BRNO UNIVERSITY OF TECHNOLOGY

Faculty of Electrical Engineering
and Communication

MASTER'S THESIS

Brno, 2023

Bc. Michal Špaček

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

NÁVRH A REALIZACE SYSTÉMU PRO SLEDOVÁNÍ POHYBU OČÍ

DESIGN AND IMPLEMENTATION OF EYE TRACKING SYSTEM

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE Bc. Michal Špaček

AUTHOR

VEDOUCÍ PRÁCE doc. Ing. Radim Kolář, Ph.D.

SUPERVISOR

BRNO 2023

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

Master's Thesis

Master's study program **Biomedical Engineering and Bioinformatics**

Department of Biomedical Engineering

Student: Bc. Michal Špaček ***ID:*** 211690

Year of study: 2 ***Academic year:*** 2022/23

TITLE OF THESIS:

Design and implementation of eye tracking system

INSTRUCTION:

1) Study and describe the basics of eye tracking, available hardware solutions, methods of image data analysis and calibration methods used in this field. 2) Using available cameras and lenses, design a basic system that will be suitable for eye tracking. 3) Implement or use an available implementation of a method for pupil tracking, or for tracking illumination reflections from the cornea. 4) Design a presentation of the stimuli on a computer monitor and perform measurements on a group of volunteers. 5) Process the data and present the results using a suitable visualization. 6) Evaluate the implemented system, list its pros and cons, and suggest further possibilities for its extension or improvement.

RECOMMENDED LITERATURE:

[1] Mahanama R. et al. Eye Movement and Pupil Measures: A Review, Frontiers in Computer Science, 3(733531), 2022

[2] Li Z. et al. A Review of Main Eye Movement Tracking Methods, Journal of Physics: Conference Series, 1802(4), 2021

Date of project specification: 6.2.2023 ***Deadline for submission:*** 22.5.2023

Supervisor: doc. Ing. Radim Kolář, Ph.D.

prof. Ing. Valentine Provazník, Ph.D.
Chair of study program board

WARNING:

The author of the Master's Thesis claims that by creating this thesis he/she did not infringe the rights of third persons and the personal and/or property rights of third persons were not subjected to derogatory treatment. The author is fully aware of the legal consequences of an infringement of provisions as per Section 11 and following of Act No 121/2000 Coll. on copyright and rights related to copyright and on amendments to some other laws (the Copyright Act) in the wording of subsequent directives including the possible criminal consequences as resulting from provisions of Part 2, Chapter VI, Article 4 of Criminal Code 40/2009 Coll.

ABSTRACT
The aim of this thesis is to design and implement an eye tracking system. Eye tracking is a technology that allows for the measurement and analysis of eye movements and gaze behavior, providing insights into attention, perception, and cognitive processes. The thesis discusses the necessary theory related to eye tracking, including an overview of the anatomy and physiology of the human eye, as well as the different types of eye tracking, methodologies, applications, and more. For the practical part of this work, a solution design was created, which served as the basis for the implementation of both the hardware and software components. A dataset was collected, consisting of videos capturing eye images and their movements. Subsequently, the dataset was analyzed using the developed software solution, and the results of the individual analysis components were discussed.

KEYWORDS
eye tracking system, pupil detection, gaze estimation, IDS monochromatic camera, infrared light, eye movement analysis

ABSTRAKT
Cieľom tejto práce je design a implementácia systému na sledovanie očí. Sledovanie očí je technológia, ktorá umožňuje meranie a analýzu pohybov očí a správania pohľadu, poskytujúc informácie o pozornosti, vnímaní a kognitívnych procesoch. V práci je diskutovaná potrebná teória k sledovaniu očí, konrétne prehľad anatómie a fyziológie ľudského oka a potom už samotné sledovanie očí, jej delenie, metológia, využite a ďalšie. Pre praktickú časť tejto práce bol vytvorený návrh riešenia, podľa ktorého sa ďalej implementovali obe časti praktickej práce, teda hardwarová a softwarová časť. Bol nameraný dataset, ktorý obsahoval videá zobrazujúce snímky očí a ich hýbanie. Následne bol dataset analyzovaný pomocou softwarového riešenia a diskutovali sa dosiahnuté výsledky z individuálnych častí analýzy.

KĽÚČOVÉ SLOVÁ
sledovanie očí, detekcia pupily, určenie pohľadu, IDS monochromatická kamera, infračervené svetlo, analýza pohybov oka

Typeset by the `thesis` package, version 4.07; `http://latex.feec.vutbr.cz`

Rozšírený Abstrakt

Sledovanie pohybu očí je výkonnou technológiou, ktorá umožňuje meranie a analýzu pohybov očí a pohľadu, čo poskytuje poznatky o pozornosti, vnímaní a kognitívnych procesoch. Otvára nové perspektívy o ľudskom správaní, ktoré neboli doteraz preskúmané. Systémy sledovania očí sa uplatňujú v rôznych oblastiach, ako sú psychológia, neuroveda, interakcia medzi človekom a počítačom, výskum trhu, testovanie použiteľnosti, a taktiež veľmi často v oblasti medicíny.

Táto práca sa delí na viaceré kapitoly. Prvé dve popisujú teóriu tejto práce, konkrétne ľudské oko a sledovanie očí. Následne prichádza návrh a implementácia systému, na čo nadväzuje kapitola výsledkov a diskusie.

S cieľom navrhnúť a implementovať systém sledovania očí sa táto práca zaoberá teoretickými základmi a praktickými aspektami sledovania očí. Teória začína porozumením zložitostí ľudského oka, skúmaním jeho anatómie, fyziológie a rôznych typov pohybov očí. Tieto poznatky tvoria základ sledovania očí a umožňujú nám oceniť zložitosť sledovania pohybov očí. Posúvame sa ďalej do oblasti sledovania očí samotného. Definujeme, čo obnáša sledovanie očí a skúmame jeho široké uplatnenie v rôznych oblastiach. Či už ide o porozumenie správania používateľov v interakcii medzi človekom a počítačom alebo dešifrovanie kognitívnych procesov v neurovede, sledovanie očí sa ukázalo ako nedoceniteľný nástroj. Diskutujeme tiež o metrikách a meraniach používaných v problematike sledovania očí, čo nám umožňuje zhromažďovať významné údaje a získať cenné poznatky. S pevnými teoretickými základmi prechádzame do praktickej časti tejto práce.

Navrhované riešenie systému sledovania očí sa začína formovať, zahŕňajúc hardvérové a softvérové komponenty. Venujeme pozornosť dôkladnému návrhu hardvérovej súpravy, ktorá zahŕňa kameru, infračervené svetlo, počítač a monitor. Každý komponent je starostlivo testovaný a overovaný, pričom sme riešili problémy, ktoré sa vyskytli a vykonávali sme testovacie merania na malých skupinách účastníkov. Počas druhej fázy zhromažďovania údajov sme identifikovali problém s infračerveným svetlom a nahradili ho vhodným riešením. Súčasťou samotného meranie bola prezentácia stimulov. Táto prezentácia navádzala subjekt, kam sa má pozerať, a teda, kam majú smerovať jeho oči. V prezentácií boli urobené aj pauzy, počas ktorých so mohol subjekt oddýchnuť a nežmurkal v signále, ktorý sa používa na analýzu. Tieto pauzy v prezentácií sa totiž pri analýze odstraňujú. Týmto spôsobom bola docielená úspešná filtrácia artefaktov žmurknutia. Po získaní datasetu sme sa pustili do implementácie softvérových komponentov. To zahŕňalo aplikáciu rôznych metód spracovania a analýzy údajov. Prvým krokom bolo detekovať oči a ich okolie, aby sme získali nové oblasti záujmu. Z týchto oblastí záujmu sa detekovala zrenica pomocou dvojitej filtrácie a thresholdingu, z ktorých sme extrahovali

súradnice stredu každej zrenice pre každý snímok. Tieto údaje tvoria jadro tejto práce a boli ďalej segmentované na základe prezentácie stimulov. Taktiež dôležitým krokom bola konverzia pixelových hodnôt do reálnej metriky, v práci sa využíva centimeter.

Analýza sa delí na dynamickú a statickú časť. V dynamickom rozboru sme sa sústredili na skúmanie najkratších a najdlhších reakčných časov na stimuly, ako aj najrýchlejších a najpomalších pohybov očí po obrazovke. Výsledky sme vizualizovali v grafe znázorňujúcom rýchlosť pohybu očí v čase. V statickej analýze sme hodnotili účinnosť mriežkového stimulu v strede kruhu pri udržiavaní fixácie bez menších pohybov očí. Výsledky naznačovali pozitívny efekt. Zvyšné časti statickej analýzy, ktoré zahŕňali analýzu textu a obrázkov, poskytli zmiešané výsledky. Cieľom analýzy textu bolo určiť potrebný čas a počet fixácií pre čítanie riadku textu, zatiaľ čo analýza obrázka generovala heatmapu. Iba jedna heatmapa bola úspešná, zatiaľ čo ostatné preukazovali významné nepresnosti. Tieto rozdiely sa pripisujú chybám pri prevode súradníc pixelov na reálne metriky. Táto chyba by mohla vzniknúť počas fázy zberu údajov, najmä pri nadmernom pohybe hlavy účastníkov, čo komplikovalo kalibráciu a presné určenie konverzného faktora. Napriek týmto obmedzeniam implementovaný systém sledovania očí plní svoj účel. Jeho používateľsky prívetivé rozhranie poskytuje pohodlie a jednoduchosť použitia, aj keď môže obetovať presnosť zhromaždených údajov.

Možné úpravy by mohli zahŕňať zameranie sa na konkrétny typ analýzy zo škály možností a hlbšie preskúmanie toho, ako ju využiť. Navyše, rozšírenie datasetu a vykonanie štatistických porovnaní by mohlo poskytnúť ďalšie poznatky a príležitosti na zlepšenie. Zhrnutím je táto práca úspešne dosiahla svoje ciele navrhnutím a implementáciou komplexného systému sledovania očí. Poskytla cenné poznatky o pohybe očí a súvisiacich kognitívnych procesoch. Aj keď systém preukazuje jednoduchosť použitia, má obmedzenia v presnosti údajov. Budúce práce by mohli zahŕňať zdokonalenie konkrétnych častí analýzy alebo rozšírenie datasetu pre širšie štatistické porovnania a ďalšie vylepšenia

Author's Declaration

Author:	Bc. Michal Špaček
Author's ID:	211690
Paper type:	Semestral Project
Academic year:	2022/23
Topic:	Design and implementation of eye tracking system

I declare that I have written this paper independently, under the guidance of the advisor and using exclusively the technical references and other sources of information cited in the paper and listed in the comprehensive bibliography at the end of the paper.

As the author, I furthermore declare that, with respect to the creation of this paper, I have not infringed any copyright or violated anyone's personal and/or ownership rights. In this context, I am fully aware of the consequences of breaking Regulation § 11 of the Copyright Act No. 121/2000 Coll. of the Czech Republic, as amended, and of any breach of rights related to intellectual property or introduced within amendments to relevant Acts such as the Intellectual Property Act or the Criminal Code, Act No. 40/2009 Coll. of the Czech Republic, Section 2, Head VI, Part 4.

Brno . . . .

author's signature*

*The author signs only in the printed version.

ACKNOWLEDGEMENT

I'd like to thank and acknowledge my professor, doc. Ing. Radim Kolář Ph.D., for their invaluable help and support, as I worked on this thesis. I am very thankful for their time, patience, and willingness to share their knowledge and skills. Their help and advice have been very important to me as I've worked through the challenges and difficulties of this project.

Contents

List of Figures

Introduction

Eye tracking is a powerful technology that allows the measurement and analysis of a subject's eye movements and gaze behavior. It does this by following the subject's eyes as they move. It provides essential new perspectives on human attention, perception, and cognitive processes that have not been previously explored. Eye tracking systems are able to collect useful information about where a person is looking, how long eyes fixate on certain objects or areas of interest, and how a person's gaze behavior relates to their cognitive processes because of tracking the precise movements of a person's eyes. The fields of psychology, neuroscience, human-computer interaction, market research, and usability testing are just some of the areas that have found uses for eye tracking systems.

The aim of this thesis is to design and provide the implementation of an eye tracking system. The theoretical foundations necessary to understand eye tracking as a whole will be discussed firstly. It will start with the anatomy and physiology of the eye and then move to the concept of eye tracking, including its definition, applications, and various ways it can be utilized. From theory, the focus will shift to the practical aspect, which is based on the proposed solution. The practical part itself will be divided into hardware and software component. The hardware section will describe the individual steps that make up the eye tracking setup. It will then cover the data collection and dataset creation process. The methods for data processing will be described, leading to the actual data analysis, which will be divided into dynamic and static analyses. Finally, the results will be presented, followed by a necessary discussion to clarify the findings. An evaluation of the implemented system will be conducted, including an overview of its advantages, disadvantages, and possible modifications.

Objectives of the Thesis

1) Study and describe the basics of eye tracking, available hardware solutions, methods of image data analysis and calibration methods used in this field.
2) Using available cameras and lenses, design a basic system that will be suitable for eye tracking.
3) Implement or use an available implementation of a method for pupil tracking, or for tracking illumination reflections from the cornea.
4) Design a presentation of the stimuli on a computer monitor and perform measurements on a group of volunteers.
5) Evaluate the implemented system, list its pros and cons, and suggest further possibilities for its extension or improvement.

1 Human Eye

The human eye, which is a wonder of biomedical engineering, is the cause of sight, something that gives us the ability to see and thus perceive the world around us in remarkably fine detail. It allows us to take in information about our surroundings and interpret it in our brains. The complex structure of the eye, which includes various parts, is essential for the visual process. The sophisticated mechanisms of the human eye can offer important insights into not only the physiological processes of vision but also the interactions between people and their surroundings.

1.1 Anatomy of the Human Eye

A pair of eyes make up the peripheral part of the visual system. It is a complex organ that facilitates spatial orientation, mediates the perception of the largest amount of information about the surrounding environment, and enables the perception of light and color. The eyeball (bublus oculi) and accessory eye organs (organa oculi accessoria) combine to form the visual organ, which is contained in the orbit. [1]

Eyeball

The eyeball is roughly spherical in shape and measures about 23 mm in height and width and 24–26 mm in anterior–posterior diameter. It is made up of two sphere segments with different curvature radii. The radius of curvature of the larger posterior segment (sclera) is 11–12 mm, whereas the radius of curvature of the smaller anterior segment (cornea) is 7-8 mm. The larger posterior part of the eyeball is located deep within the orbit; only the smaller anterior part of the eyeball is visible in the eye opening between the open eyelids. [1]

The eyeball consists of the wall of the eyeball and the contents within the eyeball. The wall of the eyeball is composed of three layers. The outer fibrous layer (tunica fibrosa bulbi), which is formed by the sclera and the cornea, then the middle vascular layer (tunica vasculosa bulbi), also known as the uvea, which is made up of the choroid at the posterior, the ciliary body in the middle, and the iris at the front, and the inner sensory layer (tunica interna s. sensoria bulbi), which forms the retina. The contents of the eyeball are transparent and clear structures that transmit light rays and focus them so they fall on the retina. They are the optical environment of the eye and consist of the lens, retina, and both the front and back sections of the eye chamber. [1]

Orbit

The orbit, which surrounds the eye and is made up of portions of several skull bones to resemble a four-sided pyramid with its apex pointing back into the head, protects it from mechanical harm. A layer of orbital fat surrounds the eyeball and its functional muscles and functions somewhat as a cushion, allowing the eyeball to smoothly rotate about the rotational center, which is essentially a fixed point. Additionally, the eye is supported by canals and fissures that transmit blood vessels and nerves. [2]

1.1.1 Components of the Eye

To understand how diseases and other conditions can affect vision, it is essential to understand the anatomy and function of the various eye parts. Every component of the eye plays a distinct part in the visual process, and any interference with or harm to these structures can cause vision issues. New therapies and interventions can be created to stop or treat vision loss by better understanding the mechanisms of vision.

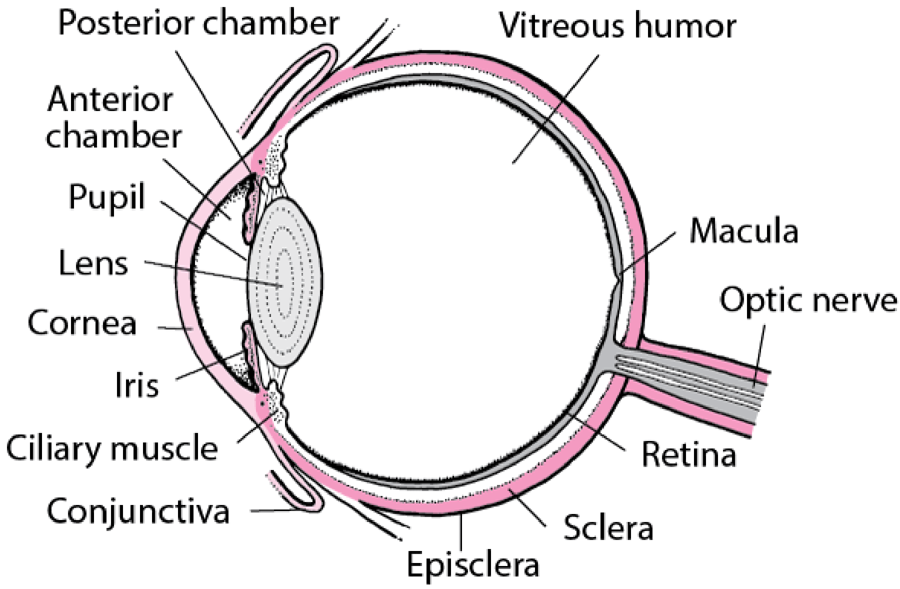

Fig. 1.1: Human eye anatomy sketch, taken from [3]

Sclera

The sclera, together with the cornea, forms the outer fibrous layer of the eye, which is a tough and sturdy covering of the eyeball. It provides as the site of attachment for the extraocular muscles and gives stability to the shape. Through its anterior portion, the cornea, the light rays enter the eye.

The sclera is a strong and tough fibrous membrane, mainly composed of lamellarly arranged bundles of collagen and elastic fibers, and occupies the posterior 5/6 of the eyeball. It provides a protection for the deeper eye structures. The sclera ranges in thickness from 0.3 to 1.5 mm, with the posterior portion being thicker and the weakest region being directly below the extraocular muscle insertions. Almost no blood vessels are found in the sclera. It combines with the cornea at the front part. The central retinal artery and vein enter the eye through this perforated disc (lamina cribrosa sclerae), which is located in the posterior segment of the sclera, medially to the optic disc. [1]

Cornea

The cornea is a convex, dome-shaped structure that bridges and closes a circular opening of approximately 12 mm in diameter at the anterior edge of the sclera. It occupies the anterior 1/6 of the eyeball. The cornea's thickness varies by about 1 mm, being slightly thicker at the edges than at the apex, which is about 0.5 mm thick. The cornea's transparency, or its permeability to light rays, is ensured by the arrangement of its individual layers. Thus, the cornea represents the entrance section of the so-called optical medium of the eye, and from the perspective of the refractive index of light, it is its most important component. The cornea contributes about +40 D to the total optical power of a healthy eye (+60 D). This value remains constant throughout life. The cornea is the most sensitive tissue and is densely innervated by sensory nerves. [1]

Choroid

The choroid (vascular layer), the ciliary body, and the iris make up the three components of the uvea. It lies beneath the surface layer and is relatively thin. The majority of its structure is made up of loose connective tissue with lots of pigment cells. It contains blood vessels that supply the majority of the eyeball. It also acts as an insulating layer against light and heat thanks to its pigment content and strong blood vessels. In some parts of it, smooth muscle cells are accumulated, which participate in regulating the amount of incoming light and changing the optical power of the lens (accommodation of the eye).

The choroid makes up the majority of the middle layer of the eye wall, taking up the back two-thirds of it. It appears as a black-brown membrane that is thin (0.2–0.4 mm) and filled with blood vessels. The sclera connects to its superficial side, which is separated from it by a thin layer of loose connective tissue through which numerous blood vessels pass. The smooth inner surface of the choroid is close to the retina. This part of the choroid is heavily pigmented and absorbs light rays, preventing reflections and over-illumination of the eye. It is formed by a network of capillaries. Thus, the choroid functions as a dark chamber for the optic layer of the retina, and the pigment cells of the retina, the cones and rods, are nourished by the capillary networks. The perforated disc corresponds to the opening at the back of the choroid through which the optic nerve fibers, retinal artery, and vein pass. The ciliary body and anterior choroid margin seamlessly converge. [1]

Ciliary Body

The ciliary body has the shape of a ridged ring that is attached to the inner side of the sclera. The ciliary body thickens and merges with the iris toward the front, while its posterior border thins and smoothly transitions into the choroid. On the inner side of the ciliary body, numerous two-type cilia are arranged, and besides other functions, they also serve a secretory function, meaning they secrete aqueous humor into the posterior chamber of the eye. The suspensory ligament fibers of the lens are attached to the ciliary body. The underlying structure of the ciliary body is a fibrous stroma, in which numerous smooth muscle cells are located, forming the musculus ciliaris. The muscle cells are arranged radially, longitudinally, and circularly. The musculus ciliaris forms a ring-like sphincter, which, upon contraction, releases the suspensory apparatus of the lens. The lens, due to its elasticity, changes shape and optical power. [1]

Iris

The iris forms the most anterior part of the middle layer of the eye wall. It has the shape of a disc with a centrally located opening called the pupil. The lateral edge of the iris merges into the ciliary body, while the medial edge encloses a circular opening, the pupil. The anterior chamber of the eye is located between the anterior surface of the iris and the posterior surface of the cornea. The anterior surface of the iris is variably colored depending on the amount of pigment, which also determines the color of the eye. The posterior chamber of the eye is located between the posterior surface of the iris and the anterior surface of the lens. The pigment layer of the retina (pars caeca retinae) extends onto the posterior surface of the iris, giving it a black coloration, and also extends to a small extent over the pupillary border of the iris onto its anterior surface, forming a narrow ring around the pupil. The iris serves as a light diaphragm in the eye. Depending on the amount of light, the retina adjusts its illumination through the pupillary reflex to achieve optimal values. The pupillary reflex is mediated by two systems of smooth muscle cells. The pupillary sphincter muscle (musculus sphincter pupillae), whose contraction causes the constriction of the pupil (miosis) and thus the reduction of the amount of light entering the eye, and the pupillary dilator muscle (musculus dilatator pupillae), which upon contraction dilates the pupil (mydriasis) and thereby allows more light to enter the eye. The color of the iris and the pattern on the front surface of the iris are used in genetics and computer systems for identification purposes. Unlike fingerprints, which serve the same purpose in identifying individuals, the iris does not change with age. [1]

Retina

The retina covers the eye's interior. It fills the entire eyeball cavity all the way to the iris' pupillary margin. Its inner surface is attached to the vitreous body, and its outer surface rests against the middle layer of the eye wall. The optic portion of the retina (pars optica retinae) and the blind portion of the retina (pars caeca retinae) are two structurally and functionally separate portions of the retina.

The optical portion of the retina is made up of a thin, brittle membrane that rests on the choroid. There are various structures on the back of the eye. The macula lutea is located lateral to the posterior pole of the eye. The fovea centralis, which is located in the center of the macula lutea and is the location of the sharpest vision, is located at the top of the optical axis of the eye. In the macula lutea, only tightly clustered cone-shaped photoreceptor cells are accumulated. Furthermore, there is a blind spot, which is located about 4 mm medially from the macula lutea. This is the point where the optic nerve exits the eyeball. The blind spot does not contain any photoreceptor cells. In addition to the optic nerve, veins that supply the surroundings, including the macula lutea, also enter the retina at this point.

The structure of the retina can be divided into two parts: the outer and inner layers. Along with the choroid, the outer layer functions as an insulating layer to absorb incoming light and stop it from reflecting inside the eye. The inner layer contains the primary sensory cells and neurons that collect information from the sensory cells and transmit it through the optic nerve from the eye to the brain. The inner layer of the eye is more complex, so a chapter on the optic pathway of the eye will further discuss it.

The pigment epithelium of the ciliary body and iris is another name for the blind spot of the retina, which is located in front of the ora serrata. It has no sensory cells and is only covered in a layer of pigment. [1]

Lens

The content of the eyeball consists of transparent and clear structures that transmit light rays and focus them onto the retina. These are the optical media of the eye, which consist of the lens (lens cristallina), vitreous body (corpus vitreum), and contents of the anterior and posterior chambers (camera bubli anterior et posterior), filled with aqueous humor (humor aquosus).

The lens is biconvex in design. The radius of curvature of its surfaces varies. Compared to the posterior surface, the anterior surface is less curved. While the posterior surface of the lens is located inside the vitreous body, the anterior surface of the lens borders the posterior chamber of the eye. The eye's diameter is approximately 9 to 10 mm. The lens's primary optical purpose is to transmit light and focus it on the retina. About 80% of the overall refraction is contributed by the cornea, and the lens fine-tunes the light's focus on the retina. Although the human lens is colorless at birth, there is a gradual increase in yellowish pigmentation with age. It has one of the highest protein concentrations of any tissue, which helps with this. During accommodation, the optical power of the eye changes as a result of the eye's dimensions changing. Adult eyes without accommodations have an optical power of about +20 D. [1] [4]

Vitreous and Aqueous Humors

The vitreous and aqueous humors, two fluid-like substances that fill the human eye, help to keep the eyeball's shape and ocular pressure stable. Aqueous humor is a water-like fluid that lies in front of the lens. The balance between these two processes controls the secretion and absorption of this low-viscosity fluid, which in turn controls the volume and pressure of the intraocular fluid. When secretion and reabsorption are out of balance, proper functioning is affected, and disease states may develop. The anterior chamber and the posterior chamber are the other two chambers in the eye in addition to the vitreous chamber. They are two slit-like spaces between the back surface of the cornea and the front surface of the lens. Both are filled with aqueous humor. The eye chambers communicate with each other through the pupil. [1] [5]

The fluid-like gel found in the vitreous chamber of the eye is called vitreous humor. It is made up primarily of water (98–99%) with traces of hyaluronic acid, glucose, anions, cations, ions, and collagen. The basal membrane, which is made up of supporting Müller cells of the retina and epithelial cells of the ciliary body, is where the vitreous humor is condensed toward the surface facing the retina. [1] [6]

1.1.2 Extraocular Muscles

The effector organ for the eyes' voluntary and reflexive movements is the extraocular muscles. Because the fovea, the region of high acuity vision, occupies a relatively narrow angle of visual space, controlling one's gaze requires very precise coordination of the extraocular muscles. The optical axes of both eyes converge forward to ensure that the image of the object a person is looking at falls on the macula of both eyes. The complex cooperation of both eyes is ensured by the activity of the central nervous system. Extraocular muscles are thin, strap-like muscles that often arise from the anulus tendineus communis, a common circular tendon. From this common origin, each muscle travels in a separate direction in the direction of the eyeball's surface, where it inserts through short, flat tendons. The rectus muscles (musculus recti) and oblique muscles (musculi obliqui) are two types of extraocular muscles. The extraocular muscles also comprise the levator palpebrae muscle. [1] [7]

Musculi Recti

The superior, medial, inferior, and lateral rectus muscles are among the four straight muscles of the eyeball. They all start from the same tendinous ring (anulus tendineus communis) and extend forward before joining the sclera in the eye's anterior portion with a brief, flat tendon that is 4-6 mm long. The superior and inferior rectus muscles rotate the eyeball upward and downward with a small medial movement, whereas the lateral and medial rectus muscles rotate the eyeball horizontally to their respective sides. [1]

Musculi Obliqui

The superior oblique muscle and the inferior oblique muscle are the two oblique muscles of the eye. Along with the rectus muscles, the superior oblique muscle originates from the anulus tendineus communis. The origin of the inferior oblique muscle sets it apart from the other extraocular muscles. It begins on the bottom portion of the orbit. When gazing forward, the superior oblique muscle turns the eye inward, then downward, and finally outward. On the other hand, when looking ahead, the inferior oblique muscle turns the eye outward before elevating and abducting it. [1]

Musculus Levator Palpebrae Superioris

Like the other extraocular muscles, the levator palpebrae muscle, commonly referred to as the eyelid lifter, begins at the common tendinous ring at the tip of the orbit. By tugging on the top tarsal plate, it lifts the upper eyelid. [1]

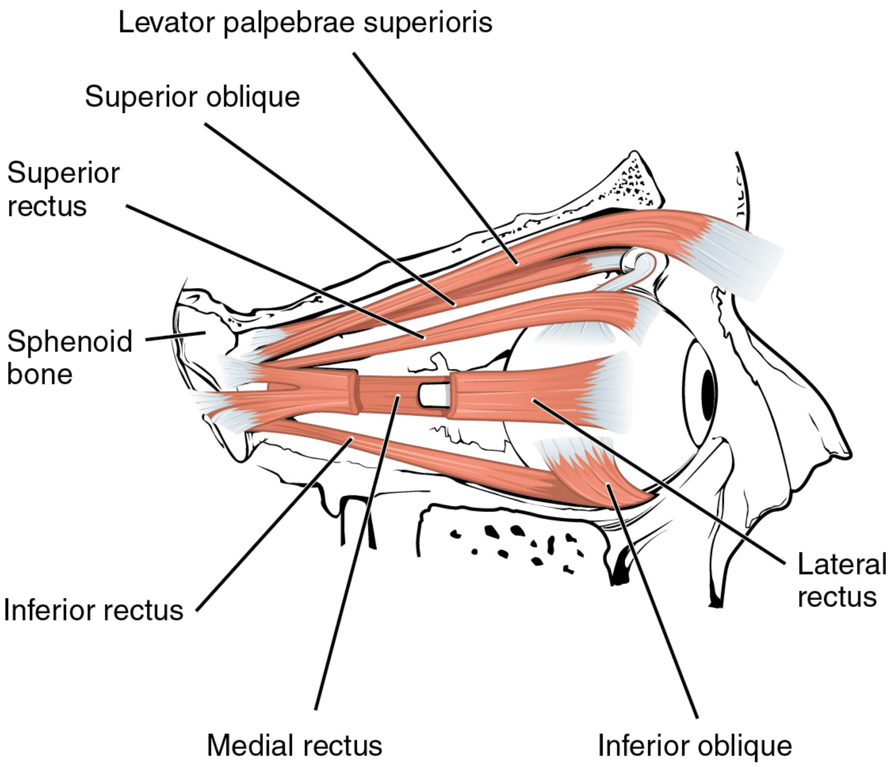

Fig. 1.2: Extraocular muscles of the eye, taken from [8]

1.2 Physiology of Vision

The visual system in humans is a sophisticated and intricate process that enables us to perceive the environment around us. The fundamental organs of vision, the eyes, collaborate with the brain to analyze and interpret visual data. From the collection of light by the eyes through the processing of this data by the brain, the process of vision entails numerous complex steps. It is essential to comprehend the physiology of vision in order to appreciate the complexity of this process and the different elements that can affect how we perceive the world.

Light and the eye

The eye's anatomy is built to give the neuronal retina a visual representation of the surrounding environment. The only transparent human bodily parts are the clear ocular media (cornea, aqueous humor, crystalline lens, and vitreous humor). The structure and arrangement of the live cells in the cornea and lens limit scatter, whereas other tissues greatly distort the transmitted light. The retina can also receive some near-infrared wavelengths in addition to visible wavelengths. The anterior components of the eye strongly absorb ultraviolet radiation, and only a very little amount (of the order of 1% or less) reaches the retina. [9]

Photoreceptors and Vision

The retina of the eye contains specialized cells called photoreceptors that react to light and give us the ability to see. Rods and cones are the two different types of photoreceptors. These light-sensitive cells are located next to the retinal pigment epithelium, a cell layer that is essential for photoreceptor survival, at the back of the retina. Rods are highly sensitive to light and operate under dim lighting conditions. Cones are able to function in ambient and bright lighting, react quickly to changes in light intensity, and are in charge of providing color vision and excellent visual acuity. 120 million rod and 6 million cone cells make up the human retina, with the latter being more prevalent in the centre or macula area of the retina. [10]

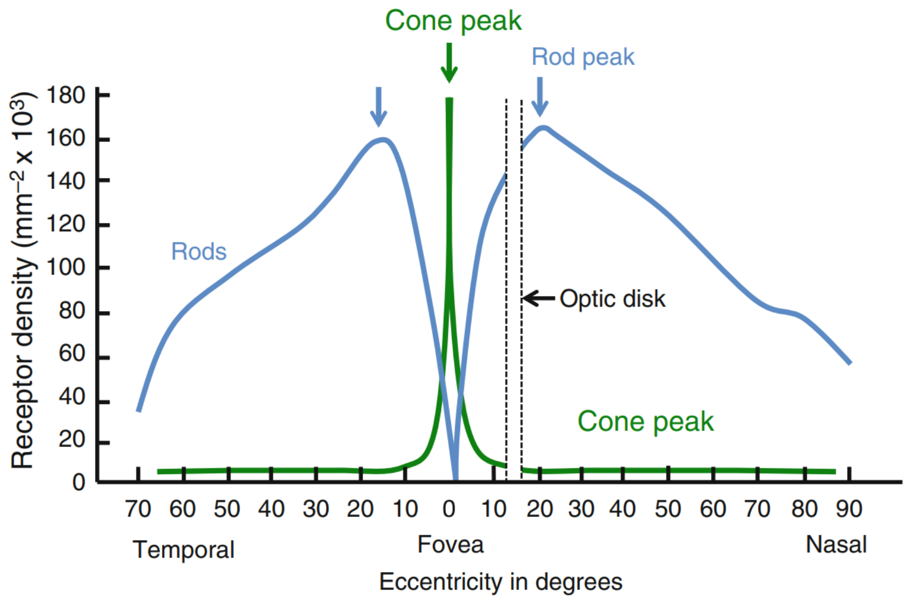

Fig. 1.3: Occurrence of rods and cones on the retina, taken from [11]

Visual Processing in the Brain

The principal area of the brain responsible for receiving, integrating, and processing visual information received from the retinas is known as the visual cortex. It is located in the most backward part of the brain, in the primary cerebral cortex's occipital lobe. On the basis of structure and function, the visual cortex is divided into five distinct regions (V1 to V5). The thalamus is the first place where visual information from the retinas that is flowing to the visual cortex passes through before synapsing in a region known as the lateral geniculate. The lateral geniculate then sends this information to V1, the primary visual cortical region. The calcarine sulcus is the geographic center of V1, also referred to as the primary visual cortex. [12]

The visual cortex in each hemisphere receives information from the opposing eye. In other words, information from the left eye is processed by the right cerebral areas, and vice versa. Receiving, dividing, and integrating visual data is the main function of the visual cortex. Other parts of the brain are then delivered the processed information from the visual cortex to be studied and used. This highly specialized technique enables the brain to instantly detect items and patterns without exerting a great deal of conscious effort. [12]

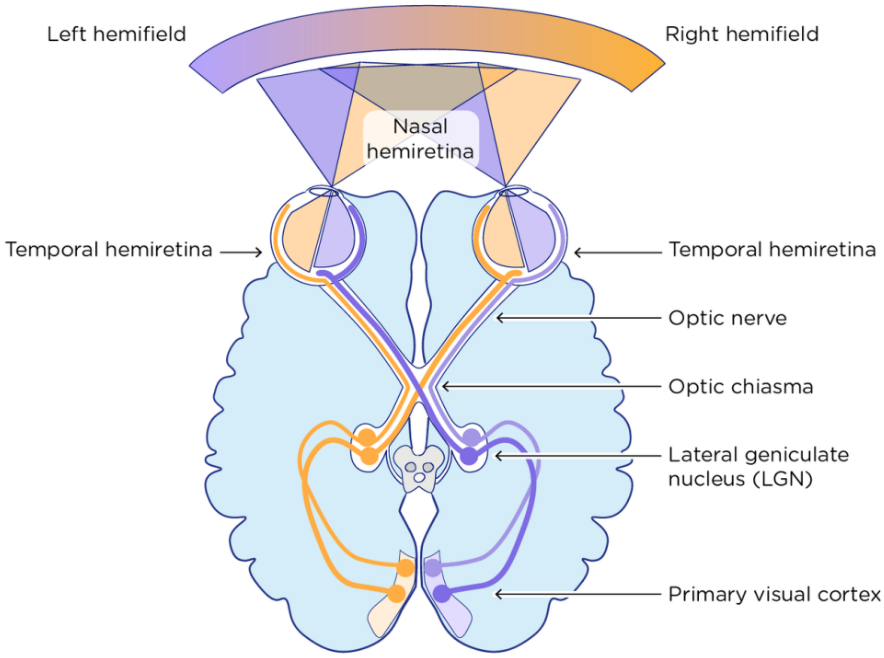

Fig. 1.4: Contralateral control of vision, taken from [13]

Color Vision

The retinal cone photoreceptors, which convert electromagnetic energy into electrical voltages, are where color vision begins. A complex network of retinal cells converts these voltages into action potentials. Three distinct color-opponent channels carrying the information are delivered from the lateral geniculate nucleus (LGN) to the visual brain. The perception of a wide range of distinct hues is made possible in the cortex by the combination of input from the three retino-geniculate channels. [14]

Visual Acuity and Contrast Sensitivity

Fundamental measurements of visual perception include contrast sensitivity and visual acuity. A Snellen chart or other comparable tests are frequently used to measure visual acuity, which is the capacity to discern minute details and see clear pictures. It is frequently utilized in clinical evaluations and serves as an essential marker of general visual function. On the other hand, contrast sensitivity gauges the capacity to recognize variations in contrast between bright and dark areas. It establishes one's capacity to recognize small color differences and separate items from their surroundings. The optical characteristics of the eye, neuronal processing in the visual system, and environmental variables all have an impact on visual acuity and contrast sensitivity. In order to diagnose and treat visual impairments, provide visual aids, and assess the success of visual treatments, it is crucial to comprehend these measurements. [1]

Binocular Vision and Depth Perception

Binocular vision is the capacity of not just humans to combine the visual information from both eyes to sense depth and three-dimensional information. It depends on the two eyes' overlapping visual areas, which allow them to merge their own images and add depth clues. Accurate depth perception—the capacity to assess the relative proximity and location of objects in the environment—is made possible by binocular vision. Stereopsis, in which the brain examines the minute variations in the pictures received by each eye to obtain depth information, is the principal process for depth perception in binocular vision. For activities like evaluating distances, recognizing the form of objects, and navigating in three-dimensional space, binocular depth perception is particularly crucial. Binocular vision problems can make it difficult to perceive depth, which can make it difficult to conduct daily tasks like driving and sports with proper spatial awareness. [1]

Visual Adaptation

The basic mechanism by which the visual system adapts to shifting environmental situations over time is called visual adaptation. Despite changes in illumination and input characteristics, it enables humans to perceive a constant and consistent visual experience. From the retina to the higher-order visual cortex, the process of visual adaptation takes place at several levels of the visual system. One of the most well-known instances of visual adaptation is the eyes' response to variations in brightness, such as when moving from a dark to a bright environment. Visual adaptation, in which the visual system adjusts to the dominant color environment to produce color constancy, is another important factor in color perception. The sense of motion, depth, and other visual qualities also depends on visual adaptation. Visual adaptation, in its broadest sense, makes sure that our visual system is adaptable and receptive to the constantly changing visual stimuli in our surroundings. [1]

1.3 Pathology and Defects of the Eye

The complex and sensitive human eye is a vital organ for vision. The structure and function of the eye can, unfortunately, be impacted by a wide range of conditions, which can result in different types of visual impairment. The ability of the eye to precisely track moving objects, which is essential for varies activities, may be impacted by some of these conditions. Some of these prevalent eye conditions that can affect visual perception and interfere with eye tracking will be discussed in this chapter. The complexity and fragility of the visual system can be better appreciated by comprehending how these conditions affect eye tracking.

These conditions can be either a challenge or a burden for an eye tracking system. For example, when detecting the pupil for identification purposes, these conditions could make it more difficult to perform accurate tracking. However, on the other hand, eye tracking systems can be focused specifically to detect these conditions and subsequently analyze them. With the proper attention and customization, eye tracking systems can be designed to work effectively even in the presence of these pathologies.

Refractive Errors

One of the main causes of blindness and a significant factor in the burden of eye diseases worldwide are uncorrected refractive errors. Myopia, hyperopia, and astigmatism are the most common categories of refractive errors in children and young adults; older adults may also have these conditions, with presbyopia adding another risk factor for vision loss in this population. [15]

Myopia

The myopia (nearsightedness) is a refractive error that causes distant objects to appear blurry while close objects remain clear. Refractive errors in the cornea, lenses, or a combination of the two may result in refractive myopia. Axial myopia is a form of nearsightedness where the elongation of the eye results in an increase in the axial length of the eye. Myopia can be treated with contact lens or glasses, but if untreated, it can result in other vision issues. Nearsightedness can change the size and shape of the eye, which can affect the accuracy of gaze tracking, which may have an impact on eye tracking systems. [16]

Hyperopia

Because the focal plane of light entering the eye in hyperopia (farsightedness) is posterior to the retina, it causes objects farther away to appear sharper and up-close objects to appear blurrier. This farsightedness in refractive hyperopia may be brought on by refractive errors in the cornea, lens, or both. A short axial length for the eye causes the farsightedness in axial hyperopia. Refractive correction, such as glasses or contact lenses, is another method of treatment. [16]

Presbyopia may or may not be compared to the hyperopia, as they both have similar conditions, but they differ in a way of development and treatment. The hyperopia occurs due to misshapen the cornea or the eyeball is short, but the presbyopia occurs as the muscles of the eye become less elastic with age. Thus the presbyopia is an aging condition.

Astigmatism

Astigmatism is a refractive error that results from corneal and/or lens optical aberrations that deviate from a refractive configuration that is perfectly spherical. Additionally, astigmatism is frequently present in many connective tissue diseases. Because of this, it might be challenging for the eye tracking system to detect eye movements precisely. Myopia or hyperopia-correcting glasses or contact lenses can be used to treat the condition. [16]

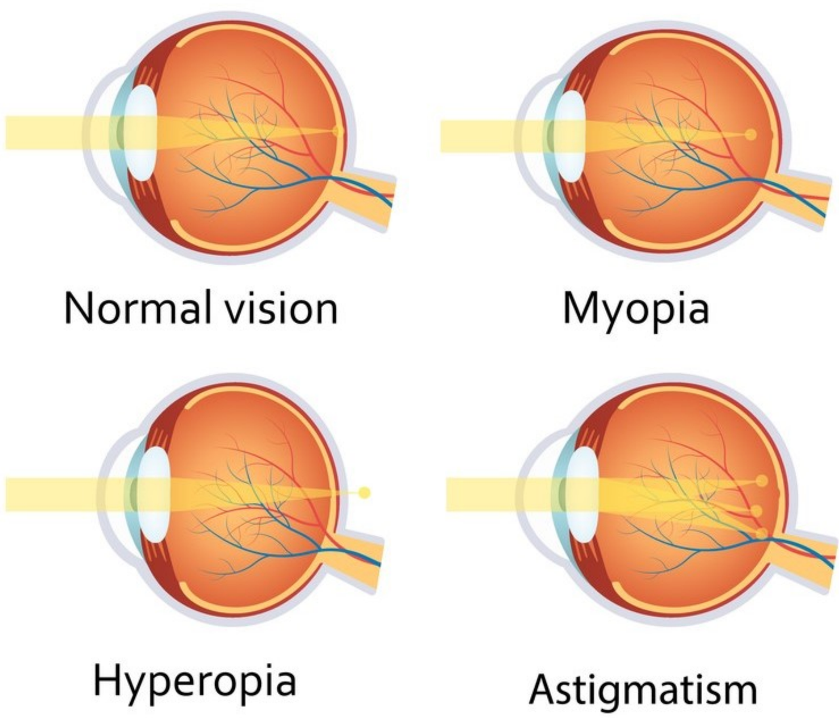

Fig. 1.5: Emmetropia and Refractive Errors, taken from [17]

Cataracts

When the eye's lens develops a cataract, transparency is lost as a result of the lens becoming cloudy. Based on their causes, which can be age-related, pediatric, or secondary to other factors, the different types of cataracts are categorized. One of the main causes of lens opacity is oxidative stress. According to where the cloudiness develops within the lens, age-related cataracts can be classified as nuclear, cortical, or posterior subcapsular. Cloudiness that develops at or soon after birth is referred to as congenital and infantile cataracts, and it may be unilateral or bilateral. Drugs can also cause cataracts; corticosteroid use over a long period of time is strongly linked to the development of posterior subcapsular cataracts. Other causes of cataracts include physical and chemical injury, as well as exposure to various forms of radiation. [18]

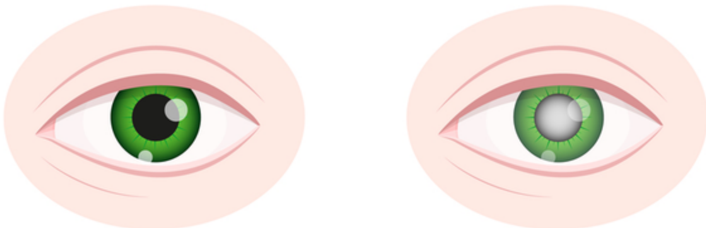

Fig. 1.6: Healthy eye (left) and a developed cataract (right), taken from [19]

Glaucoma

Glaucoma is a group of disorders that can cause damage to the optic nerve and loss of vision. The damage is often caused by increased pressure inside the eye, known as intraocular pressure. Glaucoma encompasses a range of conditions that share a common feature: the potential to progress and cause vision loss. The clinical hallmark of glaucoma is a distinct type of optic neuropathy called glaucomatous optic neuropathy. The most common type of glaucoma is primary open-angle glaucoma, which causes changes to the optic disc and loss of peripheral vision. As the disease progresses, it can also cause loss of visual acuity and even blindness. There are two major categories of glaucoma based on the appearance of the drainage pathway in the eye. In open-angle glaucoma, the drainage is restricted despite its normal appearance, while closed-angle glaucoma is caused by tissue physically blocking the drainage angle. [20] [21]

Macular Degeneration

Age-related macular degeneration is a chronic and progressive disease that affects the central retina, caused by a combination of genetic and environmental factors, and is one of the leading causes of vision loss globally. The disease usually leads to visual impairment in its later stages due to either neovascular or geographic atrophy processes. In neovascular age-related macular degeneration, the choroidal neovascularization breaks through to the neural retina, causing fluid, lipids, and blood leakage and subsequent fibrous scarring. In geographic atrophy, progressive atrophy of the retinal pigment epithelium, choriocapillaris, and photoreceptors occurs. The most severe visual loss from age-related macular degeneration occurs in the advanced stages of these forms of the disease. [22]

Diabetic Retinopathy

A significant contributor to blindness is diabetic retinopathy (DR), a microvascular complication of diabetes. Over 60% of people with type 2 diabetes and nearly all people with type 1 diabetes develop some form of retinopathy after 20 years of having diabetes, making DR more likely as time goes on. Non-proliferative and proliferative DR stages can be distinguished from one another.

Microaneurysms and retinal hemorrhages are the nonproliferative DR's first detectable symptoms. The development of cotton-wool spots, venous beading, and intraretinal microvascular abnormalities are all symptoms of progressive capillary nonperfusion. Further retinal ischemia leads to proliferative DR, which is characterized by the development of new blood vessels on the retina's or the optic disc's surface. These abnormal blood vessels have the potential to bleed, which could cause vitreous hemorrhage, follow-up fibrosis, and tractional retinal detachment. The development of hard exudates at the central retina and increased vascular permeability are the two main features of diabetic macular edema (DME), which can happen at any stage of DR. The main cause of vision loss in diabetics today is diabetic macular edema. While secondary interventions, like laser photocoagulation, may stop the further progression of DR and vision loss, primary interventions, like intensive glycemic and blood pressure control, can lower the incidence of DR. [23]

Retinal Detachment

The condition known as retinal detachment causes vision loss when the neurosensory retina separates from the underlying retinal pigment epithelium. The retina's close proximity to the RPE is necessary for normal retinal function. Rhegmatogenous retinal detachment is the most typical type, and it happens when vitreous fluid leaks into the subretinal space through a tear in the neurosensory retina. Trauma, inflammation, and other factors can all contribute to retinal breaks. Chronic detachment can also happen silently. It is imperative to seek treatment right away, which may entail surgery, to prevent irreversible vision loss. [24]

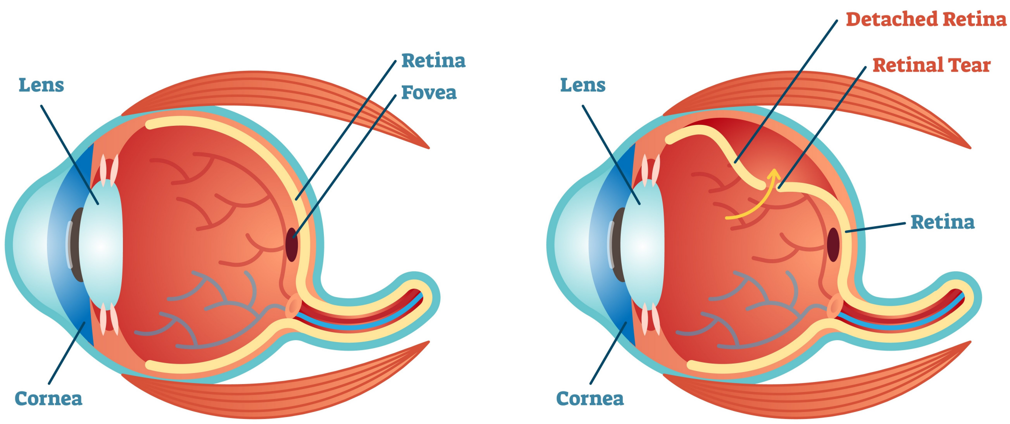

Fig. 1.7: Healthy eye (left) and a retinal detachment (right), taken from [25]

1.4 Oculomotor Events

Any type of eye movement or activity involving the eye muscles is referred to as an oculomotor event. These occurrences include both voluntary and involuntary eye movements, such as fixational eye movements, smooth pursuit movements, vestibular-ocular reflexes, and optokinetic nystagmus. These movements are crucial for a number of visual tasks, including tracking a moving object, maintaining steady vision while turning your head, and switching your focus between various visual targets. Vision issues, including diplopia (double vision), blurred vision, and trouble moving the eyes, can be caused by abnormalities in oculomotor events. [27]

1.4.1 Types of Eye Movements

For the purpose of creating precise and dependable eye tracking systems, it is essential to comprehend the many kinds of eye movements and the underlying mechanics that underlie them. Eye movements can be used to diagnose visual and neurological diseases and provide important information about visual attention, cognitive functions, and motor control. A full knowledge of the underlying mechanics of these motions is also necessary for appropriate diagnosis and treatment of eye movement abnormalities.

Saccades

Rapid eye movements, known as saccades or saccadic movements, are used to move the fovea to a different spot in the visual environment. Saccades are quick eye movements that are used to rapidly scan a visual scene. The speed of saccades can range anywhere from 200 to 600 degrees per second on average, with larger saccades reaching speeds of up to 900 degrees per second. The typical range of saccades in tasks is typically between 15 and 30 degrees, and the typical frequency of saccades is between 2 and 5. The fovea is moved to a new location in the visual environment with each saccade, shifting the focus to a different area of the visual scene. Saddadic eye movements can be voluntarily made, triggered by visual stimuli in a reflexive manner, or used as a corrective motion linked to optokinetic or vestibular movement. [26] [27]

Fixation

Fixations are eye movements that stabilize the retina over a stationary object of interest. Tremor, drift, and microsaccades are the minuscule eye movements that define fixations. Fixation is defined as a brief period of time during which a person's visual gaze is fixed in one location. The tracked object's image appears on the retina's fovea when the gaze is fixed. This vision has a very sharp field of view that encompasses about one degree of visual angle. The following object's image is held roughly stable on the retina during each fixation. Eye movements vision information can be basically interpreted as an alternating sequence of saccadic movements and fixations. [26] [27]

Smooth Pursuit

Pursuit movements are involved when visually tracking a moving target. When the object of interest is slowly moving, it is possible to keep it more or less stable in foveal vision by making smooth pursuit eye movements. This pursuit system matches the eye velocity to the target velocity. The threshold of velocity at which the slowly moving object can be tracked is around 5–30 degrees per second, and if the target is moving faster so that its velocity is above the threshold, saccades are then used to compensate for the lag. [26] [27]

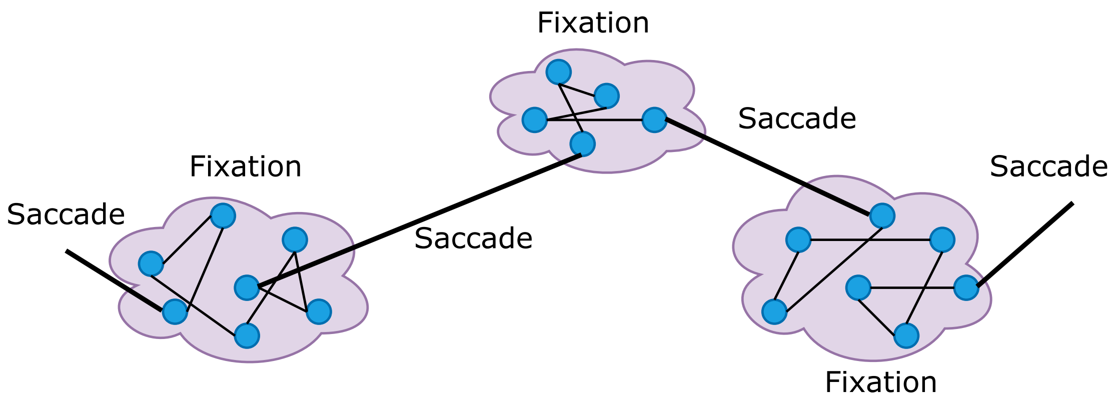

Fig. 1.8: Fixations and Saccadic movements

Vergence

Up until this point, the movements described are all referred to as conjugate eye movements, meaning that the eyes move in the same direction to fixate an object. Fortunately, we have the option to fixate on objects in various depth planes while maintaining binocular vision with opposing eye movements. Vergence eye movements are disconjugate. Ocular vergence is the result of these eyes moving in directions that are simultaneously opposite. Both convergence and divergence can result from these vergence movements, which can happen in either direction. Focusing from far to near causes convergent movements, while focusing from far to near causes divergent movements. [28]

Fixational Eye Movements

Although the eyes appear to be steady when they are focused on a goal, they are actually moving, although this movement cannot be seen by looking at the eyes. There are three different types of tiny movements that make up the fixational eye motions category. Microsaccades, microtremor, and microdrift are these. [29]

Microsaccades are miniature replicas of the rapid gaze shifts (saccades) that humans normally use to explore a visual scene. These movements have drawn a lot of attention since they are easier to identify than other fixational eye movements. For these reasons, when fixational instability is discussed, microsaccades are frequently the first thing that comes to mind. Microsaccades may refresh images to actually prevent the fading when the eyes are immobile. They are spatially random, varying over 1 to 2 minutes of arc in amplitude. [26] [29]

A tremor is an irregular, wavelike eye movement that has a low amplitude (around the diameter of the foveal cone) and a high frequency (about 90 Hz).

Tremors are difficult to accurately record since they fall within the range of recording noise due to their nature. It is possible to maintain visual sharpness during extended fixations thanks to tremors. [26]

Low-frequency eye movements called drifts happen in between saccades and microsaccades. A fixated object's retinal picture drifts across photoreceptors throughout this process. Drifts play a compensating role in sustaining precise visual fixation; they happen either when microsaccades are absent or when their compensation is insufficient. A drift typically has a dimension of 6 arc and travels at a speed of 1 arc/sec. [26]

Blinks

The eyelids effectively close and open again when a blink. Although they are simple, blinks are frequently utilized in eye tracking studies. The blink becomes a voluntary blink or a wink when it results from a voluntary activity. Non-voluntary blinks come in two varieties: reflexive blinks and spontaneous blinks. External stimuli can cause reflexive blinks as a kind of defense, whereas any unintentional blink that does not fit into one of these categories is referred to as spontaneous. Despite being used as a form of interaction, winks or voluntary blinks are not frequently accepted as a metric. In contrast, involuntary blinks signify a person's mood or a response to a stimulus. Given their linkage with one's internal state, spontaneous blinks, as opposed to involuntary blinks, are the most frequently employed type of blink as a measure. [26]

1.4.2 Visual Impairments and Eye Movement Disorders

Oculomotor system failure, which regulates eye movement and alignment, can lead to visual impairments and eye movement disorders. A variety of eye movements, such as smooth pursuit and saccades, are made possible by the oculomotor system, which is made up of a complex network of muscles, nerves, and brain areas. Any interference with this system can cause visual impairments such diplopia (double vision), fuzzy vision, and issues with reading and visual tracking. Eye movement abnormalities can have a substantial influence on quality of life by making it difficult to navigate the surroundings and carry out daily tasks.

Strabismus

The condition of strabismus involves improper eye alignment. One eye looks in a different direction than the other as a result of an imbalance in the muscles that govern eye movement. Double vision, trouble with depth perception, and occasionally amblyopia (lazy eye) can result from this. Strabismus comes in a variety of forms, such as esotropia, exotropia, hypertropia, and hypotropia. One eye moves inward, toward the nose, in esotropia, a kind of strabismus, while the other eye remains straight. Exotropia is the reverse, when one eye turns away from the nose and is straight while the other eye is turned outward. When one eye is higher or lower than the other, it is said to have hypertropia or hypotropia, respectively. Congenital strabismus is present at birth, whereas acquired strabismus appears later in life. Depending on how severe the disease is, it may also be intermittent or ongoing. [30]

Amblyopia

Amblyopia is a condition in which the processing of visual information is dysfunctional. Although the defects comprise many different forms of visual function, this dysfunction is often identified and obvious as decreased recognizing visual acuity. Amblyopia is a condition that develops when the retinal image deteriorates during a crucial time for visual development. The sensitive time during which therapy is possible could not coincide with the sensitive period during which amblyopia develops. The illness is a symptom of another clinical process, not its root cause. Amblyopia can also be considered to be the outcome of usage owing to aberrant binocular interaction (strabismic) or disuse because there is no distinct image on the retina (anisometropia or deprivation). Amblyopia therefore never happens by itself.[31]

Nystagmus

Nystagmus refers to repetitive, to-and-fro movements of the eyes that are initiated by slow phases. Different types of nystagmus can be distinguished based on the presence of fast and slow phases or alternating slow phases. In the assessment of nystagmus, a structured approach is essential, as the differential diagnosis of nystagmus is very broad. Nystagmus comes in a variety of forms, which are categorized according to their underlying causes and characteristics. There are, for example, types such as pendular nystagmus, acquired nystagmus, and others. Nystagmus's particular type and characteristics can reveal crucial details about its underlying causes and may have an impact on available treatments. This includes both benign peripheral-vestibular disorders and dangerous, potentially life-threatening central causes. [32]

Oculomotor apraxia

Although some eye movements can be preserved, patients with ocular motor apraxia have absent or significantly delayed voluntary eye movements, primarily saccades. One phenotype is distinguished by a lack of or a significant impairment in the ability to perform voluntary horizontal and vertical gaze shifting, while the slow and rapid phases of vestibular nystagmus are preserved. It is a symptom of adult-onset neurodegenerative disorders and acute brain injuries, and it is brought on by a malfunction in the cortical (and basal ganglia) control of voluntary eye movements. The congenital type is characterized by head thrust, severely defective fast phases of nystagmus, and very hypometric staircase saccades. [33]

2 Eye Tracking

Eye tracking is a technique that assists in the study of visual attention. It can determine where users are looking at a given moment in time, how long they are looking at something, and the path their eyes take using eye tracking. Numerous fields, including human factors, cognitive psychology, marketing, and the broad area of human-computer interaction, have used eye tracking. Eye tracking aids in the understanding of the entire user experience, even that which users are unable to articulate. [34]

Brief History Overview of Eye Tracking Development

Since the earliest days of measuring cognitive and physiological activities, our understanding of the human mind has remained comparatively archaic. Significant technological strides have only been introduced in the last half of the decade that have made it possible to precisely investigate visual perception and to record and show cognitive processes. [34]

The most fundamental theories of how the brain and visual system interacted were the focus of early eye-tracking studies. The research was frequently highly academic and too expensive and complex to be used for commercial purposes. Systems that track eye movement using film recordings of the eye were developed by the 1940s. In order to better understand how pilots' eyes move while using cockpit controls and instruments to land an aircraft, Paul Fitts and his colleagues started using cameras in 1947. The objective of this early eye-tracking usability study was to systematically observe how users interacted with an interface in order to enhance the system's design. [35]

A new generation of eye trackers was created as a result of the development of the video-based eye tracker in the 1960s and 1970s, and new applications for eye tracking became possible. Unfortunately, even though researchers had easier access to the technology, the participants still had to put up with a very intrusive device that required a head restraint and bite bar. The contemporary eye tracker, which is used by many researchers in the UX sector today, was invented in the late 1990s. The eye tracker was able to leave the academic setting and be implemented into commercial user experience laboratories because to new developments in hardware and software design. [34]

Overall, eye tracking is a flexible tool that provides important insights into how people behave, think, and interact with their environment across a variety of disciplines. Its applications keep growing, giving scholars, experts, and professionals useful information for improving functionality, user experiences, and general comprehension of human behavior.

2.1 Methodologies

For academics and practitioners interested in using eye tracking in their study, understanding the various eye tracking approaches is crucial. Each approach has certain benefits and is appropriate for particular research problems and experimental settings. Researchers may get precise and insightful data to study visual attention, cognitive functions, and human behavior in a variety of domains by choosing the right eye tracking technology.

2.1.1 Eye Tracking Techniques

The monitoring and analysis of eye movements and gaze behavior are made possible by the effective research technique known as eye tracking. It offers insightful information on user behavior, cognitive processes, and visual attention. The main approaches utilized in eye tracking research will be introduced in this part along with a discussion of their benefits, drawbacks, and uses.

Remote Tracking

Remote eye tracking is a non-invasive method that eliminates the requirement for direct physical contact with the participant's eyes in order to monitor and analyze eye movements and gaze behavior. It includes remotely recording and following the movement of the eyes using cameras or other sensors. This method eliminates the need for large and intrusive devices, making it a practical and pleasant approach to gather eye-tracking data. In order to gather insights on visual attention, cognitive processes, user behavior, and other topics, remote eye tracking has been widely employed in a variety of domains, including psychology, human-computer interaction, market research, and virtual reality. It has created new options for researching human visual perception and behavior and offers a flexible and adaptable method for observing eye movements in real-world settings. [27]

Head-Mounted Eye Tracking

Head-mounted eye tracking is a method in which the eye-tracking equipment is worn on the head, usually in the form of personalized glasses or a headset. This makes it possible to quantify eye movements and gaze behavior in a subtle and lifelike way. The wearable gadget has integrated eye-tracking sensors that record the participant's eye movements in real time. As the sensors are near to the eyes, head-mounted eye tracking delivers precise and accurate data. Studies of visual attention, user interactions with computer interfaces, and immersive experiences like virtual reality can all benefit from it. This approach has applications in the study of user experience, psychology, neurology, and human-computer interface. [37]

Mobile Eye Tracking

Utilizing wearable, portable sensors, cutting-edge technology called mobile eye tracking makes it possible to record and analyze eye movements. Insights into human behavior can be gained by allowing researchers to examine visual attention and gaze behavior in actual environments. Mobile eye tracking provides a more valid way to comprehending how people interact with their surroundings by capturing eye movements in natural settings. The configuration and environment of data collection are the primary differences between mobile and remote eye tracking. Mobile eye tracking makes it possible to examine eye movements in actual situations, giving researchers a better understanding of how people interact with their surroundings in realistic circumstances. On the other hand, remote eye tracking is more frequently utilized in controlled laboratory studies where accurate measurements of eye movements are needed. [27] [38]

Binocular Eye Tracking

Eye movements in both eyes may be tracked and analyzed concurrently using a technique called binocular eye tracking. It reveals details about how the eyes cooperate to process visual data. This approach is useful for investigating binocular vision and eye coordination in a variety of disciplines, including psychology, cognitive science, and clinical research. Researchers can find patterns and problems in binocular vision by examining characteristics like vergence eye movements and binocular fixations, which helps in the creation of diagnostic tools and rehabilitation methods. [27]

2.1.2 Principles of Eye Tracking

Eye tracking is a technique for monitoring and examining gaze patterns and eye movements. It is based on the fundamental ideas of recording and deciphering eye movements. These concepts can help us better understand where and how people focus their visual attention. The fundamentals of eye tracking include a number of methods and tools that enable precise observation and recording of eye movements.

Corneal Reflection

The corneal reflection method is predicated on the idea that, due to the spherical nature of the eyeball, a light source reflected off the cornea would remain mostly fixed as the eye moves. It is then feasible to determine the direction of the look by relating the locations of these light reflections to the pupil's center. The person needs to be unable to see the lights shining on their cornea; otherwise, they will draw attention to themselves. The illumination light fall into the infrared region of the spectrum as a result. An early iteration of this technique involved capturing infrared film of newborns' eyes and analyzing frame-by-frame the locations of the reflections and the center of the eyes. Since then, corneal reflection-based eye tracking has become more popular. With more powerful computers, the painstaking measurements of reflection positions have been automated, and a number of companies are now selling eye trackers that provide direct gaze points. All of them look at how infrared light reflects off the eyes by using algorithms to process pictures. This is what makes them all similar. Each system is outfitted with infrared light sources (ranging in wavelength from 750 to 1,400 nm) and a camera that records the light as it reflects on the cornea, sclera, iris, and retina. [39]

Video-based Tracking

Digital video cameras are used in the non-invasive technique known as video-based eye tracking to record and examine eye movements. It entails filming the participant's eyes while they are performing an activity, then utilizing specific software to analyze the video and extract data about the participant's gaze behavior. The participant's eyes are often lit with infrared light during video-based eye tracking, which is invisible to the human eye but detectable by the camera. For precise tracking, the infrared light improves the visibility of the pupil and other important eye characteristics. The eye tracking system for videos use complex algorithms to recognize and monitor the location of the eyes, pupil, and other eye attributes in real-time or after analysis. The device can follow the pupil's movement in order to calculate the direction and length of gaze fixations, saccades (fast eye movements), and other eye movement characteristics. Using computer vision methods, such as image processing, pattern recognition, and machine learning algorithms, to detect and track the eyes reliably and correctly is a typical method in video-based eye tracking.

For identifying and tracking characteristics like the pupil center, iris shape, or eye corners, the system may use a variety of techniques. A number of benefits come with video-based eye tracking, including non-intrusiveness, high spatial resolution, and the capacity to record a variety of eye motions and behaviors. It is frequently used in psychological research and applications as well as in market research, usability testing, human-computer interface, and other domains where comprehending visual attention and gaze behavior is important. The possible drawbacks of video-based eye tracking, should be taken into account. These include sensitivity to lighting conditions, head motions, and occlusions of the eyes. To guarantee accurate and trustworthy eye movement data, proper calibration and validation processes are required. [41] [43]

Electrooculography

A non-invasive eye tracking method called electrooculography (EOG) detects the electrical potential difference between electrodes positioned on the skin surrounding the eyes. It is predicated on the idea that the electrical potentials of the cornea and retina are different, and that eye movements cause electrical changes that may be detected. The eyes are frequently surrounded by two or more electrodes during an EOG, typically above and below one eye or on the outer corners of both eyes. The eye movements may be seen and evaluated by monitoring the difference of electrical potentials between these electrodes. A voltage gradient is produced when the location of the positive and negative charges on the cornea and retina changes when the eyes move. The electrodes record these voltage variations and turn them into electrical impulses. Saccades (fast movements between fixation locations) and gradual drifts are two types of eye movements that generate changes in the eye's resting potential, and EOG is particularly sensitive to these changes.

Compared to other approaches, it is less precise at detecting minute and delicate eye movements, but it is easier to set up and less disruptive for the participants. EOG has applications in several disciplines, including sleep research, human-computer interface, clinical evaluations of eye movement abnormalities, and cognitive neuroscience. It offers insightful data on the time and pattern of eye movements, which may be utilized to comprehend reading habits, visual attention, and other oculomotor activities. [40] [41]

The Magnetic Search Coil Technique

The magnetic search coil technique, also known as the scleral search coil method is an eye tracking approach that uses finely detailed measurements and recordings of eye movements. It entails inserting a tiny coil of wire into the sclera, the eye's white outer layer. The limbus, or point where the cornea and sclera converge, is often the location where the coil is put. The coil interacts with an external magnetic field as the eye moves, producing electrical impulses that can be recognized and measured. Real-time data regarding the eye's position and rotation is provided by these signals. For its precision and dependability in catching even the slightest eye movements, the Scleral Search Coil technique is well-known. It is frequently employed in investigations of visual perception, oculomotor control, and cognitive processes, among others, that need for accurate assessment of eye location and gaze direction. Despite the fact that the scleral search coil approach needs surgical coil insertion, it gives extremely precise measurements. This restricts its application to clinical investigations and research settings where the advantages exceed the procedure's invasiveness. [42] [43]

Infrared Oculography

Measurement of the intensity of reflected infrared light is the foundation of the eye tracking technique known as infrared oculography. With this method, the sclera's reflected light is measured after infrared light has irradiated the eye. Information regarding variations in eye position may be gleaned from differences in the quantities of infrared light that are reflected. This technology is intrusive because a light source and sensors may be attached to spherical glasses for implementation. There are restrictions on the range of eye movements that infrared oculography can measure. It can often detect eye movements of up to 20 degrees vertically and 35 degrees horizontally. Infrared oculography has a particular purpose in research systems employed in magnetic resonance imaging (MRI) tests. These technologies enable measurements of eye movement in low-light conditions. Using image processing tools, infrared oculography is also used for gaze engagement. Overall, the use of infrared oculography in diverse academic and clinical contexts offers special benefits, including the capacity to record eye movements in complete darkness. [44]

Optical Coherence Tomography

A non-invasive, high-resolution optical imaging technique called optical coherence tomography (OCT) relies on interference between signals coming from the item under study and a nearby reference signal. A cross-section image of the object, or a two-dimensional image in space (lateral coordinate, axial coordinate), may be created via OCT in real time. Confocal microscopy's ability to scan the retina of the human eye is constrained by the combination of the eye's low numerical aperture and aberration. The numerical aperture of the microscope objective affects lateral and axial resolutions in confocal microscopy. The retina of the human eye may be scanned with at least 100 times greater axial resolution than that possible with confocal microscopy since with OCT, the axial resolution is mostly governed by the light source. Even when employing tiny diameter beams and imaging aberrated eyes, such a high depth resolution is still possible. Decoupling the depth resolution that can be achieved from the value of the microscope objective's numerical aperture was a significant factor in the creation of OCT. [45]

Pupil-corneal reflex tracking

In the pupil-corneal reflex tracking method, an infrared light source is used, as it can improve the contrast between the edge of the pupil and the iris as well as create a reflection on the inner and outer surfaces of the cornea and lens. When light from the outside enters the pupil of the eye, the surface of the cornea reflects and forms light. When the eye is moved, the angle of light entering the eye changes, but the position of the light spot does not change, so this is used as the reference point. On the contrary, the center of the pupil will always follow the direction of the gaze. The vector made up of the center of the pupil and the bright spot of the cornea changes with eye movement and can be used to analyze the direction of vision. The advantage of this method is that it is relatively accurate, is easy to measure with, and is by far the most widely used method. [41]

Dual-Purkinje-Image Method

The Purkinje image is formed by the reflection of several optical interfaces of the eye. The first Purkinje image is the image reflected from the anterior surface of the cornea. The second Purkinje image is reflected from the posterior surface of the cornea, and the reflection of the second image is weaker than the reflection of the first one. The two are almost identical. The next one, the third Purkinje image, is reflected from the anterior surface of the lens, and the image reflected from the posterior surface of the lens is called the fourth Purkinje image. If the eye is shifted, the first and fourth Purkinje images move together for exactly the same distance, but if the eye is rotated, the two images will move a different distance because the curvature centers of the two mirrors are not the same. Thus, the eye rotation can be accurately measured by monitoring the spatial separation of the two Purkinje images. To use this method correctly, the head of the participant needs to be fixed. Another disadvantage is that the second, third, and fourth images are very weak. The price is high, as the laboratory requirements are really high and this method is not always universal. [41]

Machine Learning-based Tracking

Machine learning algorithms are used to evaluate eye movement data in an eye tracking system. Based on the patterns and characteristics that are retrieved from the data, predictions or classifications are then made. Large datasets of labeled eye movement data may be used to train machine learning algorithms, where the output labels correspond to particular eye movement occurrences or states and the input features are recorded eye movement measurements. These labeled samples are used by the algorithms to train models that may be applied to unforeseen eye movement data. Following that, the machine learning models may be used to forecast a variety of eye movement characteristics, including gaze direction, fixation length, saccade amplitude, and even higher-level cognitive states like levels of concentration or weariness. These hypotheses can shed light on how people behave, make decisions, and perceive the world. Extraction of pertinent characteristics, such as pupil size, gaze coordinates, velocity, or acceleration, from the raw eye movement data is a typical strategy in machine learning-based eye tracking. The machine learning algorithms learn to translate the input features to the desired output labels using these characteristics as input. The precise job at hand and the properties of the eye movement data determine the machine learning method to use. Support vector machines, random forests, convolutional neural networks, recurrent neural networks, and deep learning architectures are common techniques used in eye tracking systems. Eye tracking systems based on machine learning provide a number of benefits, such as the capacity to manage intricate and non-linear correlations in eye movement data, the flexibility to accommodate individual variances, and the possibility for real-time or online processing. They can also be utilized for customized eye tracking, in which case the system gradually learns and adjusts to a user's particular eye movement patterns. [41]

2.2 Measurements

Eye movement measurements are the processes and procedures applied to quantify and examine different facets of eye movements. These measures offer important new understandings of visual attention, cognitive functions, and interactions between people and their surroundings. Researchers and practitioners can better understand human behavior, perception, and decision-making by recording and analyzing eye movement data. Eye movement measures come in a variety of typical forms that are often utilized in both academic and practical contexts. Individual eye movements are already described in detail in section 1.4.1 Types of Eye Movements.

2.2.1 Eye Movement Metrics and Parameters

Quantitative measurements are utilized to study and explain different features of eye movements using eye movement metrics and parameters. These measures offer insightful information on the dynamics and traits of eye activity.

Gaze Point

The gaze point designates where the eyes are fixed or focused on a screen or other visual input. The coordinates of the point in a two-dimensional space are commonly used to define it. The gaze point, a fundamental eye movement metric, indicates where a person is fixating or staring when they are seeing a visual input or scene. It speaks about the precise location in space where the line of sight meets the outside world. Eye movement research places a high value on analyzing and comprehending gaze locations because they provide important details about cognitive functions and behavior. Here are some critical elements pertaining to the significance of gaze points. [46]

Fixation

Fixation duration is the amount of time that the eyes are generally steady and concentrated on a certain place or area of interest. It offers details on cognitive involvement with a specific stimuli and attentional processing.

Researchers used films to transmit both positive and negative feelings, and they found that the two emotional states analyzed differed considerably in terms of the quantity and length of fixations. In terms of stress situations, socially anxious individuals did not show initial orienting bias and were more likely than non-anxious participants to focus on furious faces (having also longer fixation durations). Contrarily, non-anxious participants had a greater likelihood of fixation at cheerful faces within the two-second window after the commencement of the stimulus. Arousal and tension are also believed to have an impact on the distribution of gaze locations. Arousal studies have used gaze characteristics such gaze direction, gaze congruence, and the magnitude of the gaze-cuing effect. Furthermore, fixation instability has been linked to trait anxiety in both voluntary and stimulus-driven circumstances, although it is more evident under threat. Contrary to the neutral image, those who are more anxious often fixate on the emotional image first. [46]

Saccade

The distance traveled by the eyes during a saccadic eye movement is known as the saccade amplitude. Rapid, voluntary eye movements known as saccades are used to transfer the gaze from one place to another. The magnitude of the visual leap between focused spots may be calculated from the amplitude, which depicts the spatial extent of the saccade. The amount of time needed to complete a saccadic eye movement is referred to as saccade duration. It represents the rate of eye movement and can shed light on how well visual information is scanned and processed.

Additionally, saccadic duration, acceleration, and velocity have all been linked to arousal. In high arousal conditions, saccadic velocity has been thought of as an indicator of rising cognitive load and arousal. A precise time course for the rise in involuntary movements was seen under situations of arousal, along with a considerable increase in involuntary saccades. The antisaccade task, which disrupts saccadic control, can be used to identify specific inhibitory deficiencies associated with arousal. [46]

Pupil

The size of the pupils, which is related to emotional, cognitive, or sexual arousal, is a reflection of autonomic involuntary activity. The fluctuation in pupil size has been successfully used as a stress and arousal measure. When exposed to stressful stimuli in a lab setting, pupil width rises. The pupillary reaction seems to be stronger in people who report feeling more stress overall and when the visual stimuli are pictures with a negative valence. As opposed to emotionally neutral sounds, positive or negative arousing noises may also cause an increase in pupil size. Due to the high subject variability in the pupil diameter changes, irregular pupil size changes occurred under a variety of settings, and the peak point values varied depending on the task's level of difficulty. Mean pupil diameter has been demonstrated to positively correlate with the cognitive burden in a number of different activities, and pupil area has been shown to be substantially correlated with the user's continuing task difficulty. Along with an increase in cognitive load, the pupil size standard deviation also rises. The size of the pupil appears to expand in direct proportion to how much effort is required when doing intellectually taxing tasks. [46]

Blink

The amount of time that the eyes are closed as a result of a blink is referred to as blink duration. Periodic blinking might impede the processing of visual information. The length of a blink can be a good indicator of how often and how long your eyes blink.

A decrease in blink frequency indicates a tendency to focus on a particular area of a picture, which means that the viewer is making an effort to keep his or her eyes open for a longer period of time to observe the desired location. Blink rate can therefore provide useful information about the viewer's propensity to pay more attention to a particular location in a picture. During times of stress or other emotional arousal, spontaneous eye blinks occur more often and take on noticeably different patterns. This can partly be related to the blood being diverted to the periorbital eye muscles, which facilitate quick eye movements. However, when performing more demanding tasks (such reading a challenging text), eye blinks become less frequent. Eye blinking rates and stress levels are significantly correlated. The frequency of eye blinks temporarily increased in response to advertisements that induced artificial emotional reactions and automobile crash simulators that induced more genuine emotional responses. Given that there are noticeably more brief blinks in situations of high visual load, blink duration is a sensitive predictor of cognitive workload. Additionally, it was found that the blink time tended to shorten as the job got harder. The video game's slowest pace was also demonstrated to enhance blink length and frequency, and as effort grew, blink frequency dropped. From low to medium cognitive stress, the blink rate dropped; however, there was no additional change in high cognitive load. [46]

2.3 Applications

Eye tracking is a flexible technology that has uses in many different industries. Eye tracking is a very accurate method for recording and analyzing eye movements and gaze patterns, and it offers important insights into how people behave, think, and interact with their surroundings.

Medical and Clinical Settings

Eye tracking is important in many different aspects of medical and therapeutic research. It is widely used in ophthalmology to identify and track eye problems and disorders. Eye tracking aids in measuring eye movements, evaluating visual function, and identifying anomalies in gaze patterns. Early identification of eye conditions including glaucoma, macular degeneration, and diabetic retinopathy can be aided by it. Eye tracking is useful in both therapeutic situations and neurological research. It sheds light on the oculomotor system and aids in the diagnosis and follow-up of illnesses including multiple sclerosis, Parkinson's disease, traumatic brain injury, and stroke. Eye movement analysis can help determine how a disease is developing, how well a therapy is working, and how neurological disorders affect visual perception. Eye tracking helps investigate Autism Spectrum Disorder (ASD) and associated developmental problems. It assesses social attention and gaze behavior, revealing social interactions, communication patterns, and sensory processing variances. Early autism identification, diagnosis, and intervention can be helped by eye tracking. ADHD patients are studied using eye tracking. Eye movements during cognitive activities measure attention, distractibility, and impulsivity. Eye tracking helps diagnose ADHD and monitor therapy. These uses of eye tracking in clinical and medical research show how important it is for comprehending cognitive functions, behavior, and enhancing patient care. [47]

Psychology and Cognitive Science Research

Research in psychology and cognitive science frequently employs eye tracking to examine many facets of human perception, attention, and cognition. By recording and examining eye movements throughout various activities and stimuli, it offers useful insights into the underlying cognitive processes. Eye tracking is a tool used by researchers to examine visual attention, including how people direct their gaze in response to stimuli and how attentional mechanisms affect perception and decision-making. Using parameters like fixation lengths, saccade patterns, and regression movements during reading tasks, it is also used to study reading behavior. Researchers can better understand the mechanisms underlying visual search, memory functions, face perception, emotion detection, and other cognitive phenomena by using eye tracking in their investigations. Eye tracking technology advances psychology and cognitive science by giving accurate and objective measures of eye movements, which eventually results in a better understanding of human cognition and behavior. [27]

Human-computer Interaction and User Experience Design

Eye tracking makes it possible for more intuitive and natural user interfaces, which is essential for human-computer interaction (HCI). By employing their eye movements, gaze, and blink patterns, people may interact with and control computer systems. Individuals with motor limitations may find it easier to browse and engage with computer interfaces thanks to HCI systems that integrate eye tracking technologies. Additionally, eye tracking enables gaze-based interaction, in which users may scroll or move through material by merely gazing at things to choose them. HCI systems may modify and customise the interface to improve user experience by comprehending users' gaze behavior. By detecting users' gaze in the virtual world, eye tracking is used in virtual and augmented reality apps to enable realistic and immersive experiences. Additionally, it aids in user experience research and usability testing by offering insightful data on how people interact with software, websites, and apps. In HCI, eye tracking aids in the creation of user interfaces that are more effective and pleasant, thereby enhancing usability and user satisfaction. [27]

Market Research and Advertising

By giving insights into customer behavior and preferences, eye tracking has changed market research and advertising. Businesses may track customer attention and gaze patterns to have a better understanding of how people interact with marketing and products. Eye tracking aids in the analysis of design components, location optimization, and fine-tuning of marketing campaigns. Additionally, it offers useful information on how consumers make decisions, which helps to improve shop settings and boost sales. By detecting regions of interest and enhancing the user experience, eye tracking aids in the optimization of website designs and user interfaces in digital advertising and user experience. Eye tracking is an effective method for market research since it offers accurate and unbiased information on visual engagement and attention. It aids companies in honing their marketing techniques and producing ads that are more compelling and powerful. The most attention-grabbing components are identified via eye tracking, and ad placements are optimized. Eye tracking is used to evaluate the impact and efficacy of commercials. Additionally, it aids firms in comprehending how customers behave in retail settings, what things catch their attention, and how they make decisions. By refining product designs and producing more logical and user-friendly interfaces, eye tracking helps to increase consumer satisfaction and market success. [27]

3 Implementation

The practical application of this thesis is the primary topic of this chapter. The suggested solution is discussed in the first part of the chapter. This section provides a concise summary as well as the anticipated strategy, suggestions for the process, and intended results. The next step in the process is the actual implementation, which is then divided to its software and hardware components. It will be discussed how each stage progressed, the many approaches that have been implemented, the obstacles that have been faced during implementation, and how they are being gradually resolved.

3.1 Suggested Solution

The practical outcome of this master's thesis was the design of an eye tracking system. Before proceeding with the actual implementation, it was important to determine specific requirements that the system should meet. The focus of this work is on eye tracking, specifically the monitoring of eye parts such as the pupil and iris, as well as their movements during specific tasks.

The eye tracking system would make use of a camera that was positioned in front of the participant, with the camera's primary focus being on the region of the participant's face that included the eyes and the areas around the eyes. The subject would be able to sit in a comfortable posture without being physically attached or restricted in any way, such as by employing head mounts, thanks to the fact that the entire system would be based on an approach that is user-friendly. In addition to the camera, a crucial component is the infrared light source. This type of light source is frequently utilized in eye tracking owing to the fact that it is non-intrusive. This is because people are unable to sense infrared light. In the same way that the camera is placed in front of the subject, the infrared light will be placed in front of the subject in order to evenly illuminate the region. A computer and a monitor complete the remaining components of the technical setup. The computer will be used for a variety of purposes, including the manipulation and adjustment of the camera, the recording of videos, and the running of presentations that contain tasks for the patient to do while eye tracking is being performed. The computer monitor will be used to display these presentations.

After the setup has been properly prepared, configured, and validated, measurements will be taken on a number of different subjects. There is no specific limit or ideal number of participants, although as a general rule, more data leads to better findings. There is no limit or ideal number of subjects.

The next stage in the practical application of this thesis is the analysis of the dataset. This will allow the eye movement videos of the individuals to be prepared for the dataset. The Python programming language will be used to carry out the study, in addition to a wide variety of other libraries and methodologies that are now at the disposal. The generation of numerous individual outputs, such as the dynamic part, the static part, and others, will be the primary focus of the analysis of the dataset. These individual outputs will be further discussed in the sections that will follow. After then, the findings of the investigation are going to be detailed in the next chapter.

3.2 Hardware Implementation

The proposed solution that was described in the previous section was used as the basis for the implementation of the hardware part for practical section of the thesis. A camera, an infrared light source, a computer, and a monitor were the components that made up this system. In this subsection, a more in-depth description of the hardware component will be provided, including the selection process and the reasoning behind the individual components that were selected. In addition to that, it will describe the difficulties that had to be conquered in order to complete the preparation of the hardware component.

Camera

In the preparation of the hardware practical part, the first component that was addressed was the camera. The camera needed to fulfill a number of predetermined requirements. Because the purpose of the experiment was to measure eye movements, which can happen at a relatively high rate, the camera needed to be able to record images with high acquisition frame rate. A high quantum efficiency for the transmission of infrared light was yet another feature that was sought after. In order to accurately capture the rapid movements of the eye, it is essential to have a very fast scene detection process. Because of this, the camera will be able to effectively track the eye movements of the subject without experiencing a significant amount of lag or motion blur. For this particular task, a camera that possesses both a high frame rate and a quick response time would be ideal. Because the infrared illumination is applied so often, the quantum efficiency of infrared light is something that must be taken into consideration. If the quantum efficiency of the camera sensor is high, this indicates that it is able to detect and record the infrared light that is reflected by the eye more efficiently. In applications that involve infrared eye tracking, this results in improved sensitivity as well as image quality.

For the purposes of this work, the UI-3060CP-M-GL Rev.2 camera from IDS imaging company was chosen. It is a monochrome camera with a resolution of 1936 by 1216 pixels. At this maximum resolution, the camera achieves a frame rate of 166 frames per second. To achieve higher frame rates, it is necessary to lower the camera's resolution and to limit the region of interest. The quantum efficiency of this camera reaches a maximum of 80% at a wavelength of 500 nm. However, for the specific requirements of this work involving infrared light, the quantum efficiency is 20% at 800 nm, and the value decreases with increasing wavelength. The camera communicates with the computer via a USB cable using the uEye Cockpit software environment, where all the necessary parameters can be adjusted according to the requirements, including video recording and storage.

Fig. 3.1: Illustration of UI-3060CP-M-GL Rev.2 camera, taken from [48]

In addition to the camera, it was necessary to test and select a lens for the camera. Up to ten different lenses were tested, and the chosen lens is the Schneider TELE-XENAR 2.2/70-0902. This lens is suitable for measuring distances around one meter to one and a half meters, which is the desired distance between the subject and the camera. It also has adjustable focus and aperture (range from 70-90 mm) that can be fixed, and it can utilize aperture control if needed.

A filter that is attached to the camera and positioned in front of the lens of choice was the final component that was required to get the camera ready for measurement. The filter prevents light rays with wavelengths shorter than 800 nm from entering the camera; consequently, only light rays with longer wavelengths are permitted to enter. For the purposes of eye tracking, this ensures that the camera receives only light that falls within the desirable range of infrared light.

Source of Infrared Light

Infrared light is electromagnetic radiation with a longer wavelength than visible light. It is located on the electromagnetic spectrum between microwaves and visible light, which places it in the infrared region. Since the wavelength of infrared light falls outside the spectral range that can be perceived by the human eye, this type of light cannot be seen by humans. It finds application in a wide variety of fields, including eye tracking techniques, communication systems, therapeutic applications, detection of thermal radiation, and security systems. Near-infrared, mid-infrared, and far-infrared are the three different categories that can be used to classify infrared light according to the length of its wavelength. Near-infrared light, which has a wavelength range of approximately 700 nm to 2500 nm, is the only type of light that is utilized for the purposes of this work. [50]

There are many different justifications as to why eye tracking typically makes use of infrared light. Since infrared light is not within the visible spectrum, it cannot be seen by a human because our eyes are not sensitive enough to detect it. This makes it possible for eye tracking systems to illuminate the eyes without causing any discomfort to the user's vision or interfering with their ability to see normally. Infrared light is also considered safe for use in applications involving eye tracking, which is another reason. Even with extended contact, there is no evidence that it could cause any kind of damage to the eyes or pose any kind of risk. Infrared light, on the other hand, has the potential to improve contrast and accuracy when identifying eye movements and facial features. This precise measurement and analysis of gaze direction, pupil size, and other eye-related parameters is made possible by the eye tracking system, which captures the infrared light as it reflects off the cornea and is captured by the system.

The preparation of an infrared light source was necessary because of the factors that led to the decision to use infrared light in eye tracking, which were discussed previously. A borrowed light source was available from the school at the beginning of the work. However, due to the fact that it flickered rapidly on and off rather than producing a steady stream of light, it was insufficient to serve as a source of constant illumination. This issue was most likely brought on by the wiring of the infrared light source, which may have included a brightness sensor that detected the levels of light that were present in the surrounding environment. As a consequence of this, the light source was only able to function correctly at relatively dim levels of brightness, and it did not offer stable illumination. By covering the sensor with adhesive tape, the problem of flickering at low brightness levels was successfully resolved; However, a solution to the problem of flickering in this particular light source could not be found.

Because the issue with the first source of light could not be fixed, the focus of the work shifted to the development of a second source of light. The Kemo Electronic company provided the new potentially useful infrared radiation source that was obtained. To be more specific, the source is denoted by the label B223. On the other hand, the source was delivered as separate unassembled components that first required soldering together before they could be used. The supply requires between 12 and 14 volts of direct current to function and draws somewhere around 300 mA of current. This particular light source emits light with a wavelength that ranges from 870 to 950 nm, which happens to be the ideal range for the work being done for the thesis. In addition, once activated, the source of light maintains its level of illumination throughout the duration of its use, in addition, there is no sensor of light. [51]

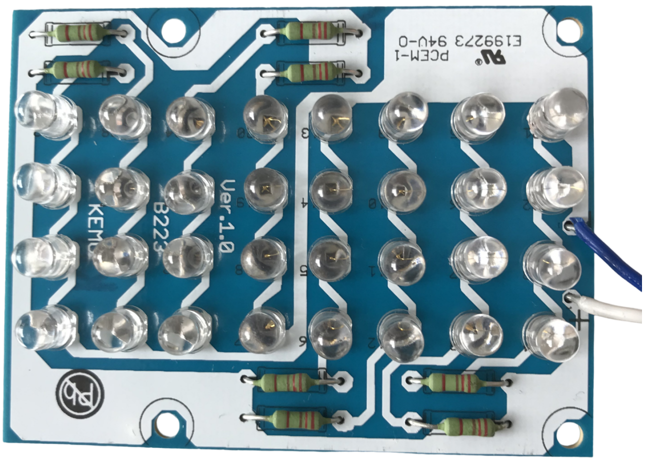

Fig. 3.2: Infrared Light Source B223

3.3 Dataset Acquisition

The preparation of the stimulus presentation, which outlines the sequence of stimuli that the subjects will be exposed to while the measurements are being done, was another significant step in the process. This step involves designing and organizing the stimuli in a structured manner to elicit specific responses or behaviors from the subjects.

Presentation of Stimuli

PowerPoint was the application that was used to create the presentation for the stimulus. The subject will be looking at a presentation while we conduct the measurement, and within that presentation are several crucial components. The initial step is known as the calibration phase, and it lasts precisely twenty seconds and is made up of ten frames. The fixation object, which is a black circle in this case, is moving across the screen in each individual frame. To be more specific, it begins in the middle of the screen, then it travels to the upper-left corner, then it travels to the upper-right corner, then it travels to the lower-right corner, then it travels to the lower-left corner, and finally it travels back to the middle of the screen.

Following this section, there will be a pause that lasts for five seconds, during which the subject will see a moving bar that indicates how much time is left in the pause. This pause has been strategically placed to facilitate relaxation, which will make it much simpler for the subject to refocus their attention on the aspects of the presentation that are still to come.

The dynamic phase comes up next in the presentation, and it's important to note that the black circle continues to serve as the fixation object throughout this phase. This section lasts for ten seconds, and it is broken up into five frames, each of which lasts for two seconds. On each frame, the fixation object appears at a different position on the screen.

After the dynamic phase concludes, there will be one more pause that is identical in length to the initial one. After this pause, we will now move on to the static portion of the presentation. During this stage, there will be a variety of tasks to complete. The first activity entails positioning a smaller, darker circle in the middle of the screen. The duration of this frame is three seconds. The circle is left intact after this, but a red grid in the shape of a cross is added to indicate where the center of the circle is located. The duration of this frame is also three seconds.

The second activity of the static phase is a frame displaying an image for a duration of four seconds. The famous Mona Lisa painting by Leonardo da Vinci is featured here in this presentation.

The final segment of the static phase, as well as the entirety of the stimulus presentation, consists of a single frame that is seven seconds long and displays two sentences, one of which is written in Czech and the other in Slovak.

Video Parameters Configuration

The next important part of creating the dataset was setting up the camera acquisition parameters. The uEye Cockpit application from IDS Imaging was used for video recording, along with the camera used in this work. The most crucial parameter to be adjusted was the frames per second (fps). A value of 200 fps was set to enable detailed analysis of eye movements. With a lower fps value, the motion would not be as smooth, and the analysis would not be feasible. To achieve 200 fps, it was necessary to reduce the resolution of the observed scene. However, an issue occurred when saving videos at a specific resolution and 200 fps. The problem was that the saved video had the fps value halved from the preset 200. This issue was resolved by setting the horizontal and vertical bidding to twice the desired value. The pixel clock was set to the maximum possible, and the exposure time was set to an automatic value, which was then fixed as the value 6.3. The gain value was set to 60%. All these values were fixed, and the option for automatic adjustment during video recording was disabled.

Synchronization of video recording and stimulus presentation

The synchronization of the video recording and the stimulus presentation was the last task that needed to be completed before the actual measurements. An AutoHotkey script was developed and used in order to achieve the desired level of synchronization. The script started the video recording by using the uEye Cockpit application, waited for two seconds, and then began presenting the stimuli to the subject. The necessary synchronization between the video recording and the stimulus presentation was achieved through the use of this process. The desired level of synchronization was attained through the utilization of the AutoHotkey script, which made it possible to achieve precise alignment of the stimulus presentation with the sections of the recorded video that corresponded to it. This synchronization was necessary for subsequent data analysis because it made it easier to precisely analyze eye movements in relation to the stimuli that were being presented. Because of this, the video was able to be broken up into its component parts so that each could be analyzed separately.

Content of Dataset

The process of creating a measurement and dataset was broken down into three distinct stages. During the first stage, it was determined that there was a problem with the infrared light source, so a new one had to be installed in its place. The second measurement uncovered an issue with the fps, which was also fixed after being discovered. As part of the third measurement, videos that will be used to construct the dataset for this study were gathered. The dataset contains videos contributed by 10 different subjects, with multiple videos contributed by each subject. Each video is consistent with the parameters that were specified. The dataset is going to be analyzed in greater depth in the subsequent chapter.

3.4 Software Implementation

The software part of the practical work will focus on loading and processing the dataset, which consists of recorded videos from individual subjects. The software component can be divided into several smaller parts that depend on each other and will be described sequentially in this section of the chapter. To be more specific, the software incorporates processes such as initial loading, eye detection, pupil detection and tracking, data processing, and at last data analysis. The software solution was developed using the Python programming language, specifically version 3.9. The Spyder application from Anaconda was used for actual programming of the solution. The results obtained from this subsection will be described in the following chapter, "Results and Discussion".

Initial Data preprocessing

At the beginning of the program, it is important to gather basic parameters of the data and videos. Parameters such as the width, height, frame count, duration, and frames per second of the video are obtained and stored in variables for further use in the program. Additionally, a variable is created to indicate the subject's name, allowing all achieved results to be saved with this variable for identification purposes.

3.4.1 Eye Detection

The first primary task was to determine the positions of both eyes and their surroundings in the data. This was achieved by extracting two smaller regions of interest from a single video frame, one for the left eye and another for the right eye. Haar cascade classifier function with pre-trained Haar features, freely available from OpenCV, was used for this purpose. The eye regions are detected initially to simplify the subsequent detection of both pupils within these smaller regions of interest. Typically, the regions of interest are detected from the first video frame, but if only one eye and its region are detected, the process is repeated on the next frame until both regions of interest are found. The original frame and the two new regions of interest are shown in Figure 3.3.

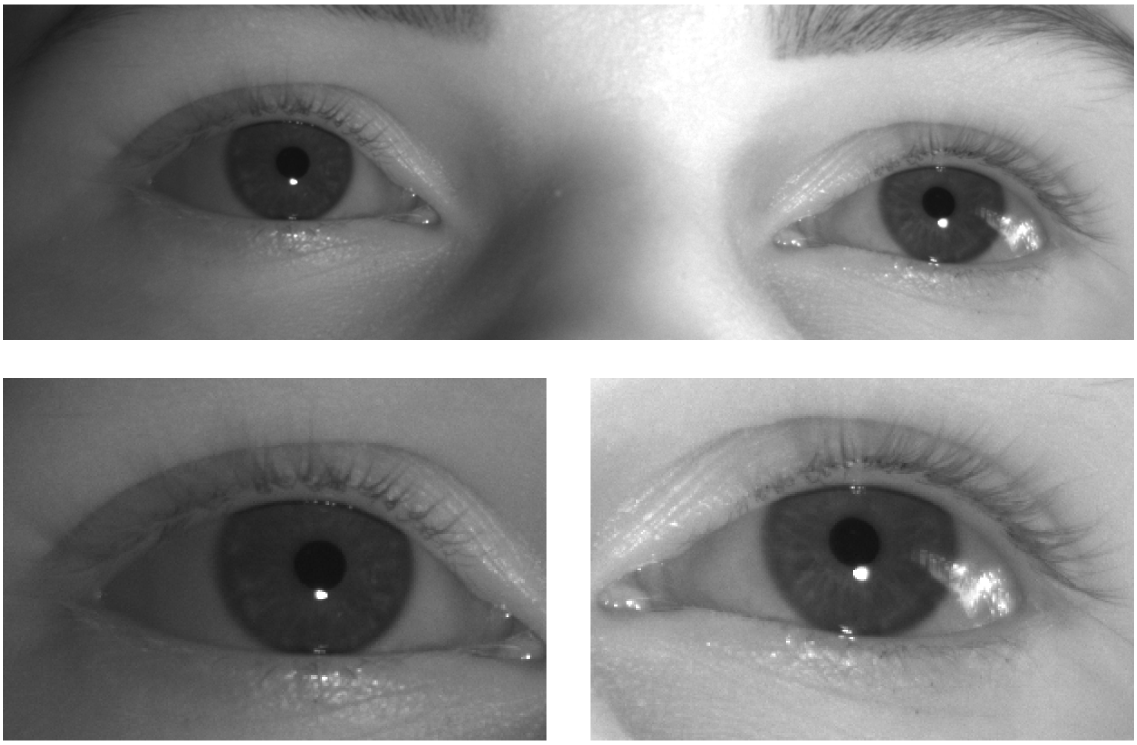

Fig. 3.3: Top image: Original frame of video (slightly cropped for a privacy purpose), Left image: New ROI of Detected Left Eye, Right image: New ROI of Detected Right Eye (both Left and Right images are sized up for better presentation)

3.4.2 Pupil Detection and Tracking

After achieving success in detecting the eyes and generating new regions of interest, the next objective was to successfully detect and track the movement of both pupils, independently for each eye. The method of double filtering was utilized to get things started in the detection process. The Gaussian filter was the first filter that was applied; it is a low-pass filter, so it removed high-frequency components from the signal before proceeding. Size of used kernel was (11,11) and both sigmas were in default value. After that, a median filter with a mask sized 3 was applied so that noise could be removed while edge information was kept intact. Following the application of the double filtering, the minimum value of contrast was found and it was saved to a variable. Image thresholding was performed after this using the variable of minimum contrast plus a constant equaled to 10 to obtain the pixels in the image that were the darkest. These pixels represent the area of the pupil that was being viewed. Subsequently, the pupils were detected using contours, as it is a segmentation task. The centroid of the contour was calculated to determine the center point, where a red circle could be drawn for better visualization. This process was repeated for each individual frame, and the x and y positions of the contour center that represented the pupil's position were saved in two variables. This process was repeated for the right eye as well, which resulted in four variables for the x and y positions of both pupils for every frame in the video. The data analysis process makes further use of these variables.

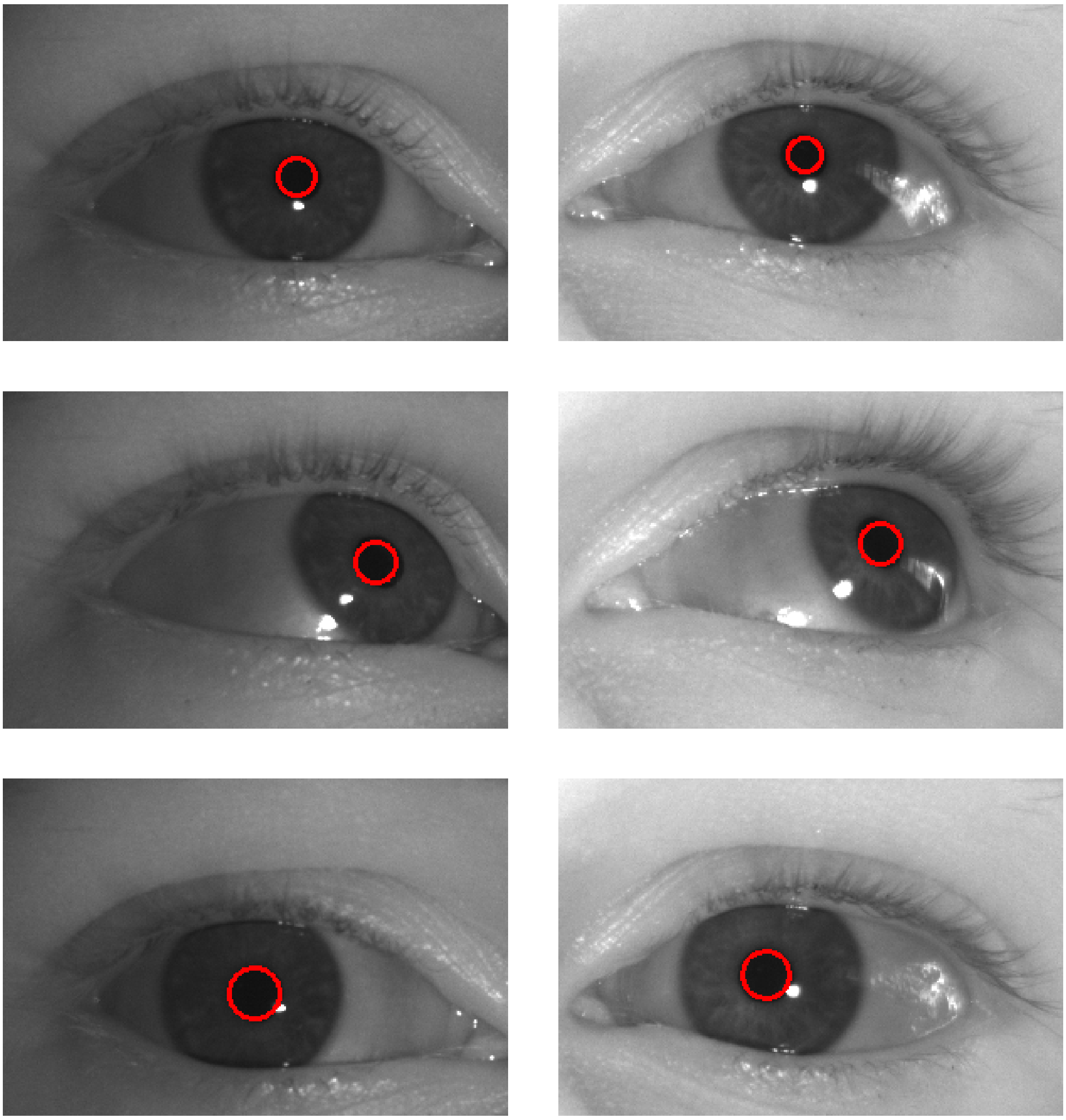

Fig. 3.4: Detected Pupils for both eyes. Red circle indicating detected pupils. Top row: Subject is looking in the center, Middle row: Subject is looking in the Top Left Corner, Bottom row: Subject is looking in the Bottom Right Corner.

Pupil Size Estimation

At the same time as attempting to detect the pupil, a binary image created by thresholding was utilized to calculate the remaining pixels representing the darkest areas of the image, indicating the pupil. This process was performed for both eyes, resulting in the pupil size in each frame for both eyes throughout the entire video. The obtained results will be presented and discussed in the following chapter.

3.4.3 Video Segmentation and Pixel to Metric Conversion

Video segmentation and the conversion of pixel values to metric values are two distinct goals that are required for further data analysis.

Video Segmentation

It is essential to divide the video into temporal segments according to the individual parts and tasks that are presented in the stimulus presentation. This is done through video segmentation. A correct segmentation is essential in order to guarantee that the analysis is carried out independently on each segment and that the data from one segment does not overlap into data from another segment. The use of a JSON file that contains the temporal intervals for segmentation was necessary in order to accomplish this goal. After importing the JSON file into the code, the data is then extracted and partitioned into these temporal segments so that further data analysis can be performed. These temporal intervals are used to divide the previously mentioned four variables that are the result of pupil detection and tracking. This provides information about the position of both pupils within each individual segment. The useful temporal segments include system calibration, the dynamic part, and three static parts: circle analysis, image analysis, and text analysis. The remaining parts, which include pauses, the beginning of the presentation, and the end of the presentation, are not relevant and are not subjected to analysis.

Pixel to Metric Conversion

The conversion of pixel values to real-world metrics, specifically centimeters in this work, is an important task as it aids in various other tasks. One immediate consequence of this conversion is the ability to determine the subject's gaze on the screen, i.e., precisely where they are looking. This enables the analysis of the image, where the information on gaze position is crucial, but that will be discussed in a subsequent subsection. Another utilization is in the visualization of resulting graphs, where the use of metrics helps readers envision the actual gaze location, movement patterns, and more.

This task directly relates to the segmentation of the video into time intervals, as it requires information about the calibration phase from the stimulus presentation. The calibration phase itself needs to be further divided into smaller intervals, with 10 intervals in total, as each indicates a different position on the monitor. Each position within the calibration interval is presented twice to achieve more accurate calibration. The positions include the center of the screen, the top-left corner, the top-right corner, the bottom-right corner, and the bottom-left corner. From each interval, the median values were computed to determine the central value. Additionally, the medians of two intervals representing the same position were also calculated to obtain a final median value. These resulted in five median values representing the median values for five different screen positions. This process was performed separately for each eye.

To achieve the conversion factor, additional parameters were needed, namely the real dimensions of the monitor and the real values of the five positions. The monitor dimensions were 62 cm in width and 29 cm in height. The real position values were as follows: center of the monitor [0, 0], top-left corner [-31, 14.5], top-right corner [31, 14.5], bottom-right corner [31, -14.5], and bottom-left corner [-31, -14.5]. It was also necessary to calculate the monitor dimensions in pixels, which was done by subtracting the top-left corner from the top-right corner for the width and subtracting the top-left corner from the bottom-left corner for the height. This provided the width and height of the monitor in pixels. The conversion factor was then calculated as the real width of the monitor divided by the width of the monitor in pixels. The same calculation was performed for the height, resulting in two conversion factors, one for each eye. These four conversion factors were obtained. The final step involved the conversion of pixel variables (x and y) for the eyes into centimeters. This was done by subtracting the pixel value of x from the central pixel value and multiplying it by the conversion factor for width. Similarly, the calculation was performed for y, resulting in the conversion of all x and y values for both eyes into centimeters.

It should be noted that the conversion was not 100% accurate, and there were small deviations observed during testing. Errors in the conversion could be attributed to several factors, but the main factor is head movement. Although the head movements were small for most subjects, they still caused deviations. However, since the system aims to be user-friendly and avoids requiring subjects to keep their heads perfectly still, small deviations are taken into account.

The last but important part of this subsection should mention another conversion, which is the conversion of frame count to time. This was simply achieved by dividing the number of frames by the frames per second (FPS) value.

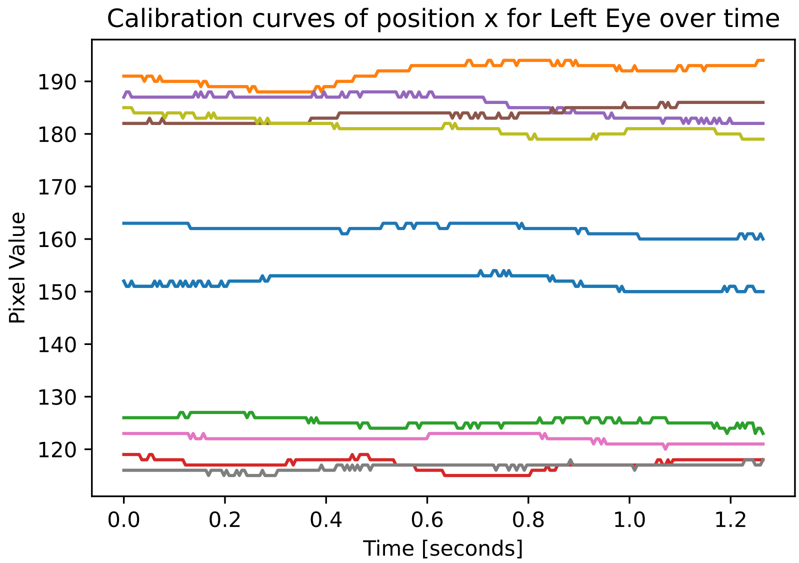

Fig. 3.5: Calibration curves of x for Left Eye. Top group of 4 curves are indicating the biggest x value, both top and bottom left corners are included there, Middle Group is consists of two centers, Bottom Group includes top and bottom right corners with smallest x value.

3.4.4 Video Analysis

In this final subsection of the Software Implementation section, the procedure of the individual data analyses is described, while the results and discussions of the analysis will be presented in the next chapter. The analysis of the video can be divided into two parts. The first part is the analysis of the dynamic segment, and the second part is the analysis of the static segment, which is further divided into various subsections: static part circle, static part image and lastly static part text.

Dynamic Part of Video Analysis

This subsection is referred to as dynamic due to the fact that the tracked object of interest that is being presented in the stimulus presentation moves around, and the task that the subject is tasked with performing is to follow it. The object in question is essentially a black circle that moves across the screen in the following five positions: from the center to the left, then to the right, followed by the upper-middle, and finally the lower-middle position. This section starts off by separating the dynamic segment from the rest of the video, and the rest of this section consists of an analysis that is solely focused on the dynamic segment. Estimating certain parameters from the signal is the objective of this section. These parameters include the distance traveled between two frames, the speed of the movement, and, finally, the acceleration of the motion.

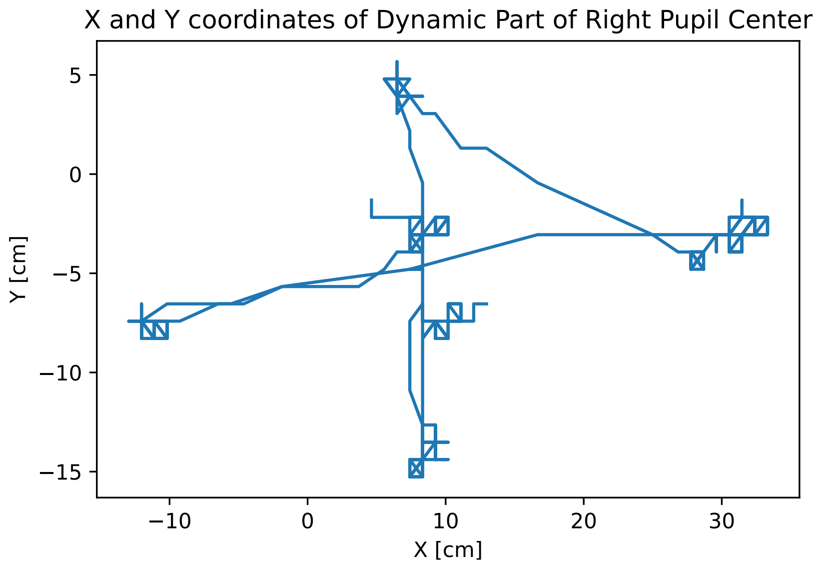

Fig. 3.6: Curve representing Right Eye Movement in the dynamic Part

Static Part Circle of Video Analysis

The first part of the static practical analysis focuses on one parameter, specifically the variance of positional data. An experiment was designed and conducted, involving two objects of interest. Both objects are black circles positioned at the center of the screen, with the only difference being that the second circle contains an additional grid in the shape of a cross, indicating the center of the circle. This part starts by selecting the segmented video footage that includes only this specific part of interest. The variance is then calculated in both axes, and a comparison is made between the segments where the circle is presented without the grid and where the circle is presented with the grid.

Static Part Image of Video Analysis

In this section, the analysis of the static part of the image was performed. The image used was Mona Lisa. The duration of this interval was 4 seconds, during which the subject was instructed to look at the image according to their preferences. The analysis was conducted only on the interval from the segmented video corresponding to this specific part of the stimulus presentation. The result of this section was intended to be a heatmap showing the positions of pixels on the selected image.

Each position was assigned a color based on the frequency of occurrence, indicating the number of pixels at that location. In this work, the process was performed as follows: firstly, a matrix was created, and the mentioned image was placed in its center with an opacity of 0.7 to preserve the contrast of the pixels. Then, the occurrences of paired positions for individual pixels were counted only within the interval for the static part of the image data. After determining the frequencies, the pixels had to be converted from the eye image to a real monitor image and then further converted to match the positions in the created matrix. Positions where multiple pixels coincided with their positions were marked with a red color, while positions with few or almost no pixels had decreasing intensity, eventually becoming completely transparent. This approach achieved a heatmap, providing insight into the exact gaze direction of the subject.

Static Part Text of Video Analysis

Text analysis was the focus of the very last part of the static practical section, which meant that it was also the very last section overall. During the presentation of the stimulus, there was a single slide that was displayed for seven seconds. This slide contained two sentences that were each split across two lines. The participants were tasked with reading the passage and demonstrating that they understood it. An interval containing data specifically from this task was extracted using segmentation as the means of analysis. The data underwent additional manual analysis, and the findings of that work will be presented in the chapter that comes after this one.

4 Results, Findings and Discussion

This chapter focuses on the results achieved during the data analysis. It will include a discussion of the significance of the individual results. Additionally, it will describe any encountered problems and their solutions. The entire implemented system will be evaluated, including its strengths and weaknesses. Possible modifications and directions for further system development will also be discussed.

Eye Detection

To begin with, it is important to mention that the chosen method and procedure for eye and periocular region detection have been widely used in the field of eye tracking for almost two decades. More about this method and the way it was used back in section 3.4.1 Eye Detection. However, even though more recent methods are available, there was no need to adopt them in this work. The chosen method proved to be reliable and accurate, performing well on both good-quality data and challenging data conditions such as poor lighting, partial occlusion, and other factors.

Pupil Detection and Tracking

The pupil detection functioned well, successfully detecting the pupil in each frame where it was present. However, there were challenges related to artifacts caused by blinking. These artifacts occurred during the interval when the pupil was not visible to the detector due to eyelid occlusion. As a result, the detector would sometimes detect the pupil at random positions within that interval.

To address this issue and remove blinking artifacts, a solution was implemented during the data collection phase. Multiple pauses were included in the stimulus presentation, and participants were instructed to blink only during these pauses. This approach proved to be successful in mitigating the blinking artifacts and improving the overall quality of the data, because the intervals of individual pauses were not used for data processing and analysis.

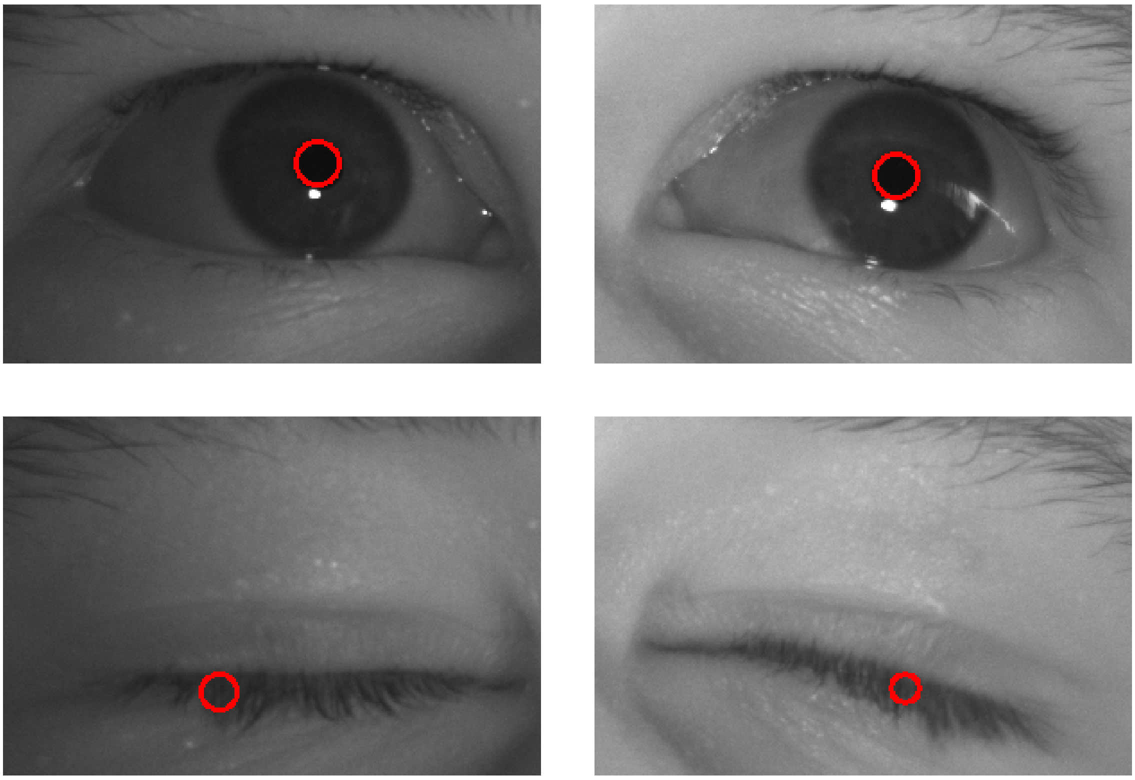

Fig. 4.1: Top image: Original frame of video (slightly cropped for a privacy purpose), Left image: New ROI of Detected Left Eye, Right image: New ROI of Detected Right Eye (both Left and Right images are sized up for better presentation)

Pupil Size Estimation

Although the analysis or estimation of pupil size was performed on the entire dataset, only the calibration interval from the segmented video was selected for its analysis and discussion. The consideration of the pupillary light reflex suggests that the pupil size should be approximately the same across different gaze directions. However, this could not be confirmed in this work due to the camera placement above and to the right of the subjects. When the subject looked towards the lower right corner or the center of the monitor, the pupil appeared larger compared to when the subject looked towards the other three corners. In these other corners, the pupil was not captured as a circular shape but rather as a flattened shape. Therefore, the average pupil size was analyzed only from the calibration portion where the subject looked at the center of the screen.

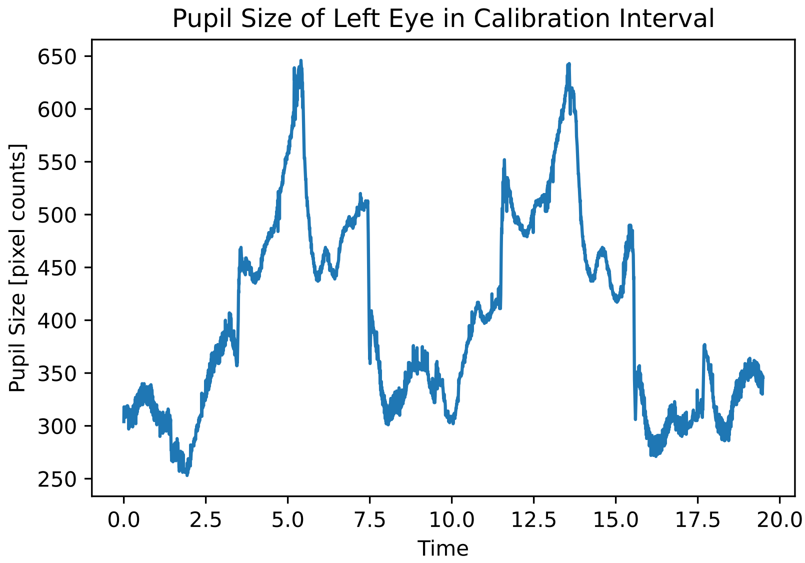

Fig. 4.2: Pupil Size for Left Eye in Calibration interval.

A table of averages was created from the first 2 seconds of calibration, where the subject looked at the center of the screen. From these averages, the overall average pupil size was obtained for the 10 subjects. The average of this work data for pupil size is 513.782.

Tab. 4.1: Table presenting mean values for each subject and then overall mean.

	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	ALL
Pupil Size Mean	342.93	302.52	665.83	581.56	526.8	727.06	632.78	608.63	399.44	350.27	513.782

Dynamic Part

In the dynamic part of the experiment, an object was observed that changed its position. The main goal of this part was to identify the largest change in the position of the object of interest, specifically from the left side of the monitor to the right side. In this small interval, the distance traveled was tracked, and the instantaneous velocity was calculated based on that. The obtained results are shown in Figure 4.3, where moments are depicted when the subjects started moving their gaze towards the right side of the screen to focus on the object of interest. This allows for several comparisons.

Firstly, the reaction time can be observed. It can be seen that Subject 7 had the shortest reaction time and thus initiated eye movement first, while Subject 9 had the longest reaction time and started the movement last. Another aspect that can be observed is the speed of the movement performed to change the gaze towards the object, indicated by the size of the peaks for each subject. The graph shows that Subject 3 had the fastest movement, while Subject 6 had the slowest. The thickness of the peaks also corresponds to the speed of the movement, with higher peaks indicating thinner movements and vice versa.

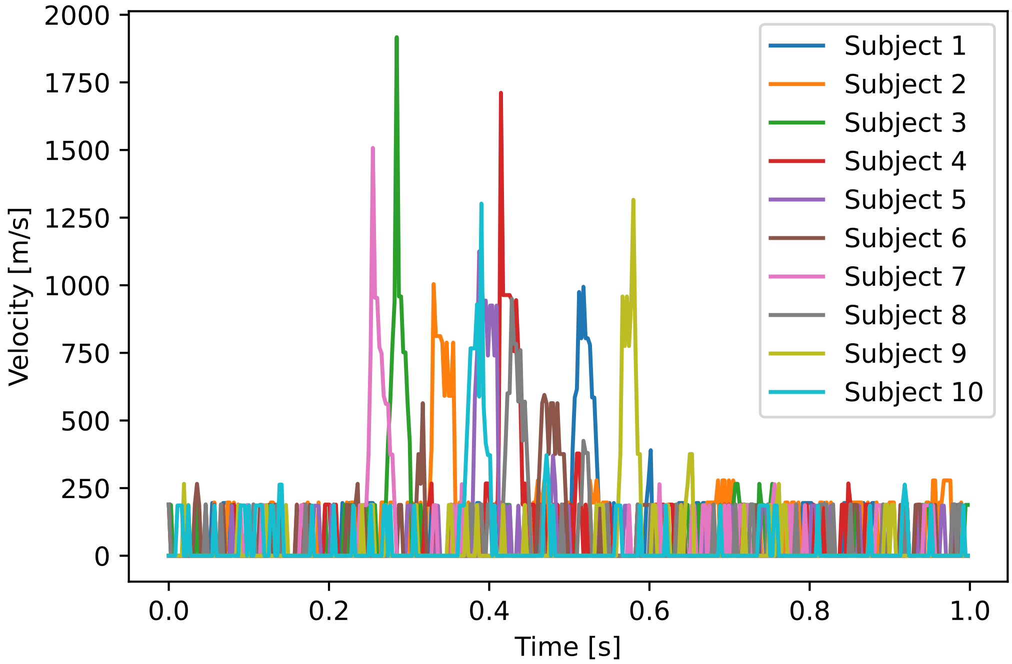

Fig. 4.3: Velocity in 1s interval for 10 subjects

Static Part Circle

In this analysis of the static part, the object of interest was a black circle for 3 seconds, followed by a black circle with a white cross indicating the center of the circle for 3 seconds. The task was to focus the gaze on the center of the circle. The achieved results are twofold: the variance in the x-direction and the variance in the y-direction for 5 subjects separately. From Table 4.2, it can be observed that for almost every subject, both variances are smaller for the grid circle. This means that the grid helps to maintain the gaze in the center of the circle and thus restricts small eye movements.

Tab. 4.2: Variance values in both axis X and Y for 5 subjects with and without grid

	S1		S2		S3		S4		S5	
	None	Grid	None	Grid	None	Grid	None	Grid	None	Grid
Variance X	1,24	0,81	1,86	2,11	1,32	0,4	3,72	1,81	5,15	0,42
Variance Y	2,22	0,31	1,57	0,66	2,36	0,7	3,1	1,1	2,21	0,61

Static Part Image

The main goal of this static part was the analysis of the image, specifically an attempt to obtain a heatmap that would indicate where and for how long a person looked at the given image. Out of the dataset generated from 10 subjects, only one correct heatmap was obtained. The remaining heatmaps did not align perfectly with the image, which could have been caused by inaccuracies in the conversion process, possibly due to head movements during calibration.

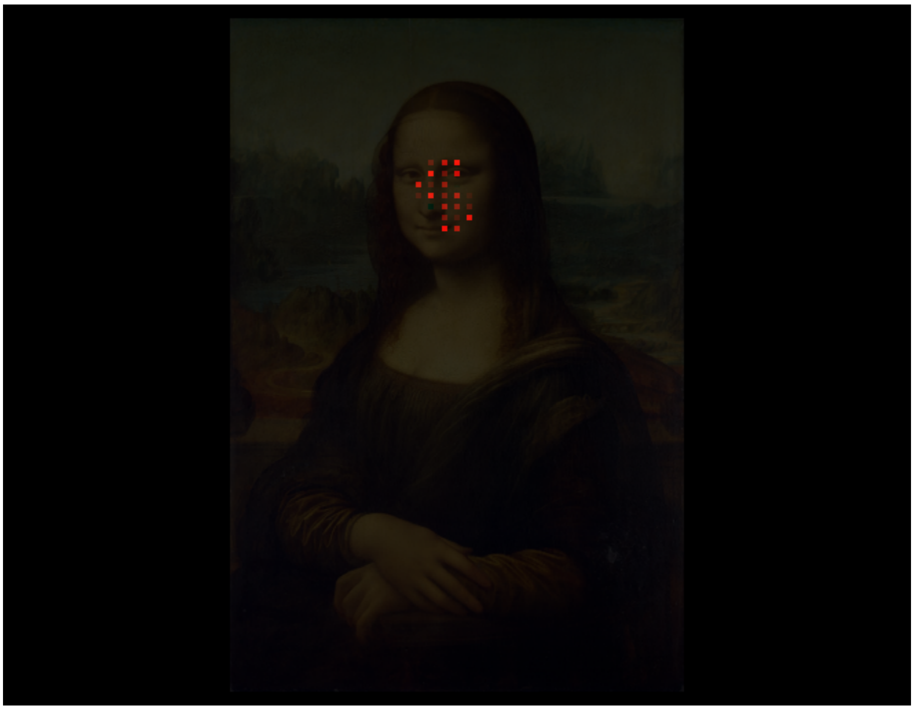

Fig. 4.4: Heatmap representing frequency and locations of fixations

Static Part Text

The static part analyzing text was primarily intended for various analyses, such as reading speed between genders. However, in the stimulus presentation, there are two sentences, one in Czech and the other in Slovak. Therefore, an attempt could be made to compare the number of fixations during reading and compare subjects from the Czech Republic with subjects from Slovakia. However, conducting such an analysis would require a large dataset with a significant number of subjects, as drawing statistical conclusions from only 10 individuals in the dataset of this study would not be reliable. Nevertheless, there was an attempt to determine the average number of fixations required to read both sentences individually and possibly the time taken. This led to the finding that the first sentence in Czech was read, on average, 0.5 to 1 second faster by all subjects, regardless of gender or nationality, compared to the sentence in Slovak. However, this finding could be biased due to manual measurements, and it is also possible that the longer length of the Slovak sentence (four additional characters) contributed to the difference. On average, it took 8 fixations, which translates to 8 mini pauses, to read one sentence. On average, it took subjects around 5 seconds to read both sentences. After the 5-second mark, the signal starts to lose its regular pattern of irregular fluctuations with occasional sharp drops followed by increases. The signal becomes highly irregular beyond this point.

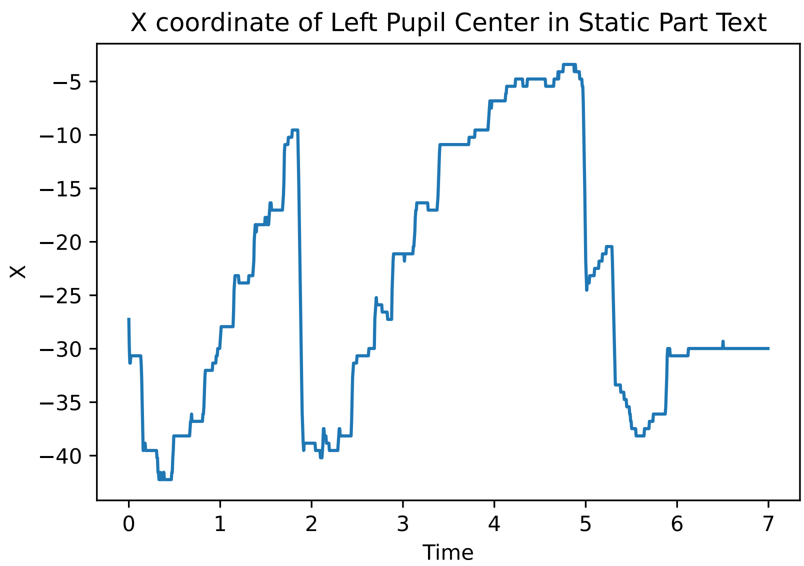

Fig. 4.5: Example of Signal during Static Part Text interval

Designed and Implemented System Evaluation

The implementation and subsequent deployment of the system were successful. A hardware setup was established that is capable of collecting good-quality data. There are no major issues with the lighting, camera, or other components. However, it is important to note that the entire system is designed to be user-friendly. This means that individuals interested in collecting and analyzing their own data do not need to spend a significant amount of time setting up the system but can comfortably measure their data in a normal sitting position.

On the other hand, this user-friendly approach introduces some level of imprecision in the data. Despite instructing the subjects to minimize movement during measurements, some minor movements may still occur. However, these slight movements do not render the data unusable. They may slightly complicate certain analyses, such as data calibration, but overall, the data remains reliable and valuable.

It is worth mentioning that the user-friendly nature of the system allows for greater accessibility and ease of use, which outweighs the minor imprecisions introduced by subject movement. The system still provides valuable insights and analysis capabilities for researchers and individuals interested in eye-tracking data.

Regarding the software implementation, the pupil detection and tracking performed well, with the exception of the artifacts caused by blinking, as mentioned earlier. The blinking issue was addressed by introducing multiple short pauses in the video. The pixel-to-real-world metric conversion, specifically to centimeters, was mostly accurate, although some deviations were observed for certain values and subjects.

The individual data analyses yielded satisfactory results. Both the dynamic analysis and static analysis, which focused on analyzing the circle and circle with a cross, produced good outcomes. One potential modification could be to develop a more advanced text analysis. However, it would require careful consideration whether to focus on detecting individual eye movements or conducting statistical experiments among different groups of subjects.

Expanding the dataset size would be a valuable improvement. Having more data would allow for the aforementioned statistical experiments and provide a more robust foundation for further analysis.

Conclusion

In conclusion, the main objectives of this thesis, which were the design and implementation of an eye tracking system, have been successfully achieved. The initial part of the work provided a theoretical introduction to ensure a comprehensive understanding of the field of eye tracking. It started with the anatomy and physiology of the eye, progressing to the various eye movements and even addressing pathological conditions. Once the fundamentals of the eye were covered, the focus shifted to the concept of eye tracking itself. This included defining what eye tracking entails, exploring its applications across different domains, discussing the available metrics, and highlighting its potential uses. With the theoretical foundation established, the thesis proceeded to present the proposed solution, which served as the basis for the implementation phase.

The implementation process began with the hardware section, where the eye tracking setup was constructed. This setup consisted of components such as camera, infrared light, a computer, and a monitor. Each component was carefully tested and validated. After conducting test measurements on a small group of participants, adjustments were made, and a second round of measurements was performed. During this data collection phase, an issue with the infrared light was identified and subsequently replaced. Following the acquisition of the dataset involving 10 subjects, the implementation of the software components commenced. This involved applying various methods for data processing and analysis. The first step was the detection of the eyes and pupils, from which the center coordinates of each pupil for every frame were extracted. This data formed the core of the thesis and was further segmented based on the presentation of stimuli. The analysis was divided into two parts: dynamic and static. In the dynamic analysis, the primary focus was on examining the shortest and longest reaction time according to stimuli, as well as the fastest and slowest eye movements across the screen. The results were visualized in a graph depicting the velocity over time.

In the static analysis, the effectiveness of a grid-shaped stimulus in the center of a circle was evaluated. It was provided in order to maintain fixation without minor eye movements . The results indicated a positive effect. The remaining static analysis, which involved text and image focus, yielded mixed results. The text analysis aimed to determine the necessary time and number of fixations to read a line of text, while the image analysis generated a heatmap. Only one heatmap was successful, while the others exhibited significant inaccuracies. The discrepancies were attributed to errors in the conversion of pixel coordinates to real-world metric. This error may have originated during the data collection phase, particularly when participants moved their heads excessively, compromising the calibration process and the accurate determination of the conversion factor. Despite these limitations, the implemented eye tracking system fulfills its purpose. Its user-friendly interface provides comfort and ease of use, although it may sacrifice some precision in the collected data. Potential modifications could include focusing on a specific type of analysis from the available options and delving deeper into its exploration. Additionally, expanding the dataset and conducting statistical comparison could provide further insights and opportunities for improvement. In summary, this thesis has successfully achieved its objectives by designing and implementing a comprehensive eye tracking system. It has provided valuable insights into eye movements and their associated cognitive processes. While the system demonstrates user-friendliness, it also exhibits limitations in terms of data accuracy. Future work may involve refining specific analysis components or expanding the dataset for broader statistical comparison and further advancements.

Bibliography

[1] *SYNEK, Svatopluk a Šárka SKORKOVSKÁ, 2014. Fyziologie oka a vidění. 2., dopl. a přeprac. vyd. Praha: Grada. ISBN 978-80-247-3992-2.*

[2] *DAVSON, Hugh a Edward S. PERKINS. Human eye. Britannica [online]. [cit. 2023-05-05]. Dostupné z: https://www.britannica.com/science/human-eye*

[3] *Blunt Eye Injuries, 2022. In: MDS manual [online]. [cit. 2023-05-05]. Dostupné z: https://www.msdmanuals.com/en-sg/home/quick-facts-injuries-and-poisoning/injuries-to-the-eye/corneal-abrasions-and-corneal-foreign-bodies*

[4] *HEJTMANCIK, J. Fielding a Alan SHIELS, 2015. Overview of the Lens. In: Molecular Biology of Eye Disease [online]. Elsevier, 2015, s. 119-127 [cit. 2023-05-07]. Progress in Molecular Biology and Translational Science. ISBN 9780128010594. Dostupné z: doi:10.1016/bs.pmbts.2015.04.006*

[5] *SUNDERLAND, Daniel K. a Amit SAPRA. Physiology, Aqueous Humor Circulation. [online]. 2023 Mar 12 [cit. 2023-05-07]. Dostupné z: https://www.ncbi.nlm.nih.gov/books/NBK553209/*

[6] *BISHOP, Paul N., 2010. Encyclopedia of the Eye. In: Encyclopedia of the Eye. Academic Press, s. 37-43. ISBN 9780123742032.*

[7] *SPENCER, Robert F. a John D. PORTER, 2006. Biological organization of the extraocular muscles. In: Neuroanatomy of the Oculomotor System [online]. Elsevier, 2006, s. 43-80 [cit. 2023-05-09]. Progress in Brain Research. ISBN 9780444516961. Dostupné z: doi:10.1016/S0079-6123(05)51002-1*

[8] *CFCF. The Extrinsic Eye Muscles right eye. In: Wikipedia Commons [online]. [cit. 2023-05-08]. Dostupné z: https://commons.wikimedia.org/wiki/File:1107_The_Extrinsic_Eye_Muscles_right_eye_lat.png*

[9] *SLINEY, David H., 2002. How Light Reaches the Eye and Its Components. International Journal of Toxicology [online]. 21(6), 501-509 [cit. 2023-05-10]. ISSN 1091-5818. Dostupné z: doi:10.1080/10915810290169927*

[10] *MOLDAY, Robert S. a Orson L. MORITZ, 2015. Photoreceptors at a glance. Journal of Cell Science [online]. 128(22), 4039-4045 [cit. 2023-05-11]. ISSN 1477-9137. Dostupné z: doi:10.1242/jcs.175687*

[11] *SKALICKY, Simon E., 2016. Ocular and Visual Physiology [online]. Singapore: Springer Singapore [cit. 2023-05-11]. ISBN 978-981-287-845-8. Dostupné z: doi:10.1007/978-981-287-846-5*

[12] *HUFF, T, N MAHABADI a P TADI. Neuroanatomy, Visual Cortex [online]. [cit. 2023-05-11]. Dostupné z: https://www.ncbi.nlm.nih.gov/books/NBK482504/*

[13] *WESTIN, Jack. Contralateral control of vision. In: Jack Westin [online]. [cit. 2023-05-11]. Dostupné z: https://jackwestin.com/resources/mcat-content/vision/visual-processing*

[14] *GEGENFURTNER, Karl R. a Daniel C. KIPER, 2003. COLOR VISION. Annual Review of Neuroscience [online]. 26(1), 181-206 [cit. 2023-05-11]. ISSN 0147-006X. Dostupné z: doi:10.1146/annurev.neuro.26.041002.131116*

[15] *ZHANG, Yan, 2015. RPE and Choroid Mechanisms Underlying Ocular Growth and Myopia. In: Progress in Molecular Biology and Translational Science. Academic Press, s. 221-240. ISBN 9780128010594.*

[16] *CHAU, Felix Y. a Dana WALLACE, 2014. Osteogenesis Imperfecta and the Eye. In: Osteogenesis Imperfecta. Academic Press, s. 289-303. ISBN 978-0-12-397165-4.*

[17] *Refractive Errors. In: Stereo Optical [online]. [cit. 2023-05-08]. Dostupné z: https://www.stereooptical.com/childrens-vision/refractive-errors/*

[18] *LIU, Yu-Chi, Mark WILKINS, Terry KIM, Boris MALYUGIN a Jodhbir S MEHTA, 2017. Cataracts. The Lancet [online]. 390(10094), 600-612 [cit. 2023-05-08]. ISSN 01406736. Dostupné z: doi:10.1016/S0140-6736(17)30544-5*

[19] *INGIMAGE. Symptom PNG. In: Pngtree [online]. [cit. 2023-05-11]. Dostupné z: https://pngtree.com/freepng/healthy-and-cataract-eye-closeup-view_8670704.html?sol=downref&id=bef*

[20] *KING, A., A. AZUARA-BLANCO a A. TUULONEN, 2013. Glaucoma. BMJ [online]. 346(jun11 1), f3518-f3518 [cit. 2023-05-08]. ISSN 1756-1833. Dostupné z: doi:10.1136/bmj.f3518*

[21] *Casson, R.J., Chidlow, G., Wood, J.P., Crowston, J.G. and Goldberg, I. (2012), Definition of glaucoma: clinical and experimental concepts. Clinical & Experimental Ophthalmology, 40: 341-349. https://doi.org/10.1111/j.1442-9071.2012.02773.x*

[22] *LIM, Laurence S, Paul MITCHELL, Johanna M SEDDON, Frank G HOLZ a Tien Y WONG, 2012. Age-related macular degeneration. The Lancet [online]. 379(9827), 1728-1738 [cit. 2023-05-08]. ISSN 01406736. Dostupné z: doi:10.1016/S0140-6736(12)60282-7*

[23] *MOHAMED, Quresh, Mark C. GILLIES a Tien Y. WONG, 2007. Management of Diabetic Retinopathy. JAMA [online]. 298(8) [cit. 2023-05-08]. ISSN 0098-7484. Dostupné z: doi:10.1001/jama.298.8.902*

[24] *STEEL, David. Retinal detachment. BMJ clinical evidence, 2014, 2014.*

[25] *Retinal Detachment. In: Nycretina [online]. [cit. 2023-05-11]. Dostupné z: https://www.nycretina.com/conditions/retinal-detachment/*

[26] *MAHANAMA, Bhanuka, Yasith JAYAWARDANA, Sundararaman RENGARAJAN, Gavindya JAYAWARDENA, Leanne CHUKOSKIE, Joseph SNIDER a Sampath JAYARATHNA, 2022. Eye Movement and Pupil Measures: A Review. Frontiers in Computer Science [online]. 3 [cit. 2023-05-09]. ISSN 2624-9898. Dostupné z: doi:10.3389/fcomp.2021.733531*

[27] *DUCHOWSKI, Andrew T., 2017. Eye Tracking Methodology [online]. Cham: Springer International Publishing [cit. 2023-05-09]. ISBN 978-3-319-57881-1. Dostupné z: doi:10.1007/978-3-319-57883-5*

[28] *PURVES, D, GJ AUGUSTINE a D FITZPATRICK, 2001. Neuroscience. 2nd Edition. Sinauer Associates. ISBN 0-87893-742-0.*

[29] *RUCCI, Michele a Martina POLETTI, 2015. Control and Functions of Fixational Eye Movements. Annual Review of Vision Science [online]. 1(1), 499-518 [cit. 2023-05-09]. ISSN 2374-4642. Dostupné z: doi:10.1146/annurev-vision-082114-035742*

[30] *GUNTON, Kammi B., Barry N. WASSERMAN a Caroline DEBENEDICTIS, 2015. Strabismus. Primary Care: Clinics in Office Practice [online]. 42(3), 393-407 [cit. 2023-05-09]. ISSN 00954543. Dostupné z: doi:10.1016/j.pop.2015.05.006*

[31] *HOLMES, Jonathan M a Michael P CLARKE, 2006. Amblyopia. The Lancet [online]. 367(9519), 1343-1351 [cit. 2023-05-09]. ISSN 01406736. Dostupné z: doi:10.1016/S0140-6736(06)68581-4*

[32] *TARNUTZER, Alexander A. a Dominik STRAUMANN, 2018. Nystagmus. Current Opinion in Neurology [online]. 31(1), 74-80 [cit. 2023-05-09]. ISSN 1350-7540. Dostupné z: doi:10.1097/WCO.0000000000000517*

[33] *TILIKETE, Caroline a Matthieu P. ROBERT, 2019. Ocular Motor Apraxia. In: SHAIKH, Aasef a Fatema GHASIA, ed. Advances in Translational Neuroscience of Eye Movement Disorders [online]. Cham: Springer International Publishing, 2019-01-04, s. 451-470 [cit. 2023-05-09]. Contemporary Clinical Neuroscience. ISBN 978-3-030-31406-4. Dostupné z: doi:10.1007/978-3-030-31407-1_22*

[34] *SCHALL, Andrew a Jennifer ROMANO BERGSTROM, 2014. Introduction to Eye Tracking. In: Eye Tracking in User Experience Design [online]. Elsevier, 2014, s. 3-26 [cit. 2023-05-16]. ISBN 9780124081383. Dostupné z: doi:10.1016/B978-0-12-408138-3.00001-7*

[35] *JACOB, Robert J.K. a Keith S. KARN, 2003. Eye Tracking in Human-Computer Interaction and Usability Research. In: The Mind's Eye [online]. Elsevier, 2003, s. 573-605 [cit. 2023-05-16]. ISBN 9780444510204. Dostupné z: doi:10.1016/B978-044451020-4/50031-1*

[36] *CARTER, Benjamin T. a Steven G. LUKE, 2020. Best practices in eye tracking research. International Journal of Psychophysiology [online]. 155, 49-62 [cit. 2023-05-16]. ISSN 01678760. Dostupné z: doi:10.1016/j.ijpsycho.2020.05.010*

[37] *COGNOLATO, Matteo, Manfredo ATZORI a Henning MÜLLER, 2018. Head-mounted eye gaze tracking devices: An overview of modern devices and recent advances. Journal of Rehabilitation and Assistive Technologies Engineering [online]. 5 [cit. 2023-05-18]. ISSN 2055-6683. Dostupné z: doi:10.1177/2055668318773991*

[38] *KUMARI, Niharika, Verena RUF, Sergey MUKHAMETOV, Albrecht SCHMIDT, Jochen KUHN a Stefan KÜCHEMANN, 2021. Mobile Eye-Tracking Data Analysis Using Object Detection via YOLO v4. Sensors [online]. 21(22) [cit. 2023-05-18]. ISSN 1424-8220. Dostupné z: doi:10.3390/s21227668*

[39] *GREDEBÄCK, Gustaf, Scott JOHNSON a Claes VON HOFSTEN, 2009. Eye Tracking in Infancy Research. Developmental Neuropsychology [online]. 35(1), 1-19 [cit. 2023-05-17]. ISSN 8756-5641. Dostupné z: doi:10.1080/87565640903325758*

[40] *KÄTHNER, Ivo; KÜBLER, Andrea; HALDER, Sebastian. Comparison of eye tracking, electrooculography and an auditory brain-computer interface for binary communication: a case study with a participant in the locked-in state. Journal of neuroengineering and rehabilitation, 2015, 12.1: 1-11.*

[41] *LI, Zhaowei, Peiyuan GUO a Chen SONG, 2021. A Review of Main Eye Movement Tracking Methods. Journal of Physics: Conference Series [online]. 1802(4) [cit. 2023-05-17]. ISSN 1742-6588. Dostupné z: doi:10.1088/1742-6596/1802/4/042066*

[42] *SHELHAMER, Mark a Dale C. ROBERTS, 2010. Magnetic scleral search coil. In: Vertigo and Imbalance: Clinical Neurophysiologyof the Vestibular System [online]. Elsevier, 2010, s. 80-87 [cit. 2023-05-17]. Handbook of Clinical Neurophysiology. ISBN 9780444529121. Dostupné z: doi:10.1016/S1567-4231(10)09006-4*

[43] *SINGH, Hari; SINGH, Jaswinder. Human eye tracking and related issues: A review. International Journal of Scientific and Research Publications, 2012, 2.9: 1-9.*

[44] *CHENNAMMA, H. R.; YUAN, Xiaohui. A survey on eye-gaze tracking techniques. arXiv preprint arXiv:1312.6410, 2013.*

[45] *PODOLEANU, A. Gh., 2012. Optical coherence tomography. Journal of Microscopy [online]. 247(3), 209-219 [cit. 2023-05-17]. ISSN 0022-2720. Dostupné z: doi:10.1111/j.1365-2818.2012.03619.x*

[46] *SKARAMAGKAS, Vasileios, Giorgos GIANNAKAKIS, Emmanouil KTISTAKIS, et al., 2023. Review of Eye Tracking Metrics Involved in Emotional and Cognitive Processes. IEEE Reviews in Biomedical Engineering [online]. 16, 260-277 [cit. 2023-05-17]. ISSN 1937-3333. Dostupné z: doi:10.1109/RBME.2021.3066072*

[47] *BRUNYÉ, Tad T., et al. A review of eye tracking for understanding and improving diagnostic interpretation. Cognitive research: principles and implications, 2019, 4: 1-16.*

[48] *IDS-imaging: UI-3060CP Rev. 2 camera [online]. [cit. 2023-01-02]. Dostupné z: https://en.ids-imaging.com/store/ui-3060cp-rev-2.html*

[49] *IDS-imaging: uEye Cockpit [online]. [cit. 2023-01-02]. Dostupné z: https://en.ids-imaging.com/ids-software-suite.html*

[50] *TSAI, Shang-Ru a Michael R. HAMBLIN, 2017. Biological effects and medical applications of infrared radiation. Journal of Photochemistry and Photobiology B: Biology [online]. 170, 197-207 [cit. 2023-05-19]. ISSN 10111344. Dostupné z: doi:10.1016/j.jphotobiol.2017.04.014*

[51] *B223 - Infrared spotlight. Kemo Electronics [online]. Germany [cit. 2023-05-19]. Dostupné z: https://www.kemo-electronic.de/en/Light-Sound/Infrared/Kits/B223-Infrared-spotlight.php*