follow a fast changing target in fixation. A delay about 4-8 frames (160-320 ms) occurred between the visual stimulus and the eye movement.

Table 4.16 Data of participants in voluntary eye movements experiment

Subject	Sex (M/F)	Age (years)	Iris Diameter (mm)	Eyeball Diameter (mm)
1st	M	24	13	29
2nd	F	26	12.5	28.5
3rd	F	25	13	29
4th	F	27	11.5	27.5
5th	M	30	12	27.5
6th	M	24	12	31
7th	F	26	11	28
8th	M	26	11	32
9th	M	45	11	29
10th	M	44	11.5	24.5

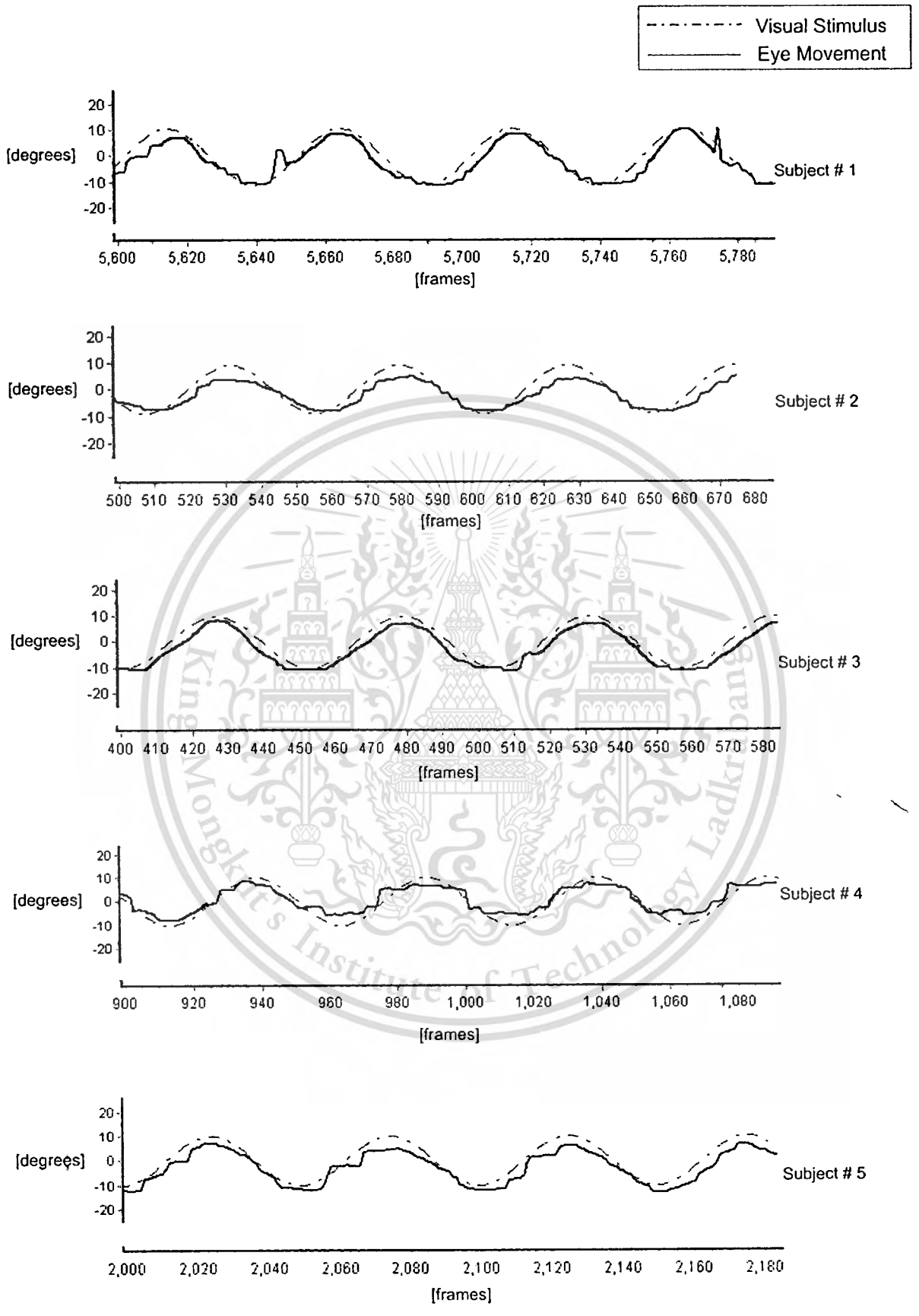


Figure 4.4 Result of vertical smooth pursuit test for subject # 1-5

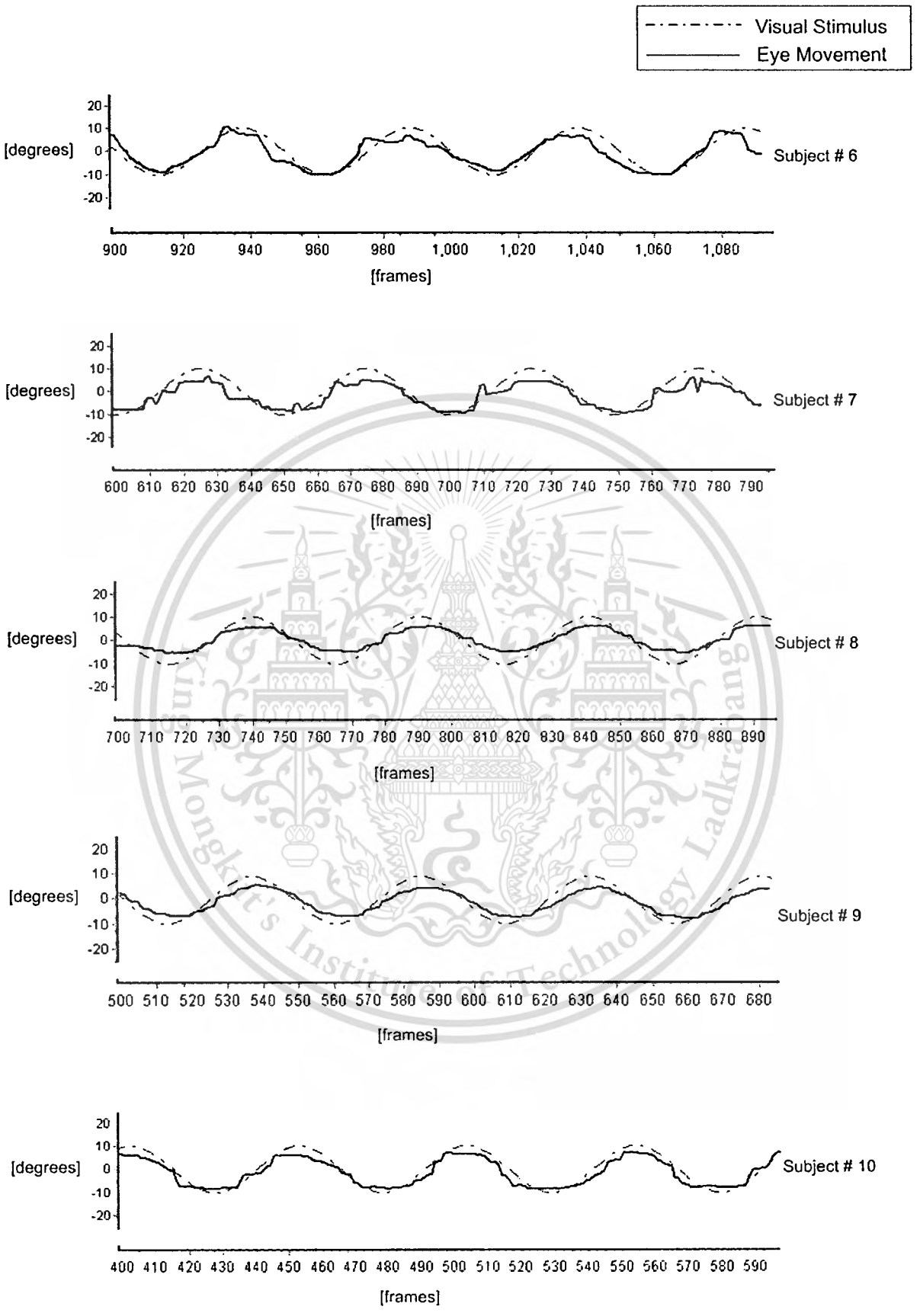


Figure 4.5 Result of vertical smooth pursuit test for subject # 6-10

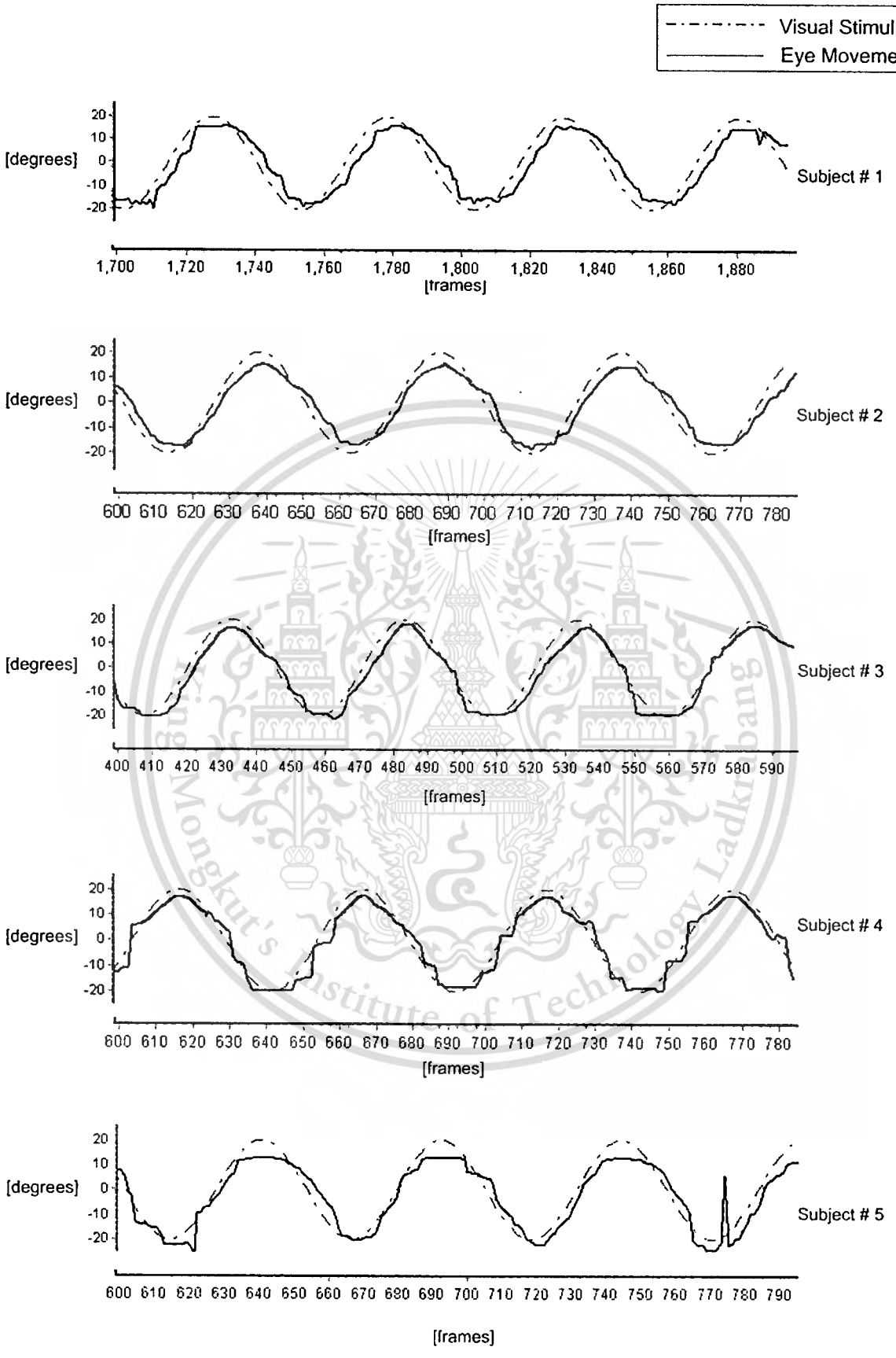


Figure 4.6 Result of horizontal smooth pursuit test for subject # 1-5

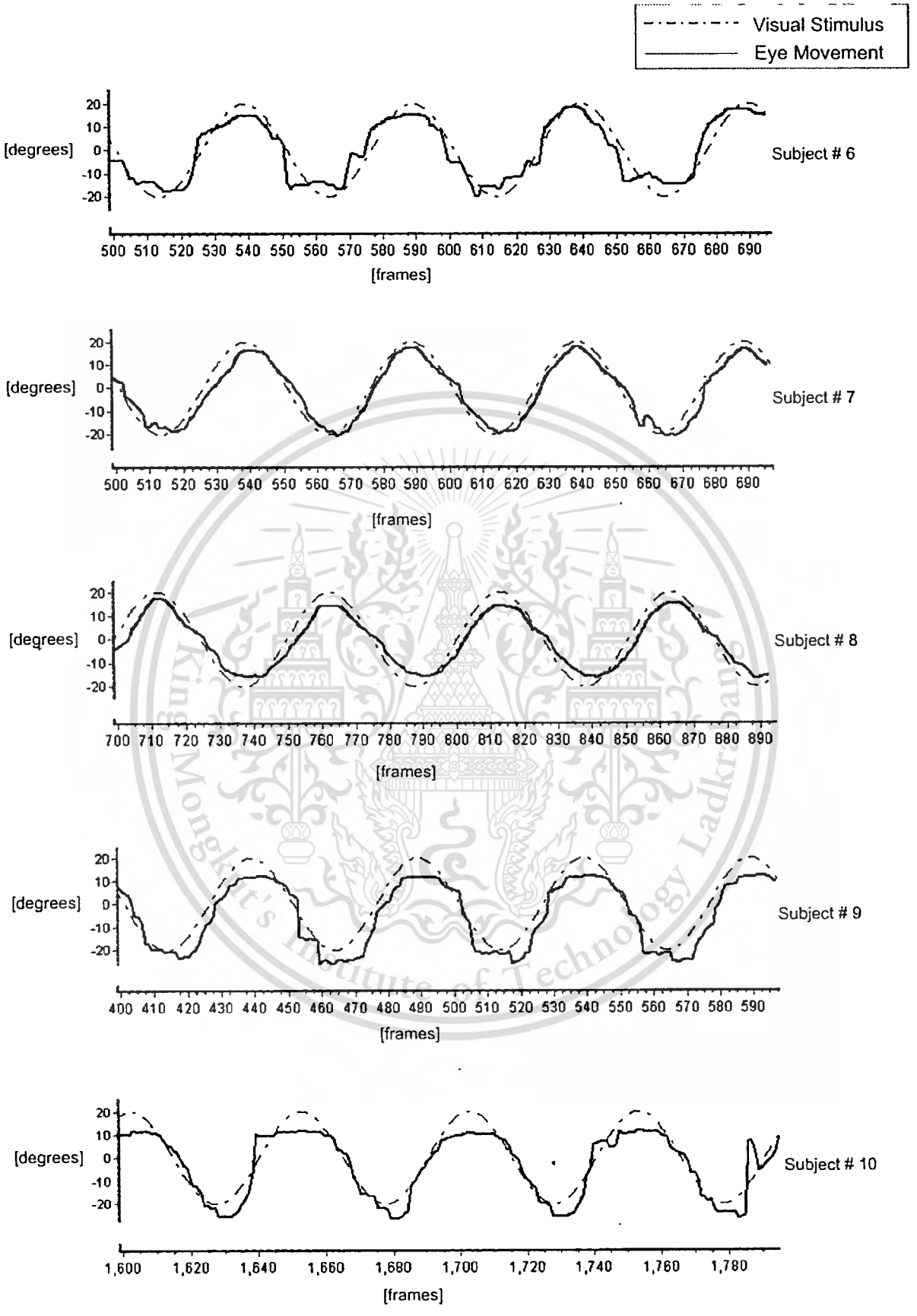


Figure 4.7 Result of horizontal smooth pursuit test for subject # 6-10

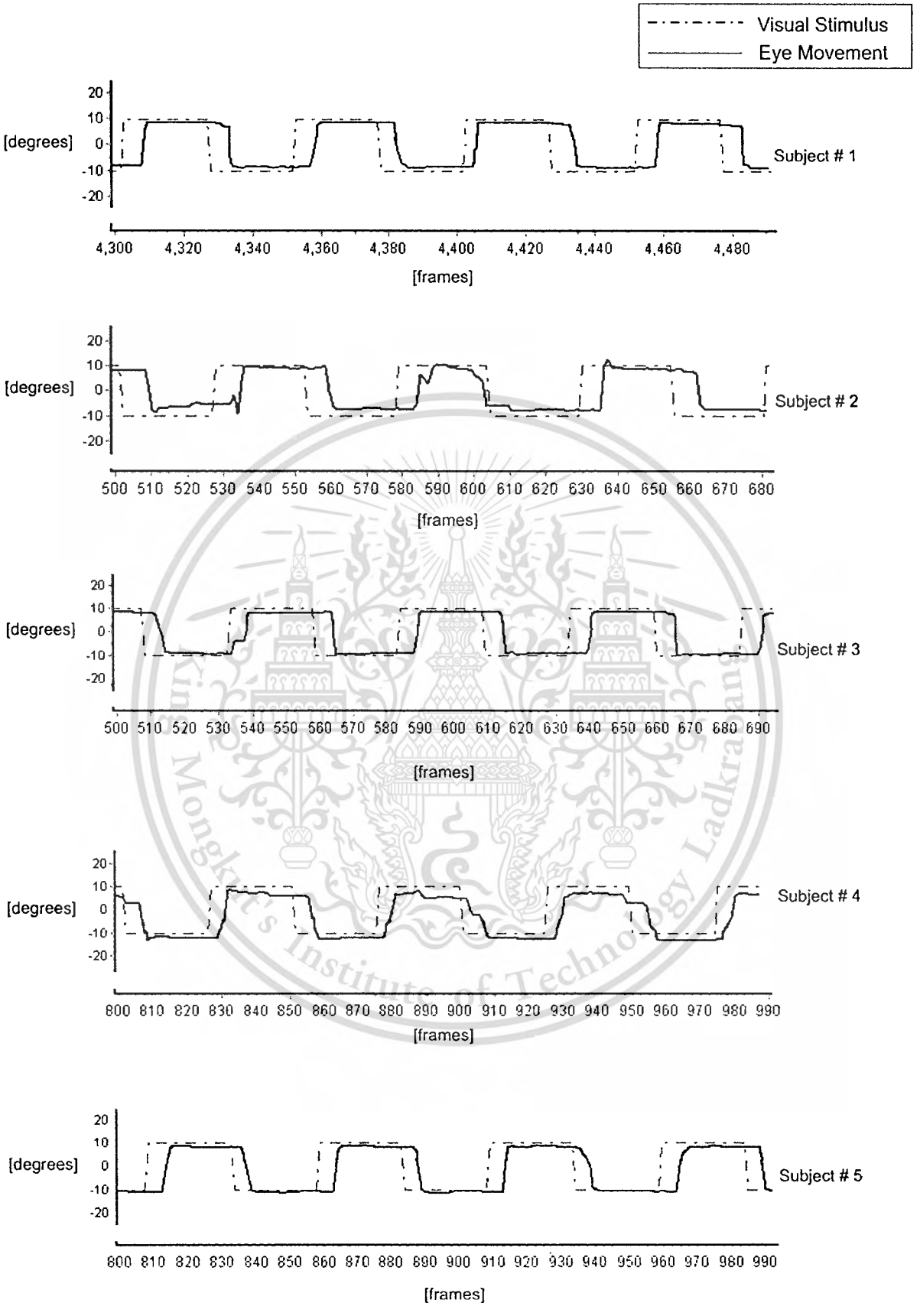


Figure 4.8 Result of vertical fixation test for subject # 1-5

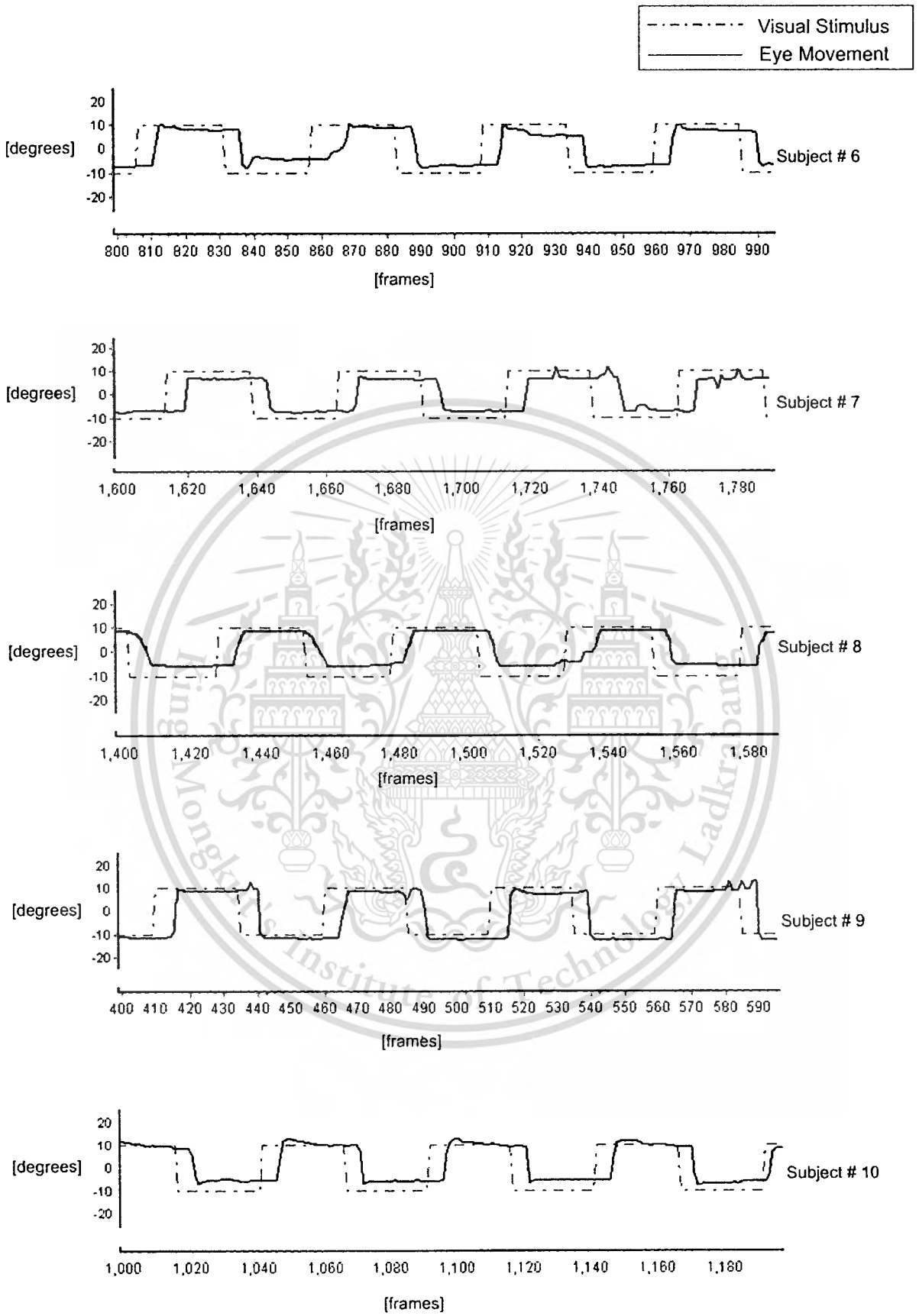


Figure 4.9 Result of vertical fixation test for subject # 6-10

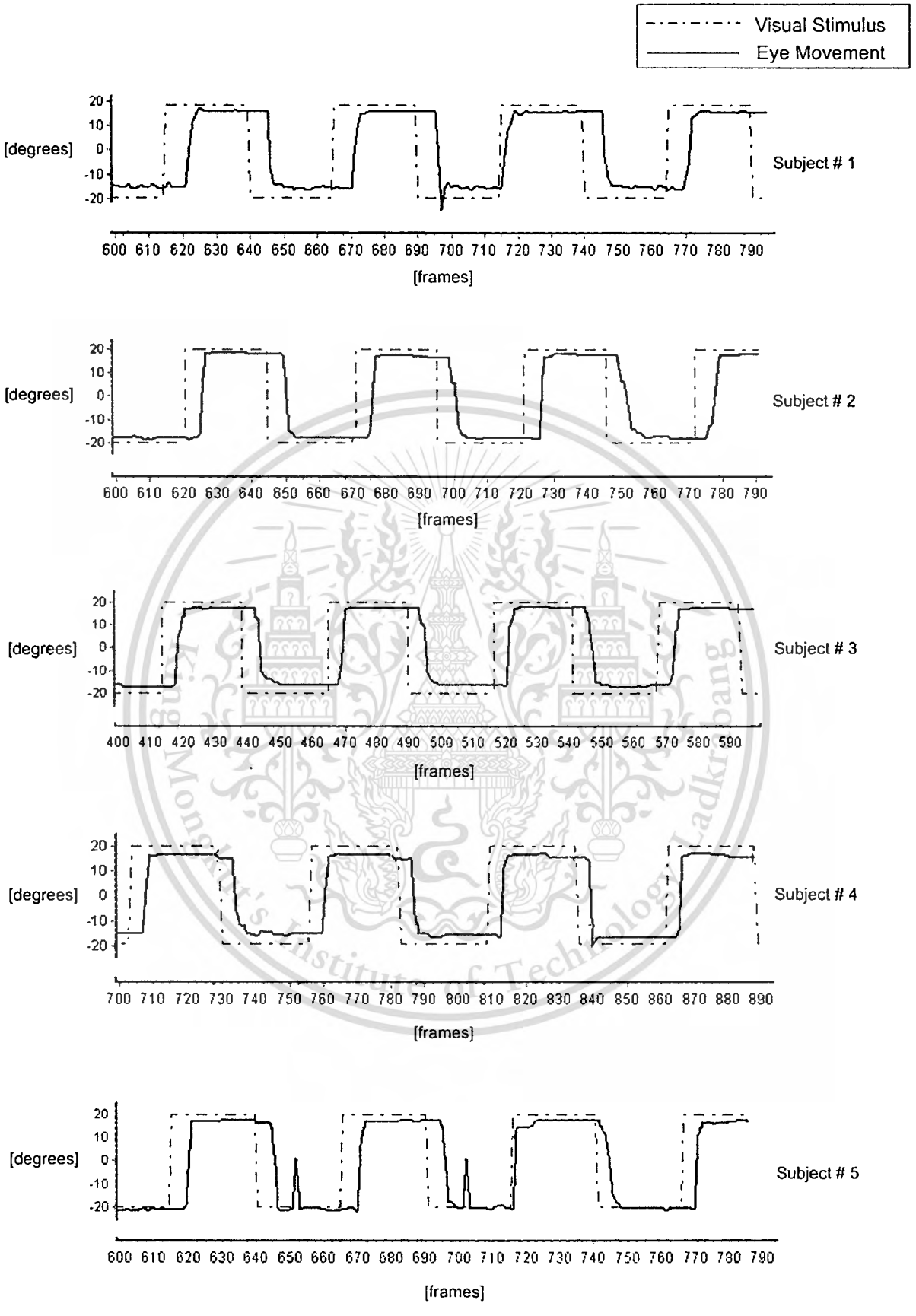


Figure 4.10 Result of horizontal fixation test for subject # 1-5

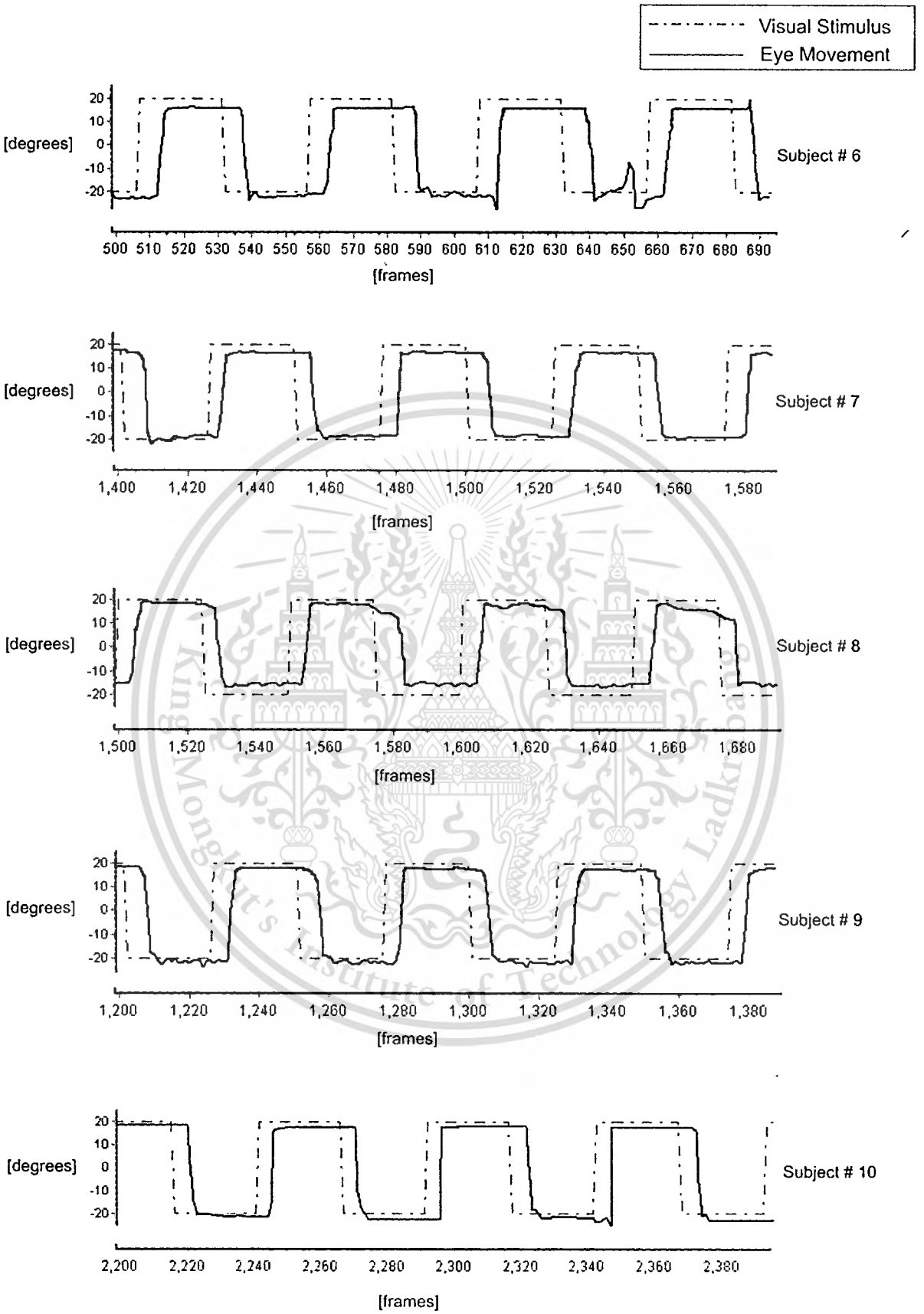


Figure 4.11 Result of horizontal fixation test for subject # 6-10

4.3.2 Involuntary Eye Movements

The eye movements tracking and visualization system was used to observe involuntary eye movements. In this case, the system was used to detect Benign Paroxysmal Positional Vertigo (BPPV) symptom. The experiment was held at Outpatient Department (OPD) Ear Nose Throat, Srinakarinviroj Hospital, Pathumthani, Thailand. A 56 years old female patient with 13.5mm eyeball radius and 12mm iris diameter was participated in this experiment. The patient had suffered BPPV since 2009. At the time experiment was performed, the patient still underwent medical therapy. The patient was asked to wear the goggle and the right eye was observed. The test was performed on complete darkness by installing the front side cover on the goggle. Dix-Hallpike test was then performed by bringing the patient from sitting to a supine position, with the head turned 45 degrees to one side and extended about 20 degrees backward. Once supine, the right eye was observed using eye movements tracking and visualization system for about 30 seconds.

The positional nystagmus was clearly noticed in torsional eye movement when the head was positioned at certain position. The positional nystagmus was caused by misplaced of calcium carbonate particles called "ear rocks" inside the inner ear. Moving the head in certain directions might cause the ear rocks to tug on hairlike sensors, triggering a type of dizziness called BPPV. The positional nystagmus occurred during frames 1100-1280 with total occurrence time about 10s as shown in Figure 4.12. After finishing the Dix-Hallpike test, the patient was brought to initial sitting position. The eye moved in normal horizontal, vertical, and torsional movements as shown in Figure 4.13. The experimental results show that the eye movements tracking and visualization system was able to detect involuntary eye movements, particularly positional nystagmus, in a clinical routine.

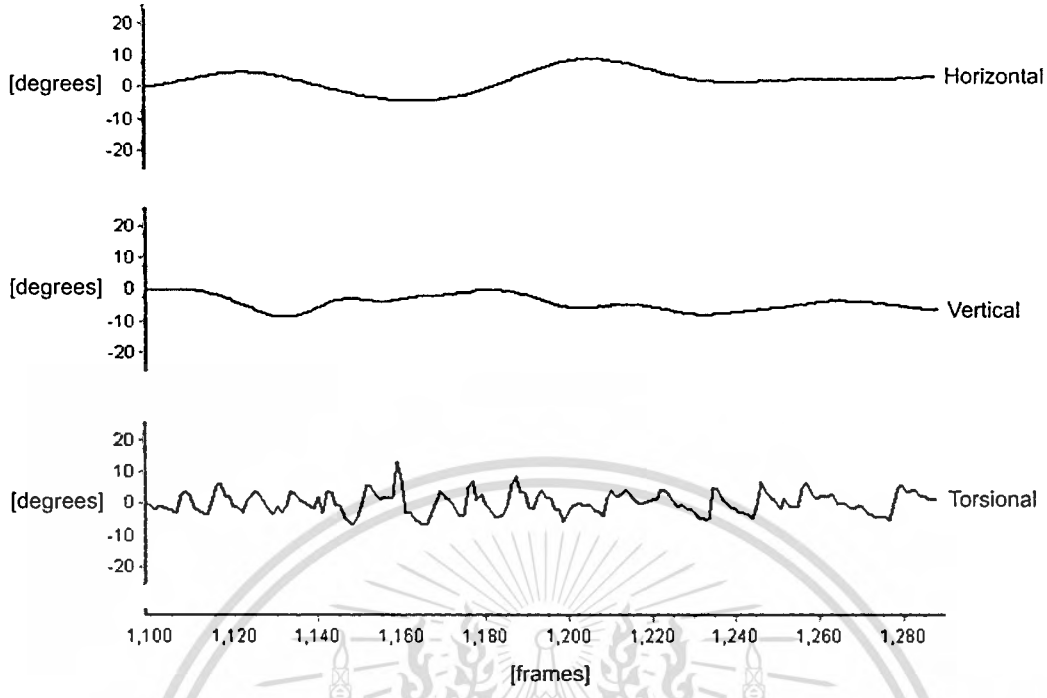


Figure 4.12 Positional Vertigo (BPPV) during frames 1100-1280

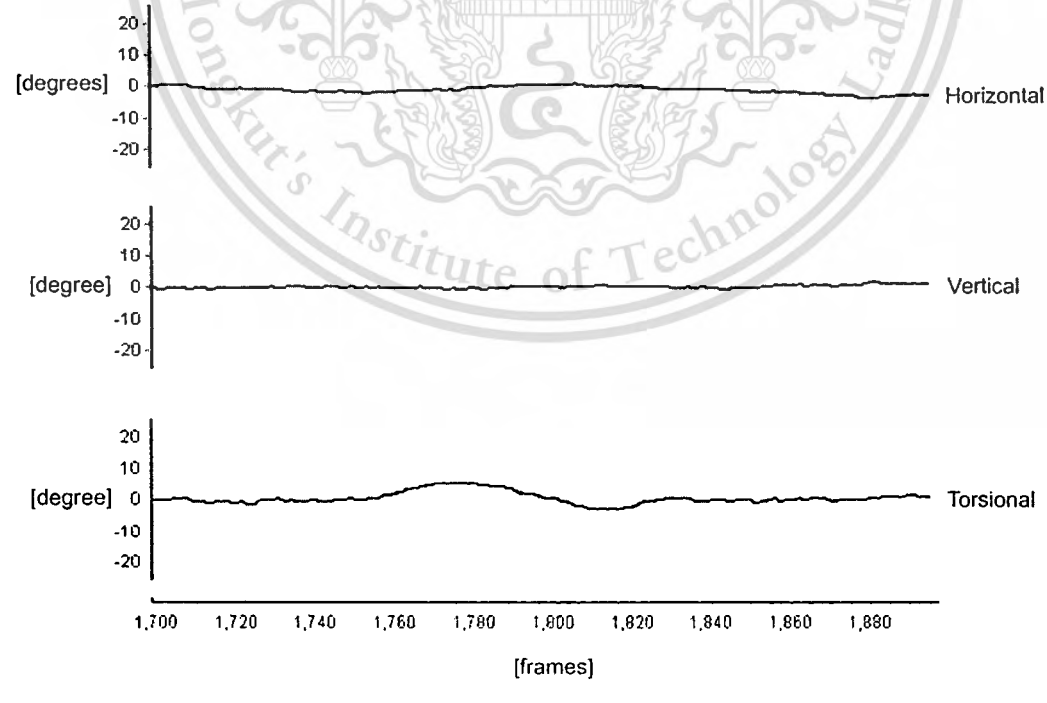


Figure 4.13 Normal eye movements during frames 1700-1880

Chapter 5

Conclusion and Future Work

5.1 Discussion

In this thesis, an inexpensive real-time 3D eye movements tracking and visualization system is proposed. The system utilizes dual cameras mounted on a consumer-grade welding goggle to track an eye. Initially, a camera calibration procedure must be performed to obtain camera parameters and lens distortion coefficients. Next, eyeball radius is obtained by direct measurement. Iris striation is gained from captured eye image. Center of mass algorithm is employed to gain 2D coordinates of pupil center. Template matching algorithm based on correlation coefficient is applied to obtain 2D coordinates of iris striation center. The 3D coordinates of pupil and iris striation centers are extracted using Direct Linear Transformation (DLT) algorithm. Furthermore, lens undistortion algorithm is added to improve the eye movements tracking result. The horizontal, vertical, and torsional angular positions of the eye are estimated from the results of 3D coordinates extraction. Finally, a real-time 3D visualization based on tracking result is performed to help clinicians understanding the characteristics of eye movements.

In order to observe the accuracy of the system, *in vitro* validation method using artificial eyeball and gimbal system was performed. The artificial eyeball was rotated in horizontal and vertical angles with excursion ranging $\pm 20^{\circ}$ at steps of 10°. At each point, five different angular torsional positions were made, resulting 125 unique positions consisted of horizontal-vertical-torsional (HVT) values. The computational time of the system was also observed to test whether the system can be used in real-time implementation. Furthermore, the system was implemented on human to detect voluntary and involuntary eye movements.

The *in vitro* experimental results show that it was not enough to implement only DLT algorithm to obtain precise horizontal, vertical, and torsional angular eye positions. Errors resulted from lens aberration could be reduced by adding lens undistortion algorithm. The lens undistortion algorithm successfully improved the accuracy of the system by decreasing the error of tracking in horizontal, vertical, and torsional angular positions about 53.13%, 63.16%, and 57.45%, respectively. Most improvement occurred particularly at the eccentric positions since the rays farther from the center of the lens were bent more or less than that those closer from as illustrated in Figure 3.29.

Experimental results of single camera method in [5] were used as comparators to verify the effectiveness of DLT and lens undistortion algorithms. The comparison results show that our system exhibited better performance in horizontal, vertical, and torsional angular positions compared to single camera method. Tracking error caused by inaccuracy in measuring physical parameters of the eye was also probed by underestimating and overestimating the eye radius with bias factors up to ±10%. Over ±10% bias in measuring eye radius, errors in horizontal and vertical angular positions reached about 2°, while error in torsional angular position was considerably more robust as it was preserved below 0.4°. The unrobustness of errors in horizontal and vertical angular positions was due to the usage of eye radius parameter in horizontal and vertical measurements as shown in Eq. (3.85) and Eq. (3.89).

Average processing time for single frame of 1500 successive frames was found to be 40.38ms. Table 4.15 shows that the iris striation tracking took the longest time compared to the other processes since iris striation tracking involves user-defined memory allocation, memory writing, and memory releasing. The iris striation tracking is initialized by determining the ROI based on the coordinates of pupil center. In practical implementation, the algorithm allocates the memory to be used by temporary image. Then, image area inside the ROI is copied into the memory and saves as a temporary image. Next, the template matching algorithm is performed to find the corresponding location of the template in the temporary image. The resulted 2D coordinates of iris striation center with respect to initial image are computed and saved. Next, the algorithm deletes the temporary image from the memory. Finally, the allocated memory is released as a free memory.

The system was used to detect voluntary eye movements stimulated by a small moving target at the monitor from 10 healthy participants. The experimental results demonstrate that a delay about 160-320 ms occurred between visual stimulus and eye movements in fixation test instead of smooth pursuit test. In fixation eye movement, the position of the target with respect to the fovea was estimated during this delay to decide how far the eye should move. Then, the difference between the initial and intended eye positions was converted into a neural command that stimulates the extraocular muscles to move the eyes in appropriate distance and direction [84]. Additionally, the system was also used to detect involuntary eye movements called positional nystagmus caused by BPPV. The positional nystagmus was clearly detected about 10s in torsional eye movement when the head was positioned at certain position. Thus, this system is appropriate to use as a real-time investigation tool in clinical routine.

5.2 Conclusion

The work presented in this thesis explains preliminary design and implementation of inexpensive real-time 3D eye movements tracking and visualization system using dual cameras acquisition. The 3D visualization of the eye is performed based on the real-time tracking results. The DLT and lens undistortion algorithms are proven reducing errors in measuring horizontal, vertical, and torsional movements. Inaccurately measuring physical parameters of the eye affects horizontal and vertical angular positions while torsional angular position is considered more robust to radius estimation error. Real-time implementation shows that the system can be used in clinical routine to detect either voluntary or involuntary eye movements.

5.3 Future Work

Even though our system exhibited a good performance, actual radius of human eye could not be obtained exactly since it was impossible to directly determine the center of the eye rotation. Instead, eyeball diameter of the human was determined by carefully measure the distance of outer part (lateral canthi) and inner part (medial canthi) of eyelid aperture (palpebral aperture) as explained in chapter 3. Regardless of how rigorously this measurement is performed, the real radius of human eye may be slightly different from the measurement result since the rotational axes of the eye may not precisely pass through the center of the eye. In the future, it may be possible to use ultrasound to improve accuracy in measuring physical parameters of the eye as explained in [85].

Another work needed to be explored is design consideration for wider clinical applications. Since there is limitation in maximum sampling rate about 25fps, this system may not appropriate to be used as measurement device for fast eye movements (i.e. saccadic eye movements) which require a video system with sampling rate more than 100fps. This limitation can be overcome by changing the low-cost CCD camera to the high-speed industrial CCD camera with average sampling rate 120fps [39]. Recent advances in data communication technology can be used to improve the system. In the future, wireless transmitters and receivers may be possible to compensate data cables in the camera.

References

[1] K. Rayner, **"Eye Movements in Reading and Information Processing: 20 Years of Research,"** *Psychological Bulletin,* vol. 3, pp. 372-422, 1998.

[2] R. E. Gans, **"Video-oculography: A new diagnostic technology for vestibular patients,"** *The Hearing Journal,* vol. 54, pp. 40-42, 2001.

[3] R. S. Remmel, **"An Inexpensive Eye Movement Monitor Using the Scleral Search Coil Technique,"** *IEEE Transactions on Biomedical Engineering,* vol. BME-31, pp. 388-390, 1984.

[4] J. Kim, **"A simple pupil-independent method for recording eye movements in rodents using video,"** *Journal of Neuroscience Methods,* vol. 138, pp. 165-171, 2004.

[5] 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, pp. 141-150, 2005.

[6] T. Dera, G. Boning, S. Bardins, and E. Schneider, **"Low-latency video tracking of horizontal, vertical, and torsional eye movements as a basis for 3dof realtime motion control of a head-mounted camera,"** in *Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference*, Taipei, Taiwan, 2006, pp. 5191-5196.

[7] J. S. Stahl, A. M. van Alphen, and C. I. De Zeeuw, **"A comparison of video and magnetic search coil recordings of mouse eye movements,"** *Journal of Neuroscience Methods,* vol. 99, pp. 101-110, 2000.

[8] T. Haslwanter, **"Mathematics of three-dimensional eye rotations,"** *Vision Research,* vol. 35, pp. 1727-1739, 1995.

[9] E. Suaste, P. Rivera, J. Leybon, J. Avila, V. Salazar, L. Leija, and H. Sossa, **"3D Visualization, Simulation, Animation, and Modeling of Eye-Pupil Movements and Visual Field for Clinical Applications,"** in *18th Annual Conference of IEEE Engineering in Medicine and Biology Society, Amsterdam*, 1996, pp. 2277-2278.

[10] M. A. Schill, C. Wagner, M. Hennen, H.-J. Bender, and R. Manner, **"EyeSi - A Simulator for Intra-ocular Surgery,"** in *Medical Image Computing and Computer-Assisted Intervention – MICCAI'99*. vol. Volume 1679/1999, ed.: Springer Berlin / Heidelberg, 1999, pp. 1166-1174.

[11] K. H. Wassill and T. Kowarsch, "**Virtual Strabismus: Movements of the Eyes simulated in 3D for Educational Purposes,**" presented at the Proceedings of the 16th European Simulation Multiconference on Modelling and Simulation, 2002.

[12] R. P. Crick and P. T. Khaw, "**Practical Anatomy and Physiology of the Eye and Orbit,**" in *A Textbook of Clinical Ophthalmology : A Practical Guide to Disorders of the Eyes and Their Management*, 3rd ed.: World Scientific Publishing Co.Pte.Ltd, 2003, pp. 5 - 42.

[13] A. M. F. Wong, **Eye Movement Disorders**, 1st ed. New York: Oxford University Press, 2007.

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

[15] M. Günther, "**Computersimulationen zur Synthetisierung des muskulär erzeugten menschlichen Gehens unter Verwendung eines biomechanischen Mehrkörpermodells,**" PhD Thesis, Natural Sciences, Physik, Eberhard-Karls-Universität Tübingen, 1997.

[16] R. S. Leigh and D. S. Zee, **The Neurology of Eye Movements**, 3rd ed. New York: Oxford University Press, 1999.

[17] K. Naoi, K. Nakamae, H. Fujioka, T. Imai, K. Sekine, N. Takeda, and T. Kubo, "**Three-Dimensional Eye Movement Simulator Extracting Instantaneous Eye Movement Rotation Axes, the Plane Formend by Rotation Axes, and Innervations for Eye Muscles,**" *IEICE Transactions on Information and Systems,* vol. E86-D, pp. 2452-2462, 13 May 2003 2003.

[18] E. E. Mhoon, L. P. Bernstein, and V. L. Towle, "**Saccular Influence on the Otolith-Spinal Reflex and Posture During Sudden Falls of the Cat,**" *Otology & Neurotology,* vol. 18, pp. 86-92, 1997.

[19] M. Brown, M. Marmor, Vaegan, E. Zrenner, M. Brigell, and M. Bach, "**ISCEV Standard for Clinical Electro-oculography (EOG) 2006,**" *Documenta Ophtalmologica,* vol. 113, pp. 205-212, 2006.

[20] J. Dewar, "**The Physiological Action of Light,**" *Nature,* vol. 15, pp. 433-435, 1877.

[21] E. Schott, "**Uber die Registrierung des Nystagmus und anderer Augenbewegungen vermittels des Saitengalvanometers,**" *Deutsches Archiv fuer Klinische Medizin,* vol. 140, pp. 79-90, 1922.

[22] I. L. Meyers, "**Electronystagnography: A Graphic Study of the Action Currents in Nystagmus,**" *Archives of Neurology and Psychiatry,* vol. 21, pp. 901-918, 1929.

[23] 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.

[24] D. A. 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.

[25] H. Collewijn, F. van der Mark, and T. C. Jansen, "**Precise recording of human eye movements,**" *Vision Research,* vol. 15, pp. 447-450, IN5, 1975.

[26] H. Collewijn, J. V. d. Steen, L. Ferman, and T. C. Jansen, "**Human ocular counterroll: assessment of static and dynamic properties from electromagnetic scleral coil recordings,**" *Experimental Brain Research,* vol. 59, pp. 1432-1106, 1985.

[27] 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, pp. 983-987, 989, 1982.

[28] E. L. Irving, J. E. Zacher, R. S. Allison, and M. G. Callander, "**Effects of Scleral Search Coil Wear on Visual Function,**" *Investigative Ophthalmology & Visual Science,* vol. 44, pp. 1933-1938, 2003.

[29] M. Hatamian and D. J. Anderson, "**Design Considerations for a Real-Time Ocular Counterroll Instrument,**" *IEEE Transactions on Biomedical Engineering,* vol. BME-30, pp. 278-288, 1983.

[30] T. Viéville and D. Masse, "**Ocular counter-rolling during active head tilting in humans,**" *Acta Otolaryngology,* vol. 103, pp. 280-290, 1987.

[31] S. T. Moore, I. S. Curthoys, and S. G. McCoy, "**VTM -- an image-processing system for measuring ocular torsion,**" *Computer Methods and Programs in Biomedicine,* vol. 35, pp. 219-230, 1991.

[32] S. T. Moore, I. S. Curthoys, and T. Haslwanter, "**Potential clinical applications of video-based eye position measurement,**" in *Engineering in Medicine and Biology Society, 1995., IEEE 17th Annual Conference*, 1995, pp. 1627-1628 vol.2.

[33] 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, pp. 28-36, January, 2985 1985.

[34] 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, pp. 338–341, March 1990.

[35] D. Ott, F. Gehle, and R. Eckmiller, "**Video-oculographic measurement of 3-dimensional eye rotations**," *Journal of Neuroscience Methods,* vol. 35, pp. 229-234, 1990.

[36] J. E. Bos and B. de Graaf, "**Ocular torsion quantification with video images**," *IEEE Transactions on Biomedical Engineering,* vol. 41, pp. 351-357, 1994.

[37] 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, pp. 445-459, 1996.

[38] S. T. Moore, T. Haslwanter, I. S. Curthoys, and S. T. Smith, "**Measurement of three dimensional eye position using image processing: a geometric approach**," in *Proceedings of IEEE International Conference in Image Processing 1994* Austin, Texas, 1994, pp. 436-440.

[39] Y. Toshiaki, K. Yasuo, A. Mio, K. Maki, and S. Kazuki, "**Three-dimensional analysis of eye movements using four times high-speed video camera**," *Auris, nasus, larynx,* vol. 32, pp. 107-112, 2005.

[40] S. C. Kim, K. C. Nam, W. S. Lee, and D. W. Kim, "**A new method for accurate and fast measurement of 3D eye movements**," *Medical Engineering & Physics,* vol. 28, pp. 82-89, 2006.

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

[42] Khronos. (2010, 22 January). **OpenGL - The Industry Standard for High Performance Graphics**. Available: http://www.opengl.org/

[43] EuroPC. (2010, 4 February). **Acer Aspire M3641 Intel Core 2 Quad Q6600**. Available: http://www.europc.co.uk/product.php?productid=130343&utm_source=googlebase&utm_medium=cpc&utm_campaign=europc

[44] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, "**Digital Image Representation,**" in *Digital Image Processing using MATLAB*, 1st ed. New Jersey: Prentice Hall, 2004, pp. 12-14.

[45] M. Sonka, V. Hlavac, and R. Boyle, "**Mathematical Morphology,**" in *Image Processing, Analysis, and Machine Vision*, H. Gowans, Ed., 3rd ed. Toronto: Thomson, 2008, pp. 657-665.

[46] M. Sonka, V. Hlavac, and R. Boyle, "**Image Pre-Processing,**" in *Image Processing, Analysis, and Machine Vision*, H. Gowans, Ed., 3rd ed. Toronto: Thomson, 2008, pp. 138-142.

[47] S. Chernenko. (2008, 10 February 2010). **Gaussian Filter.** Available: http://www.librow.com/articles/article-9

[48] G. Bradski and A. Kaehler, "**Image Processing,**" in *Learning OpenCV: Computer Vision with The OpenCV Library*, 1st ed. California: O'Reilly, 2008, pp. 109-114.

[49] R. C. Gonzalez and R. E. Woods, "**Representation and Description,**" in *Digital Image Processing*, 3rd ed. New Jersey: Prentice Hall, 2007, pp. 795-798.

[50] S. Suzuki and K. be, "**Topological structural analysis of digitized binary images by border following,**" *Computer Vision, Graphics, and Image Processing,* vol. 30, pp. 32-46, 1985.

[51] A. Rosenfeld, "**Connectivity in digital picture,**" *Journal Association for Computing Machinery,* vol. 17, pp. 146-160, 1970.

[52] T. H. Morrin and Il, "**Chain-link compression of arbitrary black-white image,**" *Computer Graphics and Image Processing,* vol. 5, pp. 172-189, 1976.

[53] R. J. Prokop and A. P. Reeves, "**A survey of moment-based techniques for unoccluded object representation and recognition,**" *CVGIP: Graphical Models and Image Processing,* vol. 54, pp. 438-460, 1992.

[54] G. Bradski and A. Kaehler, "**Moments,**" in *Learning OpenCV: Computer Vision with The OpenCV Library*, 1st ed. California: O'Reilly, 2008, pp. 252-256.

[55] R. C. Gonzalez and R. E. Woods, "**Object Recognition,**" in *Digital Image Processing*, 2nd ed. New Jersey: Prentice Hall, 2007, pp. 701-703.

[56] G. Bradski and A. Kaehler, "**Histogram and Matching,**" in *Learning OpenCV: Computer Vision with The OpenCV Library*, 1st ed. California: O'Reilly, 2008, pp. 214-218.

[57] W. K. Pratt, "**Geometrical Image Modification,**" in *Digital Image Processing*, 4th ed. Danver, MA: Wiley-Interscience, 2007, pp. 386-393.

[58] G. Bradski and A. Kaehler, "**Camera Models and Calibration,**" in *Learning OpenCV: Computer Vision with The OpenCV Library*, 1st ed. California: O'Reilly, 2008, pp. 370-375.

[59] L. G. Roberts, "**Machine Perception of Three-Dimensional Solids,**" in *Optical and Electro-Optical Information Processing*, J. T. Tippett, Ed., ed. Cambridge, MA: MIT Press, 1965.

[60] R. Hartley and A. Zisserman, "**Camera Models,**" in *Multiple View Geometry for Computer Vision*, 2nd ed. Cambridge: Cambridge University Press, 2003, p. 164.

[61] K. S. Fu, R. C. Gonzalez, and C. S. G. Lee, **Robotics: Control, Sensing, Vision, and Intelligence.** New York: McGraw-Hill, 1987.

[62] M. Sonka, V. Hlavac, and R. Boyle, "**Calibration of one camera from a known scene,**" in *Image Processing, Analysis, and Machine Vision*, 2nd ed. Toronto: Thomson-Engineering, 1998, pp. 455-457.

[63] O. D. Faugeras, **Three-Dimensional Computer Vision: A Geometric Viewpoint.** Cambridge, MA: MIT Press, 1993.

[64] H. Anton and R. C. Busby, "**Singular Value Decomposition,**" in *Contemporary Linear Algebra*, 1st ed. Hoboken, NJ: John Wiley & Sons, Inc., 2003, pp. 509-516.

[65] H. Anton and R. C. Busby, "**QR-Decomposition: Householder Transformations,**" in *Contemporary Linear Algebra*, 1st ed. Hoboken, NJ: John Wiley & Sons, Inc., 2003, pp. 417-428.

[66] Z. Zhang, "**A Flexible New Technique for Camera Calibration,**" *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, pp. 1330-1334, 2000.

[67] V. Vezhnevets and A. Velizhev. (2005, 10 February 2010). **GML C++ Camera Calibration Toolbox.** Available: http://research.graphicon.ru/calibration/gml-c++-camera-calibration-toolbox.html

[68] D. C. Brown, "**Close-range camera calibration,**" *Photogrammetric Engineering,* vol. 37, pp. 855-866, 1971.

[69] J. Y. Bouguet. (2008, 18 February 2010). **Camera Calibration Toolbox for Matlab.** Available: http://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example.html

[70] S. Thurn, "**Lecture 4: Camera Calibration,**" ed. Palo Alto, CA: Stanford University, 2007, p. 52.

[71] L. Chen, C. W. Armstrong, and D. D. Raftopoulos, **"An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique,"** *Journal of Biomechanics,* vol. 27, pp. 493-500, 1994.

[72] D. G. Everaert, A. J. Spaepen, M. J. Wouters, K. H. Stappaerts, and R. A. B. Oostendorp, **"Measuring small linear displacements with a three-dimensional video motion analysis system: Determining its accuracy and precision,"** *Archives of Physical Medicine and Rehabilitation,* vol. 80, pp. 1082-1089, 1999.

[73] D. Gorpas, K. Politopoulos, and D. Yova, **"A binocular machine vision system for three-dimensional surface measurement of small objects,"** *Computerized Medical Imaging and Graphics,* vol. 31, pp. 625-637, 2007.

[74] A. Ben-Simon, O. Ben-Shahar, and R. Segev, **"Measuring and tracking eye movements of a behaving archer fish by real-time stereo vision,"** *Journal of Neuroscience Methods,* vol. 184, pp. 235-243, 2009.

[75] Y. I. Abdel-Aziz and H. M. Karara, **"Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry,"** in *Proceeding of the Symposium on Close-Range Photogrammetry*, 1971, pp. 1-18.

[76] J. T. Enright, **"Ocular translation and cyclotorsion due to changes in fixation distance,"** *Vision Research,* vol. 20, pp. 595-601, 1980.

[77] E. Wollf, **The anatomy of the eye and orbit**, 2nd ed. London: H. K. Lewis & Co., 1940.

[78] A. D. Straw, **"Vision Egg: an open-source library for realtime visual stimulus generation,"** *Frontiers in Neuroinformatics,* vol. 2, 2008-Nov-04 2008.

[79] H. Z. Tan, R. Gray, C. Spence, C. M. Jones, and R. M. Rosli, **"The Haptic Cuing of Visual Spatial Attention: Evidence of A Spotlight Effect"** in *Human Vision and Electronic Imaging XIV*, San Jose, California, USA, 2009.

[80] T. C. Hain. (2007, 3 March 2010). **Beningn Paroxysmal Positional Vertigo**. Available: http://www.dizziness-and-balance.com/disorders/bppv/bppv.html

[81] M. R. Dix and C. S. Hallpike, **"The Pathology, Symptomatology and Diagnosis of Certain Common Disorders of the Vestibular System,"** *AMA Arch Otolaryngol,* vol. 56, pp. 341-354, December 1, 1952 1952.

[82] M. Rhodes. (2008, 02 March 2010). **Dix-Hallpike Test for Vertigo.** Available: http://www.webmd.com/brain/dix-hallpike-test-for-vertigo-15884

[83] S. G. Diamond and C. H. Markham, **"Ocular Counterrolling as An Indicator of Vestibular Otolith Function,"** *Neurology,* vol. 33, pp. 1460-1469, 1983.

[84] G. J. Augustin, D. M. Chikaraishi, M. D. Ehlers, G. Einstein, D. Fitzpatrick, W. C. Hall, E. Jarvis, L. C. Katz, J. Kauer, A.-S. LaMantia, J. O. McNamara, R. D. Mooney, M. A. L. Nicolelis, D. Purves, P. H. Reinhart, S. A. Simon, J. H. P. Skene, J. Voyvodic, L. E. White, and S. M. Williams, **Neuroscience**, 3rd Edition ed. Sunderland, Massachusetts, USA: Sinauer Associates, Inc., 2004.

[85] S. Blomdahl, **"Ultrasonic Measurements of The Eye in The Newborn Infant,"** *Acta Ophthalmologica,* vol. 57, pp. 1048-1056, 1979.

Biography

Personal Information

Name	SUNU WIBIRAMA
Nationality	Indonesian
Date of birth	October 26, 1985
Place of birth	Montpellier, France

Education

Bachelor degree

Field	Electrical Engineering
Duration	2003-2007
Department	Department of Electrical Engineering
University	Gadjah Mada University, Indonesia

Master degree

Field	Electronic Engineering
Duration	2008-2010
Department	Department of Electronics
University	King Mongkut's Institute of Technology, Ladkrabang (KMITL), Thailand

Research Interests

Software Engineering, Human Computer Interaction, Computer Vision, and Image Processing.

List of International Papers

1- Sunu Wibirama, Visan Mahasitthiwat, Supan Tungjitkusolmun, Chuchart Pintavirooj, and Kasin Vichienchom, **"Eye-Motion Tracking Using Motion Gradient,"** in *The Papers of Technical Meeting on Medical and Biological Engineering* 09-36~70, IEE Japan, March 31, 2009, pp. 121-124.

2- Sunu Wibirama, Supan Tungjitkusolmun, Chuchart Pintavirooj, and Kazuhiko Hamamoto, **"Real Time Eye Tracking using Initial Centroid and Gradient Analysis,"** in *Proceeding of The 2009 6th International Conference of Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology*, Pattaya, Thailand, May 6-9, 2009, pp. 1058-1061.

3- Sunu Wibirama, Supan Tungjitkusolmun, Chuchart Pintavirooj, and Kazuhiko Hamamoto, **"Dual Cameras Acquisition For Three-Dimensional Eye-Motion Tracking,"** in *Proceeding of 2nd Biomedical Engineering International Conference (BMEiCON 2009)*, Thailand, August 13-14, 2009, pp. 126-130.

MBE－09－36～70

電気学会研究会資料

The Papers of Technical Meeting on
Medical and Biological Engineering, IEE Japan

医用・生体工学研究会

MBE－09－36～70

２００９年３月３１日

社団法人 電気学会
The Institute of Electrical Engineers of Japan
東京都千代田区五番町６－２

MBE-09-59

EYE-MOTION TRACKING USING MOTION GRADIENT

S. Wibirama†, V. Mahasitthiwat ‡, S.Tungjitkusolmun†, C. Pintaviroj†, and K.Vichienchom†
† Department of Electronics, Faculty of Engineering
and Research Center for Communications and Technology (ReCCIT)
King Mongkut Institute of Technology Ladkrabang
Bangkok 10520, Thailand
‡ Faculty of Medicine, Srinakarinviroj University
Pathumtani 12121, Thailand
Email : kpchucha@kmitl.ac.th

ABSTRACT

An Eye-motion tracking system for acquisition and estimating of eye movement is introduced in this paper. The system consists of two infrared cameras mounted on the binocular. The acquired image is transmitted to the personal camera via the USB ports for further analysis in real time. The pre-process digital image processing is applied to the eye-motion image by first thresholding the image using adaptive thresholding. The segmented pupil image is then located using template matching. The ROI image around the eye is then processed to determine the eye movement using motion gradient. The horizontal and vertical movement of eye is then estimated using the package software.

1. INTRODUCTION

Eye-motion tracking is a central problem in visual information system and computer vision. Important application of eye-motion tracking includes diagnostic tools for assessing vestibular disease and other neurological disorders [1]and human-computer interface device [2, 3]. Typically, eye-motion tracking is performed on two dimensions to measure the horizontal and vertical motion of the eye. Only recently, attempts have been tried to access the three dimensional motion of the eye [4-6]. Horizontal and vertical eye (2D) position can be determined from pupil center coordinates, which can be computed using center of mass algorithm [7, 8]. The torsional eye position, rotation about the visual axis, can be determined by tracking the location of landmark on the eye [9, 10] or using the more-robust template matching algorithm [8]. Bruce *et. al.* [1] developed a system which uses a flying-spot laser-scanner to selectively image landmark on the eye. The horizontal scan line through the pupil reveals the left and right edge of the pupil allowing the estimation of horizontal position. Similarly, the vertical scan line through the pupil reveals the upper and lower edge of the pupil resulting in the computation of the vertical edge. The system developed by Bruce and his colleague, though, is claimed to be the promising diagnostic tool; the exposure of eye to laser, however, could be extremely hazardous and hence required extremely caution.

Naoki *et. al* [11] proposed the system to measure the horizontal and vertical position of eye movement based on projecting the weak infrared on the limbus, the borderline between the iris and the white of the eye and detecting the change in reflected light. Naoki *et. al* associated the eye motion with the head motion to provide the powerful diagnosis tools for Alzheimer's disease. The sophisticated eye-tracking system consisting of a series of optical lens and mirrors, photo-diodes and the set of infrared LEDs is presented in [12]. The eye movement position is also determined from the change in reflected infrared detected by photo-diodes. The drawback of such system is the complication for operating and the high cost.

In this paper, we proposed the low-cost versatile eye-motion tracking based on applied digital image processing and required no light-source projection. The proposed system consists of two infrared cameras mounted on the binocular. The acquired image is transmitted to the personal camera via the USB ports for real-time processing. The captured frame in the video stream is first converted to binary image using adaptive thresholding. The binarized image is then processed with basic morphological process to remove the artifact. The restrictive template matching is then performed to locate the pupil. The region of interest (ROI) is cropped around the pupil center to cover the eye area. Motion gradient in the ROI is then applied to access the motion direction of the pupil. The horizontal and vertical position of eye movement can then be determined from the motion vector. The packaging software is written in C++ builder to provide friendly user interface. The software exploited the versatile digital image processing library – OpenCV [13] to facilitate the imaging process.

This paper is structured as the following. Section 2 is devoted for system overview. Section 3 describes the digital image process for determining the horizontal and vertical motion. Section 4 explains our experimental result. The conclusions and discussion is provided in the section 5.

2. SYSTEMS

The proposed eye-tracking system is shown in figure 1(b). The system mainly consists of 3 parts: (i) binocular, (ii) Personal Computer (PC) and (iii) the packaging software. The binocular is adapted from the eye-protection eye-glass for the mechanics. The built-in filter on the front plate of the eye-glass protects the environment light that might interfere with the acquisition. The two USB cameras mounted on the front plastic plate transfers the eye image to the PC. The 2.2 GHz Intel Pentium PC with Windows OS is used in the system. Installed on the PC is the C++ Builder eye-tracking packing software. The software provides friendly graphic-user interface to control the cameras. The main function of the software includes record and re-play the video, process the captured frame of the video stream and analyze the eye movement. The shareware OpenCV library is used in the software. OpenCV is the versatile and powerful tools for all digital image process algorithms and is widely used for most of the application related to digital imaging processing.

3. METHODS

The eye-tracking process is shown in figure 2. The image frame is captured from the video stream of the camera at the rate 25 fps for one camera or roughly 10 fps for two cameras. The captured frame shown in figure 3(a) is first converted to binary image using adaptive thresholding. The thresholding result is shown in figure 3(b). The remaining artifact from the thresholding is removed using basic opening-closing morphological process. The morphological result is shown in figure 3(c). The image of pupil now appears as a round disk. To locate the pupil, template matching is applied. The pupil template is the half-circular disk to include the case of matching in the presence of eye-lid drop where the pupil appears as half-circular shape. The similar measure used in the matching is correlation coefficient which is defined as

$$\hat{\rho}_s(X,Y)=\frac{\hat{C}_s(X,Y)}{\sqrt{\hat{\sigma}_x^2\hat{\sigma}_y^2}},\ -1\le\hat{\rho}_s(X,Y)\le 1 \qquad (1)$$

where X is the eye image and Y is the template, covariance $\hat{C}_s(X,Y)$, variances $\hat{\sigma}_x^2$, $\hat{\sigma}_y^2$, and means $\bar{X}$, $\bar{Y}$ are defined by

$$\hat{C}_s(X,Y)\equiv\frac{1}{N-1}\sum_{i=1}^{N}(X_i-\bar{X})(Y_i-\bar{Y}) \qquad (2)$$

$$\hat{\sigma}_x^2\equiv\frac{1}{N-1}\sum_{i=1}^{N}(X_i-\bar{X})^2$$

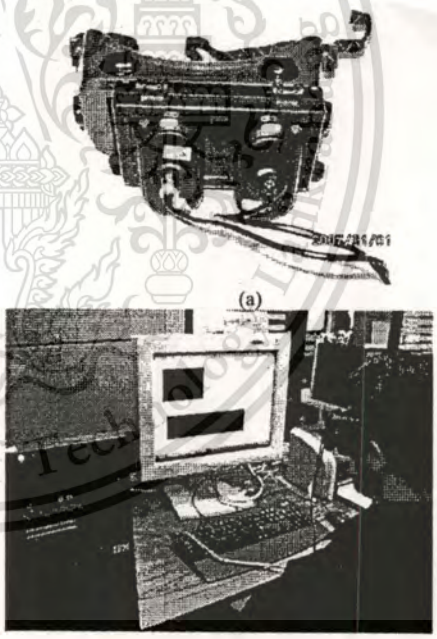
$$\hat{\sigma}_y^2\equiv\frac{1}{N-1}\sum_{i=1}^{N}(Y_i-\bar{Y})^2 \qquad (3)$$

$$\bar{X}\equiv\frac{1}{N}\sum_{i=1}^{N}X_i,\quad \bar{Y}\equiv\frac{1}{N}\sum_{i=1}^{N}Y_i \qquad (4)$$

The ROI is then cropped around the pupil location to cover the eye neighboring area. To estimate the horizontal and vertical movement of the eye, motion gradient is applied. The concept of motion gradient is to estimate the gradient vector from motion history image (MHI). The MHI is a stream of image that stores the motion occurrence by comparing the current frame with the previous frame. MHI pixels where motion occurs are set to the current timestamp, while the pixels where motion happened far ago are cleared. The gradient operator is then applied on the MHI using Sobel operator to derive G_x and G_y which is the gradient in x and y direction respectively. The gradient vector now represents the motion vector and can be estimated from the gradient magnitude and gradient orientation which is defined respectively as

$$\|G\|=\sqrt{G_x^2+G_y^2} \qquad (5)$$

$$\angle G=\tan^{-1}\frac{G_y}{G_x} \qquad (6)$$



(a)

(b)

Figure 1. (a) The Binocular; (b) Eye-motion tracking system

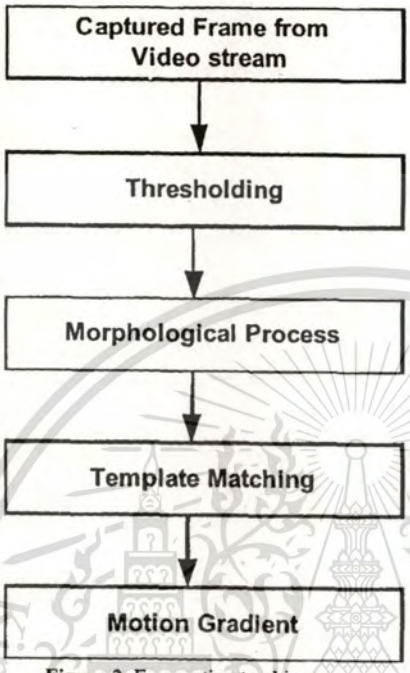


Figure 2. Eye-motion tracking process

4. EXPERIMENT AND RESULT

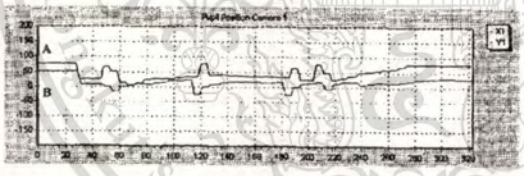


Figure 4. Horizontal and Vertical Position of Pupil

Figure 4 shows an example result of tracking pupil movement at horizontal (X) and vertical (Y) position. Graph A indicates horizontal position and graph B indicates vertical position of pupil movement. The pupil stays in stable position approximately in the first 30 frames, in frame 140-160, and frame 280-320. The graph will be usefull to check whether a person suffers vestibular disease or not. Patient who suffers vestibular disease can be noticed from large amount of oscillation at the graph.

We can also yield the direction and movement flow of the pupil by noticing the arrow in motion gradient image as shown in figure 3(c). This experiment is done in real time tracking with real patient and offline tracking using eye movement video recording file.

(a) (b) (c) (d)

Figure 3. (a) Captured frame of eye; (b) Binarized image; (c) Morphological output image; (d) Motion gradient image.

5. CONCLUSION AND DISCUSSION

This research concerns the design and construction of a low-cost versatile eye-motion tracking system. The *video stream of eye motion is captured via the USB* cameras mounted on the binocular. The captured frame is converted to binary image using adaptive thresholding. The binarized image is then morphologically filtered using opening-and-closing algorithm. The template matching is performed on the filtered image to locate the pupil. The ROI with the center at pupil center is then set to be applied further with motion gradient. The motion gradient vector is then served as eye-motion position for both horizontal and vertical position. The packaging software is written in C++ builder to provide friendly user interface for operating the system. The software exploited the well-known digital image processing library – OpenCV- to facilitate the imaging process. Future works can be done to analyze the eye-motion position for being used as the diagnostic tool for patient with vestibular disease.

REFERENCES

[1] B. A. Wilson, R. D. Jones, P. J. Bones, and T. J. Anderson, "A Flying-Spot Laser Scanner for Tracking Eye Movements," in *18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBS 1996)*, Amsterdam, 1996.

[2] S. D. Magistris, "Eye movement tracking for advanced man-machine interaction.," in *Department of Systems and Informatics*. Florence, Italy: University of Florence, 1995.

[3] Y. Kurio, "Development of eye pointer with free head-motion,"in *20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBS 1998)*, 1998.

[4] D. Zhu, M. Kunin, and T. Raphan, "Robust High-speed Binocular 3D Eye Movement Tracking System Using a Two-radii Eye Model," in *28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBS 2006)*, New York, USA, 2006.

[5] 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, pp. 201-209, 2004.

[6] M. Hatamian and D. J. Anderson, "Design considerations for a real-time ocular counter roll instrument," *IEEE Transactions on Biomedical Engineering*, vol. BME-30, no. 5, pp. 278-288, 1983.

[7] T. Vieville and D. Masse, "Ocular counter-rolling during active head tilting in humans," *Acta Otolaryngol*, vol. 103, pp. 280-290, 1987.

[8] S. T. Moore, I. S. Curthoys, and S. G. McCoy, "VTM-an image-processing system for measuring ocular torsion," *Computer Methods and Programs in Biomedicine*, vol. 35, pp. 219-230, 1991.

[9] S. Yamanobe, S. Taira, T. Morizono, T. Yagi, and T. Kamio, "Eye movement analysis system using computerized image recognition," *Arch Otolaryngol Head Neck Surg.*, vol. 116, pp. 338-341, 1990.

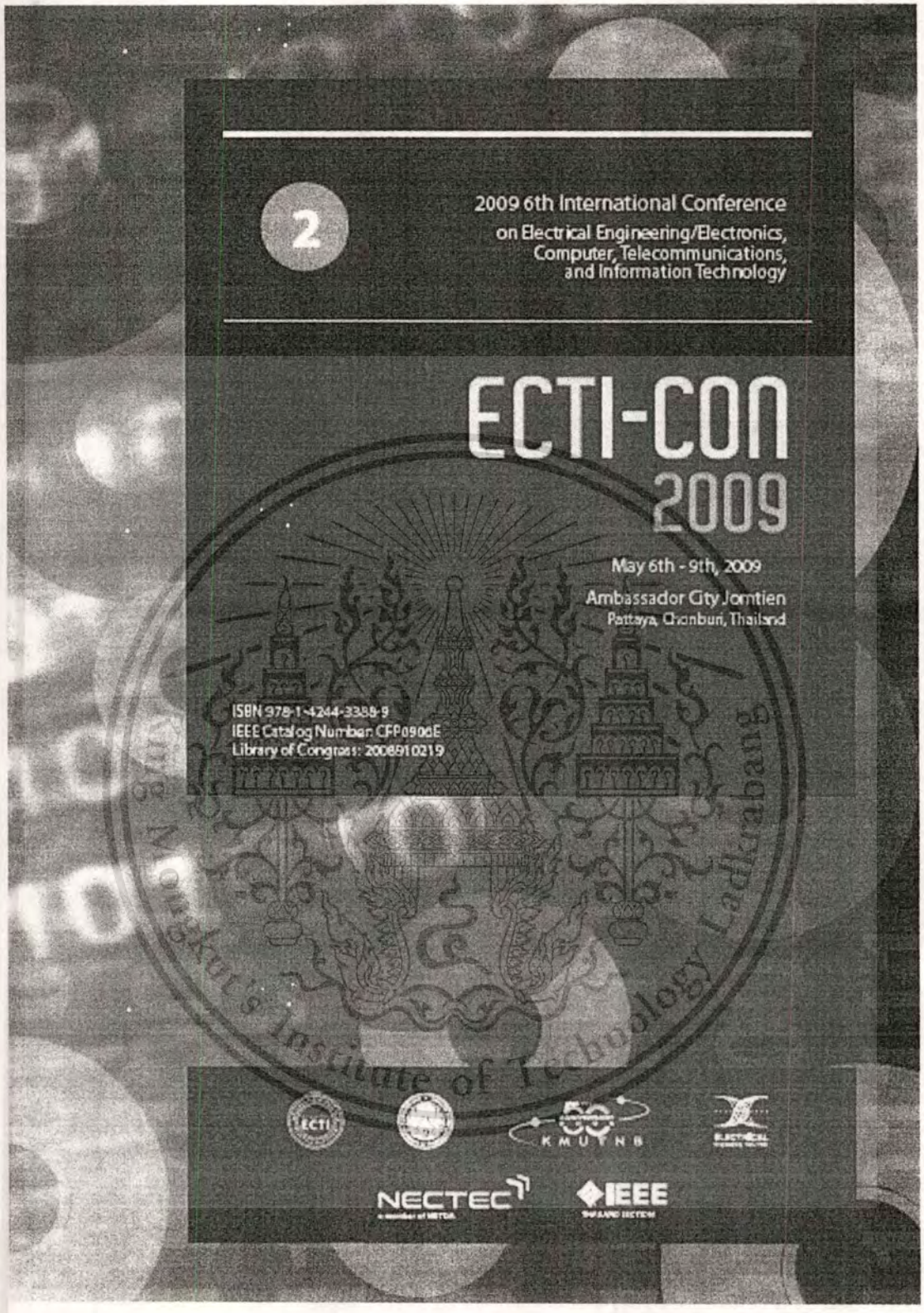
[10] J. A. Parker, R. V. Kenyon, and L. R. Young, *"Measurement of torsion from multi temporal images* of the eye using digital signal processing techniques," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 1, pp. 28-36, 1985.

[11] K. Naoki and e. al., "MEASUREMENT AND ANALYSIS OF HUMAN BRAIN FUNCTION THROUGH HEAD AND EYE MOVEMENT," in *18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBS 1996)*, Amsterdam, 1996.

[12] K. Iwamoto, K.Tanie, and T. Maeda, "Study on the Eye Movement Tracking Type Head Mounted Display -Investigation of the Image Display Method which can Present High resolution (Images at the Intentional VisualArea)," *The Transactions of the Institute of Electrics, Information ,and Communication Engineers (Japan)*, vol. 179, no. 5, 1996.

[13] http://www.sourceforge.net/projects/opencvlibrary

原稿受付日	平成 21 年 3 月 8 日



Real Time Eye Tracking using Initial Centroid and Gradient Analysis Technique*

Sunu Wibirama* Supan Tungjitkusolmun* Chuchart Pintavirooj* Kazuhiko Hamamoto**

*Department of Electronics, Faculty of Engineering,
King Mongkut's Institute of Technology Ladkrabang
Bangkok 10520, Thailand
kpchucha@kmitl.ac.th

**Department of Information Media Technology,
School of Information and Telecommunication Engineering
Tokai University, Tokyo, Japan
hama@keyaki.cc.u-tokai.ac.jp

***Abstract*-Robust and accurate algorithm in real time eye tracking system has been a fundamental and challenging problem for computer vision and biomedical engineering area. This paper proposes a new method to estimate eye-motion position and direction based on initial centroid and gradient analysis technique. The proposed method was validated by tracking eye position within high and low occlusion conditions. Experimental results show that the proposed method improves the accuracy of tracking pupil position compared to conventional window-matching method. The proposed method yields 76.07% and 87,89% accuracies while the conventional window-matching performs 44.56% and 67.89% accuracies within high and low occlusion conditions, respectively.**

I. INTRODUCTION

Eye tracking is an extensive research area in visual information system and computer vision that encompasses anything from diagnostic tool for assessing vestibular disease and other neurological disorders [1] to human-computer interface device [2, 3].

Normally, eye tracking is performed on two dimensions to measure the horizontal and vertical motions of the eye. Horizontal and vertical eye (2D) positions can be determined from pupil center coordinates, which can be computed using center of mass algorithm [4, 5]. The torsional eye position, rotation about the visual axis, can be determined by tracking the location of landmark on the eye [6, 7] or using the window-matching algorithm [4].

Haro et al. [8] proposed pupil tracking based on combining its appearance, the bright pupil effect and motion characteristics. Ji et al. [9] proposed real time subtraction and a special filter to eliminate the external light interferences. Both methods failed to track eyes when they were closed or occluded and robustness was easily affected by the external illumination interference.

In this paper, we propose a new method based on initial centroid and gradient analysis technique. The method is used to measure spatial and directional pupil position in real time tracking. We attempt to overcome accuracy problem arising in real time tracking as in [8, 9] by adding initial centroid algorithm. The method then was tested in tracking pupil motion involving high and low occlusions.

II. THE PROPOSED METHOD

Our proposed method is pattern recognition based approach as shown in Fig. 1. It comprises of three main steps: pre-processing, pupil detection, and motion direction estimation. Initial centroid algorithm is used to deal with window-matching inaccuracy when detecting pupil motion with large head movement, eye blink, and occlusion caused by eyelashes. Gradient analysis based on motion gradient method proposed by [10] is used to estimate the direction of eye movement. The method is implemented in low cost eye tracking system consists of two infrared cameras attached to special binocular, Pentium PC 2.2 GHz and eye tracking software built by Visual C++ and OpenCV image processing library.

A. Pre-processing

First of all, image frames are captured from the video stream of the camera at rate 25 fps. Then, the pre-processing step is applied. This step consists of binarization and morphological process. Binarization process is done by implementing thresholding process. The thresholding result is shown in Fig. 2(b). Morphological process with basic dilation and erosion is inserted to remove remaining artifact from the binary image. The morphological result is shown in Fig. 2(c).

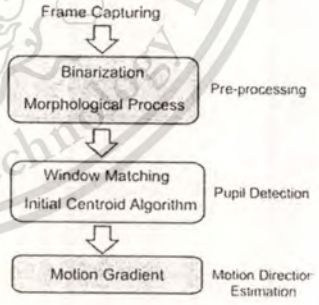


Figure 1. A flowchart diagram for our proposed method.

* This research is supported by AUN/Seed-Net programme from Japan International Cooperation Agency (JICA).

978-1-4244-3388-9/09/$25.00 ©2009 IEEE

B. Pupil Detection

The second step of our algorithm is pupil detection. It comprises of window-matching and initial centroid algorithm. Window-matching method as in [11] is based upon an analysis of the grey level pattern around a point of interest and the search for the most similar pattern in the subsequent frame. Window-matching method calculates motion between a region centered around a certain point of interest $p(x, y)$ on image I_1 that will be displaced by integer values i, j (along the horizontal and vertical directions, respectively) in the subsequent image I_2. The region on the image is represented by squared window of $N \times N$ dimension. To measure similarity, SSD (Sum of Squared Differences) function in reference [11] will be used here, which is defined as

$$R_{sq\_diff}(x,y) = \sum_{i,j=\frac{N}{2}}^{\frac{N}{2}} [I_1(x,y) - I_2(x+i, y+j)]^2 \quad (2).$$

Minimizing the result R represents minimizing the distance of similarity, then it means finding, on the subsequent image, the most similar region to the current image. Therefore, a perfect match will be 0 and bad matches will be large.

Some occlusions caused by large head movement or eye blink can be occurred during eye tracking process. Consequently, an ambiguity between the texture of Region of Interest (ROI) and other regions of the image will yield some error in tracking trajectory. To deal with this problem, improving window-matching process by applying initial centroid algorithm can be considered. The main goal of initial centroid algorithm is reducing inaccuracy of window-matching algorithm by inserting an initial position of previously tracked eye center and applying a time-stamp to prevent tracking error when eye blink or other occlusions are happened. The initial centroid algorithm can be explained as follows:

Step 1: Determine distance threshold (γ_T) and time-stamp threshold (τ_T).
Step 2: Define initial centroid $C(x,y)$. Compute pupil center and save as initial centroid $C(x,y)$. Compute ROI from initial centroid $C(x,y)$.
Step 3: Compute current centroid $C'(x,y)$.
Step 4: Compute the distance between $C(x,y)$ and $C'(x,y)$. The distance γ between initial centroid $C(x,y)$ and current centroid $C'(x,y)$ can be described as

$$\gamma = \sqrt{(x_C - x_{C'})^2 + (y_C - y_{C'})^2} \quad (3),$$

where (x_C, y_C) and $(x_{C'}, y_{C'})$ represent the coordinates of $C(x,y)$ and $C'(x,y)$, respectively.

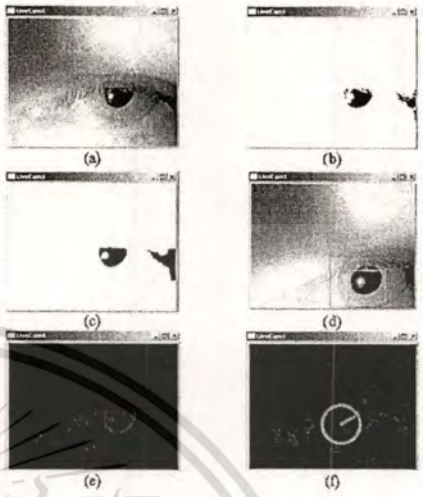


Figure 2. (a) Frame captured at 25 fps. (b) Result after binarization process. (c) Result after morphological process. (d) Implementation of Window Matching and Initial Centroid Algorithm. (e) Motion History Image. (f) Motion direction estimation.

Step 5: Compare between $C(x,y)$ and $C'(x,y)$
5.1. while $\gamma < \gamma_T$ then keep the current ROI same as the previous ROI.
5.2. while $\gamma > \gamma_T$
  5.1.1. Apply the time-stamp τ.
  5.1.2. while $\tau < \tau_T$ then keep the current ROI same as the previous ROI.
  5.1.3. while $\tau > \tau_T$ then save current centroid $C'(x,y)$ as the new initial centroid and repeat step 2 until 5.

C. Motion Direction Estimation

The last step of our proposed method is applying motion gradient which represents the direction of eye movement. The concept of motion gradient is to estimate global gradient estimation from Motion History Image (MHI) [10]. MHI is a scalar-valued image where intensity is a function of recency motion. MHI image calculation can be described as follows:

$$tMHI_\delta = \begin{cases} \tau & \text{if current silhouette at } (x,y) \\ 0 & \text{else if } tMHI_\delta(x,y) < (\tau - \delta) \end{cases} \quad (4),$$

where $tMHI$ is timed Motion History Image, τ is current time-stamp, and δ is time duration constant. The result is a scalar-valued image where more recently moving pixels are brighter.

The gradient of MHI can be calculated efficiently by convolution with separable Sobel filters [12] in the X and Y directions yielding the spatial derivatives $F_x(x,y)$ and $F_y(x,y)$. Gradient magnitude $\|F\|$ and orientation $\phi(x,y)$ at each pixel are defined by (5) and (6), respectively:

$$\|F\| = \sqrt{F_x^{\,2} + F_y^{\,2}} \qquad (5),$$

$$\phi(x,y) = \arctan\frac{F_y(x,y)}{F_x(x,y)} \qquad (6).$$

After calculating the motion gradient, we can then compute the global motion orientation [10]. Calculation of the global orientation should be weighted by normalized MHI values *tMHI* to give more influence to the most current motion within the MHI process. Global motion orientation can be computed as follows:

$$\bar{\phi} = \phi_{ref} + \frac{\sum_{x,y} angDiff\left(\phi(x,y),\phi_{ref}\right) \times norm\left(\tau,\delta,tMHI_{\delta}(x,y)\right)}{\sum_{x,y} norm\left(\tau,\delta,tMHI_{\delta}(x,y)\right)} \qquad (7),$$

where:

- ϕ : the global motion orientation.
- ϕ_{ref} : base reference angle (peaked value in the histogram of orientation).
- $\phi(x,y)$: motion orientation map found from gradient convolutions.
- $norm\left(\tau,\delta,tMHI_{\delta}(x,y)\right)$: normalized *tMHI* value (linearly normalizing the *tMHI* from 0 until 1 using the current time-stamp τ and duration δ).
- $angDiff\left(\phi(x,y),\phi_{ref}\right)$: minimum, signed angular difference of an orientation from the reference angle.

Global motion orientation in our eye tracking process is shown in Fig. 2(f).

III. Experimental Result

We evaluated performances of our method using videos of two different patients acquired from Otolaryngology (Ear, Nose, and Throat) Center, Bangkok Hospital Medical Center. The first patient had a lot of head movement while the second patient had less head movement. The first patient blinked his eyes more than the second patient. We used 1120 and 900 captured frames from patients with high and low occlusion, respectively. The experimental results show that our proposed algorithm exhibited good performance in both high and low occlusion conditions. We evaluated the performance of our proposed method using the following equation:

$$SuccessRate = Samples - error$$

$$Accuracy(\%) = \frac{SuccessRate}{Samples} \times 100\% \qquad (8).$$

TABLE I
Performance of Initial Centroid+Window Matching+ Gradient Analysis

Result / Condition	Accuracy (%)	Error (Frames)	Total Sample (Frames)
High Occlusion	76.07	268	1120
Low Occlusion	87.89	109	900

TABLE II
Performance of Conventional Window-Matching+ Gradient Analysis

Result / Condition	Accuracy (%)	Error (Frames)	Total Sample (Frames)
High Occlusion	44.56	621	1120
Low Occlusion	67.89	289	900

From our experimental results, our proposed algorithm exhibited good performance in both high and low occlusion conditions compared to conventional window-matching algorithm. Major errors in conventional window-matching algorithm are caused by ambiguity between the texture of ROI and other regions of the image when the patients blinked their eyes or moved their heads.

Fig. 3 shows various condition of eye when both proposed algorithm and conventional window-matching were tested. Eyelid occlusions often cause more error than eyelash occlusions in tracking trajectory because of mismatch between the reference pupil image and the captured image.

Fig. 4(a), 4(b), and 4(c) show trajectory results between our proposed algorithm and conventional method. We measured horizontal position as a sample. The conventional window-matching and gradient analysis technique is shown by green colour (WM+MG) while the combination of initial centroid, window-matching, and gradient analysis is shown by red colour (WM+MG+IC).

Fig. 4(d) shows the direction of pupil movement computed from global motion orientation under normal condition with less head movement. The direction was measured in degree unit.

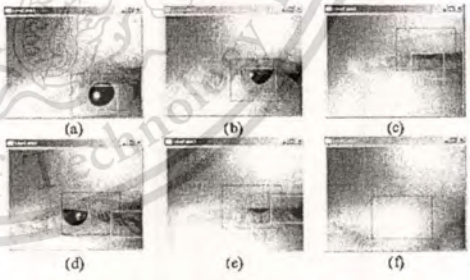


Figure 3. (a) Full circular pupil. (b) Half circular pupil. (c) Almost closed eye image. (d-e) Error pupil tracking using conventional window-matching technique (small rectangle) and success pupil tracking using initial centroid and window-matching (big rectangle). (f) Both methods fail to locate pupil on full closed eye

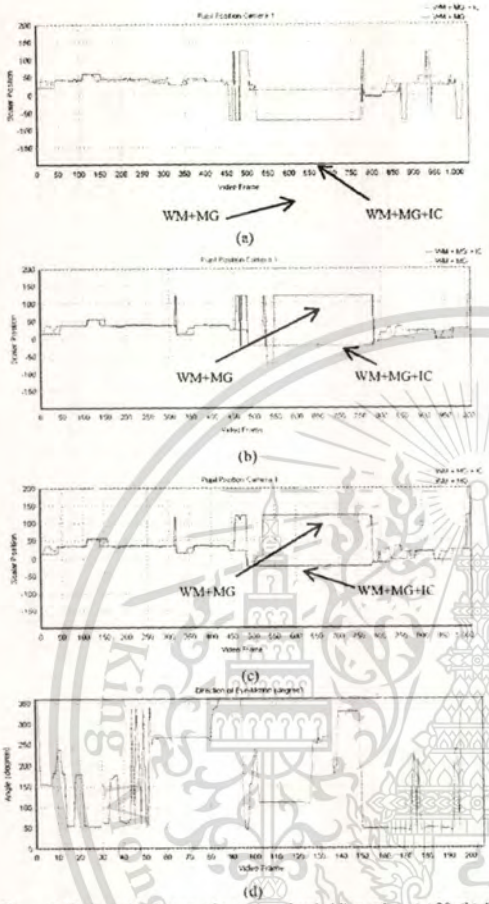


Figure 4. Horizontal trajectory of eye when thresholding value : (a) 20, (b) 40, and (c) 80. (d) Pupil movement direction (degree unit).

Fig. 4(a), 4(b), and 4(c) show that when the thresholding value in binarization process was changed, it affected directly to the accuracy of conventional technique (WM+MG). It can be implied from the drastic change of conventional window-matching and gradient analysis trajectory graph. Nevertheless, the change of threshold value would not caused a lot of error when our proposed algorithm (WM+MG+IC) was applied. The proposed algorithm trajectory graphs show that the tracking trajectory remain stable in low or high threshold value. Thus, the proposed algorithm can handle different thresholding values in reasonable range which is very usefull for different grey level image in binarization process.

IV. Conclusion

This paper presents a new method to track pupil position and direction in low cost eye tracking application. The initial centroid algorithm is added in window-matching technique to increase the accuracy when tracking in high occlusion. The results were compared to conventional window-matching which used no initial centroid algorithm. Gradient analysis technique is used to estimate eye-motion direction.

Experimental results show that the proposed algorithm offers robustness and flexibility in situations where there are a lot of factors causing unexpected error in real time tracking process. Some error in real time eye tracking process are mainly caused by eyelid and eyelash occlusions.

Future works can be done by improving the accuracy of proposed algorithm and applying the low cost eye tracking system as human computer interaction device or a diagnostic tool for vestibular disease.

Acknowledgment

Our thanks to AUN/Seed-Net programme and JICA (Japan International Cooperation Agency) for supporting research fund and equipments, Otolaryngology (Ear, Nose, and Throat) Center, Bangkok Hospital Medical Center for providing eye movement recordings.

References

[1] B.A. Wilson, R.D. Jones, P.J. Bones, and T.J. Anderson, "A Flying-Spot Laser Scanner for Tracking Eye Movements," in *18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBS 1996)*, Amsterdam, 1996.

[2] S.D. Magistris, "Eye movement tracking for advanced man-machine interaction," in *Department of Systems and Informatics*. Florence, Italy: University of Florence, 1995.

[3] Y. Kurio, T.Yagi, and Y.Uchikawa, "Development of eye pointer with free head-motion," in *20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBS 1998)*, 1998.

[4] S.T. Moore, I.S. Curthoys, and S.G. McCoy, "VTM-an image-processing system for measuring ocular torsion," in *Computer Methods and Programs in Biomedicine*, vol. 35, pp. 219-230, 1991.

[5] T. Vieville and D. Masse, "Ocular counter-rolling during active head tilting in humans," in *Acta Otolaryngol*, vol. 103, pp. 280-290, 1987.

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

[7] S. Yamanobe, S. Taira, T. Morizono, T. Yagi, and T. Kamio, "Eye Movement Analysis System Using Computerized Image Recognition," in *Arch Otolaryngol Head Neck Surg.*, vol. 116, pp. 338-341, 1990.

[8] A. Haro, M. Flickner, and I. Essa, "Detecting and Tracking Eyes by Using Their Physiological Properties, Dynamics, and Appearance," in *Proceedings IEEE CVPR 2000*, Hilton Head Island, South Carolina, June 2000.

[9] Q. Ji and X. Yang, "Real Time Visual Cues Extraction for Monitoring Driver Vigilance," in *Proceedings of International Workshop on Computer Vision Systems*, Vancouver, Canada, July 7-8, 2001.

[10] G. Bradski and J. Davis, "Motion Segmentation and Pose Recognition with Motion History Gradients," in *IEEE Workshop on Applications of Computer Vision*, 2000.

[11] P. Anandan, "A computational framework and an algorithm for the measurement of visual motion", vol.2, 1989, p. 283-310.

[12] R.C. Gonzalez and R.E. Woods, "Digital Image Processing", 2nd edition, 2002, p. 134-137.

BMEiCON 2009

August 13-14, 2009 Phuket, Thailand

PROGRAM AND ABSTRACTS

The 2nd Biomedical Engineering International Conference

IEEE IEEJ IFMBE NECTEC NSTDA THAILAND

BMEiCON 2009 August 13-14, 2009 Phuket, Thailand

DUAL CAMERAS ACQUISITION FOR THREE-DIMENSIONAL EYE-MOTION TRACKING

Sunu Wibirama[1], Supan Tungjitkusolmun[2], Chuchart Pintavirooj[3], Kazuhiko Hamamoto[4]
[1,2,3]Department of Electronics, Faculty of Engineering,
King Mongkut's Institute of Technology Ladkrabang, Thailand
[4]Department of Information Media Technology,
School of Information and Telecommunication Engineering,
Tokai University, Japan
Email: kpchucha@kmitl.ac.th

ABSTRACT

Three-dimensional eye-motion tracking using dual cameras is presented in this paper. The system consists of two arbitrary orientation USB cameras mounted on tripod. The two-dimensional positions of pupil center were obtained using template matching technique. The three-dimensional positions of pupil center were then computed using Direct Linear Transformation technique. The algorithm was validated by comparing an eye model movement in world Euclidean coordinate system and DLT system. The system yielded error 0.2072, 0.1675, and 0.1797 pixel for X, Y, and Z axes respectively with range of error from 0 to 1 pixel.

1. INTRODUCTION

Eye-motion tracking is an extensive research area in biomedical engineering and computer vision. Recent advances in understanding of the motor control of eye movements and new findings in the anatomy of extraocular muscles have greatly increased the interest in a complete three-dimensional (3D) description of eye position [1-5].

Scleral Search Coil method has been used successfully to measure horizontal, vertical and torsional components of the eye [1,2]. However, this technique is expensive and invasive, limiting the measurement duration to about 30 mins because of discomfort during recording process [3]. Parker *et al.* [4] proposed a method to measure 3D components based on digital signal processing. The measurement was done by analyzing marker on several printed eye images. Although the result was promising, this method did not support real time processing. Shih *et al.* [5] proposed a satisfactory 3D eye tracking system using multiple infrared reflections and stereo CCD cameras. The 3D gaze was measured by comparing infrared mark and pupil center. Patients were required to fix their heads to get accurate tracking result from table-mounted eye-motion tracking system. The drawbacks of such system are that it is complicated to use and associated high costs.

In this research, we introduce a real time 3D eye-motion tracking system without infrared reflections. We attempted to reduce the cost by designing an eye-motion tracking system using two images acquired from two arbitrarily oriented USB cameras. The 3D coordinates of pupil center were then estimated using DLT technique proposed in [6].

This work is organized as follows: Section II explains the proposed system and method. Experimental results are provided in Section III. Discussions and conclusions are given in Section IV.

2. PROPOSED SYSTEM AND METHOD

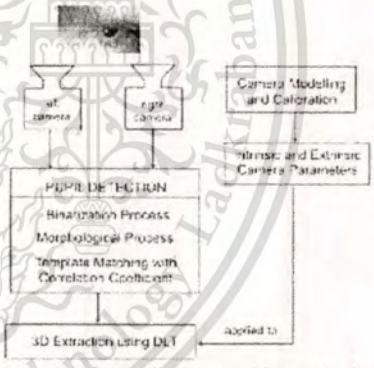


Figure 1. 3D eye coordinates acquiring method using two arbitrarily oriented cameras.

The eye-motion tracking system consists of 3 parts: (i) two arbitrarily oriented USB cameras, (ii) personal computer (PC) and (iii) the 3D eye tracking software created using Visual C++ and OpenCV library [7]. We used Oker 177 web camera (Shenzhen Golden Tiger & Dragon Technology Develop Co., Ltd, Guangdong, China) with sampling rate of 60 Hz, focus range of 30 mm to infinitive, and resolution of 2 megapixels. The two cameras were installed on small tripod in order to provide stable and comfortable environment during the experiment process. The 2.2 GHz Intel Pentium PC with Windows XP operating system was used in the system.

The 3D eye coordinates acquiring method is shown in Fig. 1. First, we applied camera modeling and calibration to obtain intrinsic and extrinsic camera parameters which would be used in DLT algorithm. The pupil detection process using template matching was then conducted to obtain 2D positions of pupil center. The 3D coordinates of pupil center were then extracted using DLT technique.

2.1. Camera Modeling and Calibration

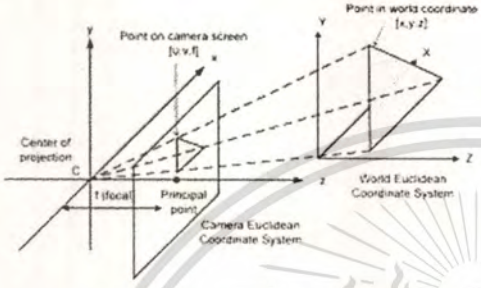


Figure 2. Perspective transformation from world Euclidean coordinate system to camera Euclidean coordinate system.

A point located in real 3D space (world Euclidean coordinate system) can be expressed as point $[x\ y\ z\ 1]^T$. A point $[x'\ y'\ z'\ 1]^T$ taken by a camera undergoes a linear transformation from 3D projective space to 2D projective space as shown in Fig.2. Point $[u\ v\ w]^T$ in camera coordinate system is obtained by examining the following homogeneous coordinate formula:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = K\begin{bmatrix} I_3 & 0_3 \end{bmatrix}\begin{bmatrix} R & -T \\ 0_3^T & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix},$$

$$\text{or} \quad \begin{bmatrix} u \\ v \\ w \end{bmatrix} = M\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \text{ where } M = [KR \mid -KRT]. \quad (1)$$

Matrices K and $[I_3 \mid O_3]$ in (1) are instrinsic parameters which define physical parameters of the camera. Matrix K is camera calibration matrix defined as

$$K = \begin{bmatrix} fa & 0 & u_0 \\ 0 & fb & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

where f is focal length in camera projection; a and b are conversion factor from physical unit to pixel unit in the x and y directions respectively; and (u_0, v_0) is a principal point in camera projection.

Matrix $[I_3 \mid O_3]$ is defined as $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$.

Matrix $\begin{bmatrix} R & -T \\ 0_3^T & 1 \end{bmatrix}$ consists of extrinsic parameter which defines camera location and orientation in world Euclidean coordinate system. Rotation matrix R expresses three elementary rotations of camera Euclidean coordinate axes which respect to the world Euclidean coordinate system. Rotation along x, y, and z are termed as pan, tilt, and roll respectively. Translation vector T gives three elements of the translation of the origin of the world Euclidean coordinate system with respect to camera Euclidean coordinate system. Matrix 0_3^T is vector [0 0 0]. To obtain matrix M, observing each known point $X = [x\ y\ z\ 1]^T$ and its corresponding 2D image point $[u\ v\ w]^T$ will yield an equation:

$$\begin{bmatrix} x & y & z & 1 & 0 & 0 & 0 & 0 & -ux & -uy & -uz & 1 \\ 0 & 0 & 0 & 0 & x & y & z & 1 & -vx & -vy & -vz & 1 \end{bmatrix} \underbrace{\begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_{11} \end{bmatrix}}_{M} = 0$$

$$\text{or} \quad AM = 0. \quad (3)$$

If n such points in world euclidean coordinate system are available, A will be of size $2n \times 12$. M can be solved by performing Singular Value Decomposition (SVD) of A to derive $A = UDV^T$. U is left singular vector, D is diagonal matrix consists of singular values of A, and V is right singular vector of A. The last column of V is the solution for M. To separate extrinsic parameter, observe that M in (1) can be written as $M = [KR \mid -KRT] = [A \mid b]$. Then we can obtain translation matrix T by solving equation $T = -A^{-1}b$. To determine R, we decompose A into a product of two matrices K and R using QR decomposition.

2.2. Pupil Detection

Figure 3. Pupil detection using plastic eye model: (a) original images, (b) segmentation using thresholding, (c) removing small artifact at pupil's boundary using morphology, (d) result of template matching with correlation coefficient.

The pupil detection process is shown in Fig. 3. The image frames were captured from two different videos of two cameras with arbitrary orientation. The captured frames shown in Fig. 3.a were first segmented to binary images using thresholding. The thresholding results were shown in Fig. 3.b. The remaining artifact at the pupil's border was removed using basic opening-closing morphological process. The morphological results were shown in Fig. 3.c.

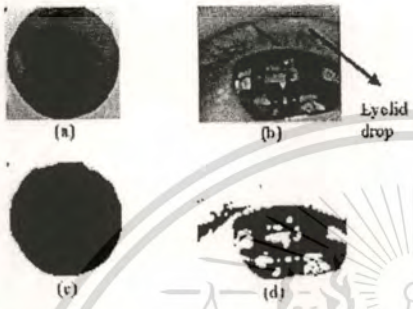


Figure 4. Template image used in experiment: (a) template of plastic eye model; (b) template of real human eye with eyelid interference; (c) template of plastic eye model after thresholding; (d) template of real human eye after thresholding.

To locate the pupil, template matching technique was then applied. The pupil template for tracking plastic eye model was circular disk as shown in Fig.4.a. The pupil template for tracking real human eye was a half-circular disk to include the case of matching in the presence of eyelid drop where the pupil appeared as half-circular shape as shown in Fig. 4.b. The templates were first converted to binary images (Fig.4.c and Fig.4.d) before used in matching process. The template image was then slided through the binary image of captured frames during the matching process. The similarity measurement used in the matching procedure was correlation coefficient $\hat{\rho}_2(X,Y)$ defined as

$$\hat{\rho}_2(X,Y) = \frac{\hat{C}_2(X,Y)}{\sqrt{\hat{\sigma}_1^2 \hat{\sigma}_2^2}}, \quad -1 \le \hat{\rho}_2(X,Y) \le 1, \tag{4}$$

where X is the eye image and Y is the template, $\hat{C}_2(X,Y)$ is covariance, $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ are variances. In this experiment, the best match was found when the template matching result achieved the maximum value of $\hat{\rho}_2(X,Y)$. The template matching result was a region inside the black rectangle as shown in Fig.3.d.

2.3. 3D Coordinates Extraction

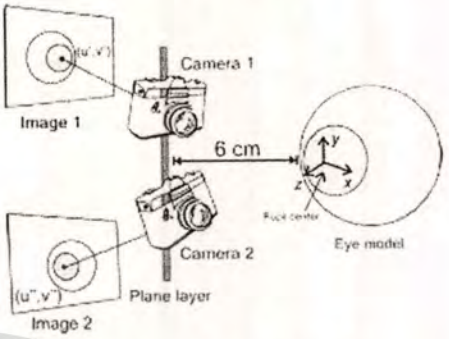


Figure 5. 3D coordinates extraction

Direct Linear Transformation is a method that extracts 3D coordinates of the object from 2D coordinates in multiple photographs taken at arbitrary poses around the object. In this research, we used two cameras which were located approximately 6 cm in front of plastic eye model as shown in Fig.5. Camera modeling and calibration process yielded intrinsic and extrinsic parameters. Parameters for the first and second cameras were inserted into matrix M' and M'' respectively. Note that (u',v') and (u'',v'') are 2D coordinates resulted from image 1 and image 2 respectively. Thus, to estimate 3D coordinates (x,y,z) of pupil center, the following homogenous equation can be solved:

$$\begin{bmatrix} u'm_3'^T - m_1'^T \\ v'm_3'^T - m_2'^T \\ u''m_3''^T - m_1''^T \\ v''m_3''^T - m_2''^T \end{bmatrix} X = AX = 0 \tag{5}$$

where $m_i'^T$ are the row i^{th} of the projection matrix M, while X is matrix $[x\ y\ z\ 1]^T$.

3. EXPERIMENTAL RESULT

To validate the 3D positions extracted from each video frame, we compared real and computed movements of plastic eye model in X, Y, and Z axes. Real movement is an eye model movement in world Euclidean coordinate system. We measured manually the X, Y, and Z movement using a ruler. Computed movements were derived from DLT technique. The experiments were repeated 5 times for each axis. From experimental results, we found that every 1 mm movement in world Euclidean coordinate system yielded = 1 pixel movement in DLT system.

TABLE I. COMPARISON RESULT OF X-AXIS MOVEMENT

Exp.	Real X-Axis Movement (mm)	Computed X-Axis Movement (pixel)	Error (pixel)
1st	5	5.1453	0.1453
2nd	5	5.2989	0.2989
3th	5	5.2972	0.2972
4th	5	5.0181	0.0181
5th	5	5.1246	0.1246
Root Mean Square of Error			**0.2072**

TABLE II. COMPARISON RESULT OF Y-AXIS MOVEMENT

Exp.	Real Y-Axis Movement(mm)	Computed Y-Axis Movement (pixel)	Error (pixel)
1st	5	4.7414	0.2586
2nd	5	5.2460	0.2460
3th	5	5.1098	0.1098
4th	5	4.9731	0.0269
5th	5	4.9882	0.0118
Root Mean Square of Error			**0.1675**

TABLE III. COMPARISON RESULT OF Z-AXIS MOVEMENT

Exp.	Real Z-Axis Movement(mm)	Computed Z-Axis Movement (pixel)	Error (pixel)
1st	5	5.6912	0.6912
2nd	5	5.3576	0.3576
3th	5	5.0199	0.0199
4th	5	5.2827	0.2827
5th	5	5.1872	0.1872
Root Mean Square of Error			**0.3797**

The experimental results show that the system yielded error approximately 0.2072, 0.1675, and 0.3797 pixel for *X*, *Y*, and *Z* axes respectively with range of error 0 to 1 pixel. The error are considered within acceptable tolerance. Most error were caused by inaccurate pupil detection which supplies wrong 2D coordinate in DLT system.

The second experiment was system implementation using real human eye as shown in Fig6. We were interested in testing the performance of pupil detection algorithm when the subject fully opened his eye (Fig. 6.a.1 and Fig. 6.a.2), closed his eye (Fig. 6.e.1 and Fig. 6.e.2), and opened only a half of his eye (Fig. 6.b.1, Fig. 6.b.2, Fig. 6.c.1, and Fig.6.c.2). We also asked the subject to wear eyeglasses to check whether the algorithm worked in the presence of interference caused by eyeglasses and some reflections of unwanted light sources (Fig. 6.f.1, Fig.6.f.2, Fig.6.g.1, Fig6.g.2, Fig.6.h.1, and Fig.6.h.2). The subject was also asked to blink his eye and to lean his head towards left and right direction. The fail cases were mainly caused by large ambiguity of black area resulted by tresholding segmentation process. The algorithm detected the eyebrow or the other areas instead of the pupil.

4. CONCLUSION AND DISCUSSION

Three-dimensional eye-motion tracking system has been presented in this paper. The DLT techniques were adopted to derive dynamic 3D information from arbitrary orientation cameras. The 3D eye tracking system was then validated in X, Y, and Z directions by comparing real and computed movements. The system was also tested using real human eye. The proposed system successfully detected 3D coordinates within acceptable error tolerance. Future work will be done to improve accuracy of pupil detection process.

REFERENCES

[1] H.Collewijn, J.van der Steen, L.Ferman, & T.Jansen, "Human ocular counterroll: Assesment of static and dynamic properties from electromagnetic scleral coil recordings," in *Exp. Brain Res.*, vol. 59, no.1, pp. 185-96, 1985.

[2] D.A. Robinson, "A Method of measuring eye movement using a scleral search coil in a magnetic field," in *IEEE Trans. Biomed. Electron*, vol. 10, pp. 137-45,1963.

[3] J.P. Ivins, J. Porrill, & J.P. Frisby, "Deformable model of the human iris for measuring ocular torsion from video images," in *IEEE Proc.Vision, Image, Signal Processing*, pp. 213-20, 1998.

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

[5] S. Shih & Jin Liu, "A Novel Approach to 3-D Gaze Tracking Using Stereo Cameras", in *IEEE Transactions on System, Man, and Cybernatics*, vol. 34, no.1, pp. 234-245, 2004.

[6] Y.I. Abdel-Aziz and H.M. Karara, "Direct linear transformation into object space coordinates in close-range photogrametry," in *Proc. Symp. Close-Range Photogrametry*, pp. 1-18, 1971.

[7] http://opencv.willowgarage.com

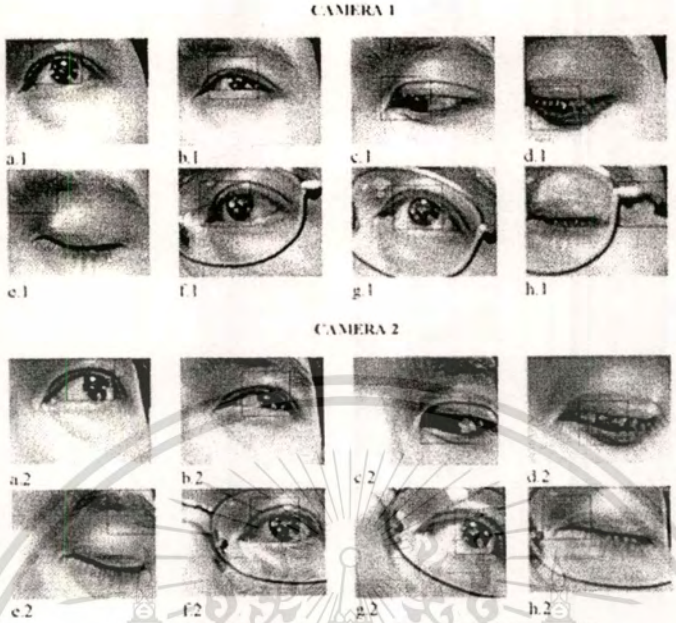


Figure 6. 2D tracking using arbitrary orientation cameras: success cases on tracking full eye (a.1 and a.2), half-closed eye (b.1 ,b.2,c.1, and c.2), almost closed eye (d.1 and d.2). Fail cases on tracking full closed eye (e.1 and e.2). Success cases on tracking full eye interfered by eyeglasses (f.1,f.2,g.1, and g.2) and fail cases on tracking full closed eye interfered by eyeglasses (h.1 and h.2).