NOVA SCHOOL OF SCIENCE & TECHNOLOGY

DEPARTMENT OF PHYSICS

MARIA DA CONCEIÇÃO FRANGO SABINO

BSc of Science in Biomedical Engineering

EYE TRACKING USING AN EYE MODEL

MASTER IN BIOMEDICAL ENGINEERING

NOVA University Lisbon

September, 2024

NOVA SCHOOL OF
SCIENCE & TECHNOLOGY

DEPARTMENT OF PHYSICS

EYE TRACKING USING AN EYE MODEL

MARIA DA CONCEIÇÃO FRANGO SABINO

BSc of Science in Biomedical Engineering

Adviser: Ricardo Nuno Pereira Verga e Afonso Vigário
Associate Professor, NOVA University Lisbon

Co-adviser: Carla Maria Quintão Pereira
Assistant Professor, NOVA University Lisbon

MASTER IN BIOMEDICAL ENGINEERING

NOVA University Lisbon
September, 2024

Eye Tracking Using An Eye Model

This document was created with the (pdf/Xe/Lua)LaTeX processor and the NOVAthesis template (v7.1.18) [1].

Acknowledgements

Em primeiro lugar, agradeço aos meus orientadores, professor Ricardo Vigário e professora Carla Quintão. Só através de toda a disponibilidade e ajuda demonstradas é que foi possível caminhar até aqui.

Em segundo lugar, ao profesor Paulo Ribeiro pelo empréstimo do suporte da queixeira e ao Pedro Rodrigues pela ajuda na concretização da mesma.

Não poderia deixar de agradecer a toda a minha família. A cada um que me moldou e motivou, desde pequenina, e me fez acreditar que é possível concretizar os desafios em que nos colocamos, sempre com carinho e amizade. Um especial agradecimento à minha mãe, ao meu pai e ao meu irmão, por toda a paciência, apoio e compreensão durante este percurso académico, ao meu Avô Licas por ser o confidente da família e a companhia das tardes ao computador, à minha Madrinha querida e aos meus primos Sofy e Tigui.

À TunaMaria, por ser fonte de crescimento e inspiração, trazendo pessoas e ensinamentos incríveis para a minha vida. À Carolina, à Inês D., à Inês R. e à Maria, pelo incrível caminho em equipa que percorremos este ano, por saberem ouvir, fazer rir e, também, consolar. Agradeço à FCT por me ter dado o privilégio de estar neste projeto que vai muito para além da música.

À Ana Júlia por saber sempre o que preciso.

Aos meus eternos amigos de Évora, ao melhor quarteto e à Inês B.

Às minhas colegas de casa, por serem conforto, amizade e apoio nos serões de estudo.

Ao Ricardo e à Carolina por ajudarem a que a apresentação fosse mais leve.

Por fim, agradeço a todos aqueles que passaram por mim durante estes anos e que, mesmo não se apercebendo, tornaram-nos mais especiais.

”

“Oh Avô,
é a lua!”

Abstract

The study of eye movements is crucial for the clinical assessment in diagnosing, monitoring, and categorizing eye diseases and neurological disorders.

With this in mind, a platform entitled RehabVisual was created, in previous work, to provide individualized treatment so that, through stimuli, it could help the visuomotor development of children and be a diagnostic and therapeutic tool for patients who have suffered a stroke.

The eye tracker developed for the platform had some flaws when detecting the iris in certain lighting conditions and when the eyelids were very closed.

Intending to combat some detection flaws and make the eye tracker more robust, this dissertation proposes adding more morphological information about the eye, in addition to the position of the iris, which has already been used.

Initially, more detailed pre-processing was carried out, trying to preserve the integrity of the structures as much as possible. Subsequently, the method employed involved analyzing the structural changes in the sclera and iris as the gaze direction shifted. This was achieved by correlating two matrices: one representing the initial position of these structures and the other containing data from each subsequent frame to be analyzed.

Two devices, a computer and a smartphone, were used simultaneously to record videos of the performances of four individuals who were all free of associated disorders.

The results of one individual have been represented in order to illustrate how detection behaves for each eye in each direction (horizontal and vertical). Next, the results of all the individuals were compared exhaustively with the stimulus. For example, one of the cases in which there had previously been a considerable loss of information was highlighted, and, with the changes made, there was a reduction in the error in horizontal detection from 41% to 7% and in vertical detection from 34% to 21% for the left eye.

The analysis helped to identify how different lighting conditions and eyelid opening, as well as structural differences in the eyes, influence the success of eye tracking. It also proved that it is possible to use the computer's camera to successfully follow the eye movement.

Resumo

O estudo dos movimentos oculares é crucial para a avaliação clínica no diagnóstico, monotorização e categorização de doenças oculares e perturbações neurológicas.

Neste sentido, foi criada uma plataforma, em trabalhos anteriores, intitulada Rehab-Visual, com o objetivo de proporcionar um tratamento individualizado para que, através de estímulos, ajudasse no desenvolvimento visuomotor de crianças e atuasse como uma ferramenta diagnóstica e terapêutica para pacientes que sofreram um AVC.

O eye tracker desenvolvido para a plataforma apresentava algumas falhas na deteção da íris em determinadas condições de luminosidade e quando as pálpebras estavam muito fechadas. Com o objetivo de combater algumas falhas de deteção e tornar o eye tracker mais robusto, esta dissertação propõe adicionar mais informação morfológica sobre o olho, para além da posição da íris, que já tem sido utilizada.

Inicialmente, foi realizado um pré-processamento mais detalhado, tentando preservar ao máximo a integridade das estruturas. Posteriormente, o método utilizado consistiu em analisar as alterações estruturais da esclera e da íris à medida que a direção do olhar se alterava. Para tal, foram correlacionadas duas matrizes: uma que representa a posição inicial destas estruturas e outra que contém os dados de cada frame subsequente que será analisado.

Dois dispositivos, um computador e um smartphone, foram utilizados, simultaneamente, para gravar vídeos de quatro indivíduos sem patologias associadas.

Os resultados de um dos indivíduos foram representados, a fim de ilustrar o comportamento da deteção para cada olho em cada direção (horizontal e vertical). De seguida, os resultados de todos os indivíduos foram comparados exaustivamente com o estímulo. Por exemplo, num dos casos em que anteriormente se registava uma perda considerável de informação, com as alterações efectuadas, verificou-se uma redução do erro na deteção horizontal de 41% para 7% e na deteção vertical de 34% para 21% para o olho esquerdo.

A análise ajudou a identificar a forma como as diferentes condições de iluminação e abertura das pálpebras, bem como as diferenças estruturais dos olhos, influenciam o sucesso do seguimento do movimento ocular. Também provou que é possível utilizar a câmara do computador, para seguir com êxito o movimento dos olhos.

Contents

List of Figures viii

List of Tables xii

Acronyms xiii

1 Introduction 1
- 1.1 Context and Motivation . . . 1
- 1.2 Objectives . . . 2

2 Literature Review About The Visual System 3
- 2.1 Anatomy Of The Eye . . . 3
- 2.2 Formation Of An Image On The Retina . . . 4
- 2.3 Functions of Eye Movements . . . 5

3 State-of-the-art 6
- 3.1 Eye Tracking Techniques . . . 6
  - 3.1.1 Scleral Contact Lens/Search Coil . . . 6
  - 3.1.2 Electro-OculoGraphy . . . 7
  - 3.1.3 Photo-OculoGraphy or Video-OculoGraphy . . . 8
  - 3.1.4 Video-Based Combined Pupil/Corneal Reflection . . . 8
- 3.2 Developed Eye Tracker of the RehabVisual Platform . . . 10

4 Methodology 12
- 4.1 Experimental Setup . . . 12
- 4.2 Data Collection . . . 13
- 4.3 Data Processing . . . 13

5 Improvements To The Eye Tracker 16
- 5.1 Automatic Calibration . . . 16
- 5.2 Image Processing . . . 17

5.2.1 Channel Separation . . . 17
5.2.2 Calculating Distances and Iris Enhancement . . . 17
5.2.3 Applying a Gaussian Filter . . . 18
5.3 Eye Detection . . . 19
5.3.1 Eye Model . . . 19
5.3.2 Eye Movement . . . 20

6 Results **23**
6.1 Data Acquired With The Computer's Camera . . . 23
6.1.1 Left Eye . . . 23
6.1.2 Right Eye . . . 24
6.2 Data Acquired With The Smartphone's Camera . . . 25
6.2.1 Left Eye . . . 25
6.2.2 Right Eye . . . 26

7 Discussion of Results **28**
7.1 Metrics Extracted by Comparison with the Stimulus . . . 28
7.1.1 Stimulus Processing . . . 28
7.1.2 Frames in Area of Interest . . . 29
7.1.3 Distance to the Stimulus . . . 31
7.2 Previous vs Current Eye Tracker . . . 34
7.3 Left Eye vs Right Eye . . . 36
7.4 Horizontal Displacement vs Vertical Displacement . . . 36
7.5 Computer's Camara vs Smartphone's Camara . . . 37

8 Conclusion and Future Perspectives **38**

Bibliography **40**

Appendices

A Visualization of Results **42**
A.1 Data Acquired With The Computer's Camera . . . 42
A.1.1 Left Eye . . . 42
A.1.2 Right Eye . . . 45
A.2 Data Acquired With The Smartphone's Camera . . . 47
A.2.1 Left Eye . . . 47
A.2.2 Right Eye . . . 50

List of Figures

3.1 Example of search coil embedded in contact lens and electromagnetic field frames. . . . 7
3.2 Example of electro-oculography. . . . 8
3.3 Example of infrared limbus tracker apparatus. . . . 8
3.4 Example of Video-Based Combined Pupil/Corneal Reflection. . . . 9
3.5 Tobii Pro Fusion . . . 10
3.6 An illustration of a corresponding gaze position (a) and a non-corresponding gaze position (b). . . . 11

4.1 Experimental setup (a) (b). . . . 13
4.2 Stimulus used (a). Areas of interest defined for the stimulus (b). . . . 13
4.3 Horizontal displacements as function of frame count . . . 14
4.4 Horizontal displacements, after process 2, as function of frame count . . . 14
4.5 Horizontal displacements belonging to the clusters . . . 15
4.6 Horizontal displacements, after process 4, as function of frame count . . . 15

5.1 Right Eye (a). Left Eye (b). . . . 16
5.2 Right Eye RGB Channels: Red, Green, and Blue Components . . . 17
5.3 Comparison of the multiplication of the normalized channel distance by each individual color channel . . . 18
5.4 Before (a) and After (b) the application of the Gaussian Filter. . . . 18
5.5 Eye Model for the Horizontal Component . . . 19
5.6 Eye Model for the Vertical Component . . . 20

6.1 Graph of the horizontal position of person 1's left eye throughout the frames captured by the computer camera. . . . 23
6.2 Graph of the vertical position of person 1's left eye throughout the frames captured by the computer camera. . . . 24
6.3 Graph of the horizontal position of person 1's right eye throughout the frames captured by the computer camera. . . . 24

6.4 Graph of the vertical position of person 1's right eye throughout the frames captured by the computer camera. 25
6.5 Graph of the horizontal position of person 1's left eye throughout the frames captured by the smartphone camera. 25
6.6 Graph of the vertical position of person 1's left eye throughout the frames captured by the smartphone camera. 26
6.7 Graph of the horizontal position of person 1's right eye throughout the frames captured by the smartphone camera. 26
6.8 Graph of the vertical position of person 1's right eye throughout the frames captured by the smartphone camera. 27

7.1 Graph of the horizontal positions of the normalized stimulus, as well as the tolerance values obtained from its radius. 28
7.2 Graph of the vertical positions of the normalized stimulus, as well as the tolerance values obtained from its radius. 29
7.3 Horizontal component of the stimulus, as well as its tolerance values, along with the result of the detection from the left eye of individual 1, through the computer camera. 31
7.4 Vertical component of the stimulus, as well as its tolerance values, along with the result of the detection from the left eye of person 1, through the computer camera. 32
7.5 Frame taken from person 2's video, acquired by the computer's built-in camera, cropped manually. 34
7.6 Graph of the horizontal position of person 2's left eye throughout the frames captured by the computer camera. 34
7.7 Graph of the vertical position of person 2's left eye throughout the frames captured by the computer camera. 35

A.1 Graph of the horizontal position of person 1's left eye throughout the frames captured by the computer camera. 42
A.2 Graph of the horizontal position of person 2's left eye throughout the frames captured by the computer camera. 42
A.3 Graph of the horizontal position of person 3's left eye throughout the frames captured by the computer camera. 43
A.4 Graph of the horizontal position of person 4's left eye throughout the frames captured by the computer camera. 43
A.5 Graph of the vertical position of person 1's left eye throughout the frames captured by the computer camera. 43
A.6 Graph of the vertical position of person 2's left eye throughout the frames captured by the computer camera. 44

A.7 Graph of the vertical position of person 3's left eye throughout the frames captured by the computer camera. 44
A.8 Graph of the vertical position of person 4's left eye throughout the frames captured by the computer camera. 44
A.9 Graph of the horizontal position of person 1's right eye throughout the frames captured by the computer camera. 45
A.10 Graph of the horizontal position of person 2's right eye throughout the frames captured by the computer camera. 45
A.11 Graph of the horizontal position of person 3's right eye throughout the frames captured by the computer camera. 45
A.12 Graph of the horizontal position of person 4's right eye throughout the frames captured by the computer camera. 46
A.13 Graph of the vertical position of person 1's right eye throughout the frames captured by the computer camera. 46
A.14 Graph of the vertical position of person 2's right eye throughout the frames captured by the computer camera. 46
A.15 Graph of the vertical position of person 3's right eye throughout the frames captured by the computer camera. 47
A.16 Graph of the vertical position of person 4's right eye throughout the frames captured by the computer camera. 47
A.17 Graph of the horizontal position of person 1's left eye throughout the frames captured by the smartphone camera. 47
A.18 Graph of the horizontal position of person 2's left eye throughout the frames captured by the smartphone camera. 48
A.19 Graph of the horizontal position of person 3's left eye throughout the frames captured by the smartphone camera. 48
A.20 Graph of the horizontal position of person 4's left eye throughout the frames captured by the smartphone camera. 48
A.21 Graph of the vertical position of person 1's left eye throughout the frames captured by the smartphone camera. 49
A.22 Graph of the vertical position of person 2's left eye throughout the frames captured by the smartphone camera. 49
A.23 Graph of the vertical position of person 3's left eye throughout the frames captured by the smartphone camera. 49
A.24 Graph of the vertical position of person 4's left eye throughout the frames captured by the smartphone camera. 50
A.25 Graph of the horizontal position of person 1's right eye throughout the frames captured by the smartphone camera. 50
A.26 Graph of the horizontal position of person 2's right eye throughout the frames captured by the smartphone camera. 50

A.27 Graph of the horizontal position of person 3's right eye throughout the frames captured by the smartphone camera. 51
A.28 Graph of the horizontal position of person 4's right eye throughout the frames captured by the smartphone camera. 51
A.29 Graph of the vertical position of person 1's right eye throughout the frames captured by the smartphone camera. 51
A.30 Graph of the vertical position of person 2's right eye throughout the frames captured by the smartphone camera. 52
A.31 Graph of the vertical position of person 3's right eye throughout the frames captured by the smartphone camera. 52
A.32 Graph of the vertical position of person 4's right eye throughout the frames captured by the smartphone camera. 52

List of Tables

7.1 Comparison of frames acquired by the computer for the left and right eyes, included in the regions of interest 29
7.2 Comparison of frames acquired by the smartphone for the left and right eyes, included in the regions of interest 30
7.3 Comparison of frames acquired by the computer for the left and right eyes, included in the regions of interest 30
7.4 Comparison of frames acquired by the smartphone for the left and right eyes, included in regions of interest 30
7.5 Average Distances between the stimulus and the eyes recorded by the computer 32
7.6 Average Distances, in pixels, between the stimulus and the eyes recorded by smartphone 33
7.7 Average distances, in pixels, between the stimulus and the eyes recorded by the smartphone. 35
7.8 Average distances, in pixels, between the stimulus and the eyes recorded by the computer. 35

Acronyms

2D Two-Dimensional *(pp. 18, 21)*

3D Three-Dimensional *(p. 12)*

DBSCAN Density-Based Spatial Clustering of Applications with Noise *(p. 14)*

EOG Electro-Oculography *(p. 7)*

FPS Frame Per Second *(p. 12)*

HD High Definition *(p. 12)*

ONK Optokinetic reflex *(p. 5)*

VOR Vestibular-Ocular Reflex *(p. 5)*

1

Introduction

1.1 Context and Motivation

At least 1 billion people, worldwide, suffer from visual impairments that could have been avoided or may be yet unaddressed. In the absence of timely detection, diminished or nonexistent eyesight can have long-term personal and economic consequences. Young children, with early onset severe vision impairment, may have lower levels of academic accomplishment and, in adults, may affects quality of life by lowering productivity, decreasing workforce engagement, and increasing rates of depression [2].

Assessing a child's visual-motor abilities, in their early years, is essential for monitoring development at a time when cerebral plasticity at its highest. Targeted visual stimulation is necessary in some circumstances to promote brain plasticity, particularly when there are developmental deficits in visual-motor skills [3].

Early research on attention was predominantly based on introspection and simple visual observations, where participants were instructed to self-report their internal cognitive processes during tasks. While these introspective methods provided initial insights, they were limited by subjectivity and a lack of replicability. [4] As the field advanced, introspection was gradually supplanted by more objective methodologies, such as reaction time experiments and controlled observational studies. Since then, the field has expanded to encompass disciplines such as computer science, cognitive neuroscience, and psychophysics.

In this context, to evaluate visual skills in babies with developmental delays and provide intervention, a platform named "RehabVisual" was designed in Raquel Machado's and Catarina Santos's master dissertations. [5] [6] In addition to these studies, Pedro Dias tested the utility of this tool in individuals who have suffered a stroke, studying its significant impact and the advantages it provides for this population. In his thesis, an eye tracker was developed, to capture eye movements, from a distance, and integrated it into RehabVisual. [7] Finally, Pedro Fonseca's dissertation tested its usability, comparing the tool's performance with the Tobii Pro Nano eye tracker [8].

Despite the promising initial results regarding the tool's capacity, to support the

intended rehabilitation, as well as its ability to characterize eye movements, it was noticed a significant decline in the quality of tracking the direction of visual attention, when the user's iris is very dark or their eyelids are almost closed. As a result, the suggested tool's clinical applications have opted to use a specialized webcam to collect images instead of a single device for both stimulation and data recording.

1.2 Objectives

This dissertation aims to build upon previous work, particularly on enhancing the eye-tracking tool, improving its ability to estimate eye direction. The main aim is to include eye model information, in order to make eye tracker more robust, enabling a single platform to handle the individual's stimulus, data collection, and performance evaluation. In this regard, the aim is to:

1. Implement an initial automatic calibration for the eye tracker, eliminating the need for manually selecting the location of the eyes.
2. Add morphological information about the eye, in addition to the iris already used, to estimate the direction of gaze more accurately.
3. Applying a stimulus and then analyzing the results in a group of non-pathologically exposed individuals over the age of 18 in order to represent a control group, minimizing possible errors in following the stimulus.
4. Compare the results collected simultaneously from the computer's built-in camera with those from a dedicated external camera, in order to understand how the difference in the positions of the devices relative to the eye affects the results obtained.

The research work described in this dissertation was carried out in accordance with the norms established in the ethics code of Universidade NOVA de Lisboa. The work described and the material presented in this dissertation, with the exceptions clearly indicated, constitute original work carried out by the author.

2

Literature Review About The Visual System

Vision is a crucial sense for human beings. It enables exploration and the understanding of the world, facilitating the interaction with one's environment [9].

This remarkable ability is made possible by the visual system, which is composed of the eyes, accessory structures (eyelids, eyebrows, eyelashes, lacrimal canals, and extrinsic muscles), optic nerves, the optic tract, and optic pathways that are crucial for processing and interpreting visual information in the visual cortex [4].

2.1 Anatomy Of The Eye

Three layers make up the anatomical wall of the eyeball: the fibrous tunic, the vascular tunic, and the retina (inner tunic) [10].

1. **Fibrous Tunic** The anterior cornea and posterior sclera form the fibrous tunic, the outermost layer of the eyeball. The colored iris is covered by a transparent layer called the cornea. The sclera, a layer of dense connective tissue, is primarily composed of fibroblasts and collagen fibers. With the exception of the cornea, the entire eyeball is encased in the sclera, which serves as an attachment point for the extrinsic eye muscles and provides shape and rigidity to the eyeball [10].

2. **Vascular Tunic** The middle layer of the eye, known as the vascular tunic, consists of the ciliary body, choroid, and iris. The choroid, which is the posterior part of the vascular tunic, lines most of the internal surface of the sclera and is highly vascularized. The ciliary body contains ciliary processes and muscles. The ciliary muscle is a smooth, circular band of muscle that adjusts the tension of the zonular fibers. This change in tension alters the curvature of the lens, allowing for accommodation of near or far vision. The iris, which is the colored part of the eye shaped like a donut, displays different colors based on the concentration of melanin: high concentrations result in brown or black eyes, low concentrations yield blue eyes, and

moderate concentrations produce green eyes. One of the primary functions of the iris is to control the amount of light entering the eye through the pupil, the opening in the center of the iris. The densely pigmented back of the eye (choroid and retina) is visible when looking through the lens, giving the pupil its characteristic black appearance [10].

3. **Retina**

   The retina, the innermost layer of the eyeball, covers the majority of the posterior section of the eye and plays a critical role in the start of visual pathways, processesing a large amount of visual information before transmitting it via optic nerve axons. This brain layer contains approximately 120 million rods, that are sensitive to low light, and 6 to 7 million cones, that are responsible for color perception and visual acuity. [10] The cones require relatively bright light to function and are capable of discriminating between different frequencies of incident light, allowing us to perceive color. In contrast, the rods are extremely sensitive to light but cannot differentiate between frequencies, meaning they primarily contribute to the perception of different shades of gray. As the light decreases, objects appear in varying tones of gray due to the predominant activation of rods. Located at the back center of the retina is the macula, with its central point, the fovea, where light is most powerfully concentrated when focusing on an object. The fovea offers maximal visual acuity due to the high concentration of cones (about 35,000) and the absence of rods. It's important to note that, during each visual fixation, the fovea captures less than 8% of the visual field, yet it accounts for 50% of the information sent to the brain via the optic nerve. The remaining portion of the visual field is referred to as the peripheral region, which captures less detailed information but is crucial for detecting motion and contrast. The optic disc, also known as the blind spot of the retina, is another important structure within the retina. It is the point where the optic nerve emerges from the retina, and it contains no photoreceptor cells, meaning it does not respond to light [9].

2.2 Formation Of An Image On The Retina

When light, regulated by the pupil, passes through the eye's lens and vitreous humor, it is refracted and converges on the retina. This stimulates the cones and rods, generating an electrical impulse that is transmitted via the optic nerves to the visual cortex. Although the image formed on the retina is reversed and inverted, the brain perceives it in an upright position. [10] [9]

2.3 Functions of Eye Movements

Eye movements serve three main functions in visual processing: [11]

1. Rapidly redirecting the gaze to place relevant information on the fovea through fixations and saccadic movements. These movements can occur up to four times per second, causing brief periods of blindness as the eyes reposition.

2. Maintaining a stable image on the retina by compensating for head or object movements through reflexes like the Vestibular-Ocular Reflex (VOR) and the Optokinetic reflex (ONK).

3. Preventing stationary objects from fading from view by using micro-saccades, tremors, and drift to stabilize the gaze.

3

State-of-the-art

3.1 Eye Tracking Techniques

Understanding the complexity of human cognition has long been challenging, with early attempts limited by the rudimentary tools available [12]. However, in the past fifty years, technological advancements have changed our ability to probe and visualize cognitive and physiological processes. One of the most significant innovations, in this realm, is eye-tracking technology, which has provided unprecedented insights into the workings of the human visual system [13].

Eye trackers, sophisticated instruments designed to measure and record eye movements, have become indispensable tools in cognitive science and psychological research. These devices have enhanced our comprehension of visual perception, attention, and cognitive processes by precisely capturing how individuals engage with visual stimuli. The integration of eye-tracking data with other physiological measures has further refined our understanding of the brain's response mechanisms [13].

3.1.1 Scleral Contact Lens/Search Coil

The use of contact lenses for eye-tracking represents one of the most precise methods for monitoring eye movements. This approach has evolved significantly over time, with early techniques paving the way for modern innovations. In the late 1800s, one of the earliest methods involved a Paris ring plaster, which was mechanically connected to recording pens and placed directly on the cornea. This method laid the groundwork for future advancements by demonstrating the potential for directly interfacing with the eye to record movements [14]. Edmund Huey made a significant contribution to this field by constructing the first eye tracker using contact lenses with an aluminum pointer attached. The pointer moved in response to eye movements, and variations in the electromagnetic field around the pointer were used to measure these movements. While this approach was quite accurate for its time, it had the considerable drawback of being invasive and uncomfortable [4].

The modern version of this technique is exemplified by the scleral search coil method,

illustrated in Fig. 3.1. In this approach, the contact lens is designed to cover both the cornea and part of the sclera, see left most frame in the Fig. 3.1. It contains wire coils embedded within. The participant is positioned within an electromagnetic field, illustrated in image on the right of the Fig. 3.1, and the movement of the eye affects the electromagnetic field around the coils. The resulting variations in the potential difference across the coils are used to determine precise eye movements. Although this method offers high accuracy, it remains invasive, causing discomfort and requiring careful handling during insertion [15].

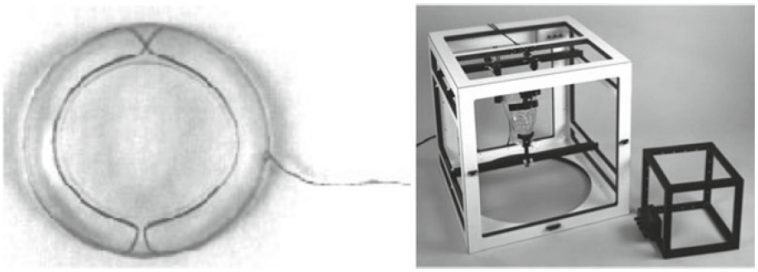


Figure 3.1: Example of search coil embedded in contact lens and electromagnetic field frames. Adapted from: [16]

3.1.2 Electro-OculoGraphy

Following a similar approach to the one presented above, and exploiting natural electromagnetic properties of the eyes, around 40 years ago, Electro-Oculography (EOG) emerged as the most popular technique for recording eye movements. This method, still in use today, operates by measuring the electric potential variations in the skin around the eyes using electrodes. The principle behind EOG is based on the natural electric potential difference between the cornea and the retina [17]. The cornea holds a positive potential relative to the back of the eye, creating a dipole with a potential difference ranging from 0.4 to 1.0 mV. As the eye moves, this dipole rotates, altering the electric potential measured on the skin around the eyes [4].

In a typical setup, three electrodes are placed on the skin near the eyes—one above, one below, and one on the temple, yet, as illustrated in Fig. 3.2, that number may be increased to acquire more precise readings. These electrodes capture the small potential differences associated with the vertical and horizontal movements of the ocular dipole. The recorded signals, which range from 15 to 200 μV with a nominal sensitivity of about 20 μV per degree of eye movement, are then used to infer the eye's position [4].

One of the key advantages of EOG is its non-invasive nature, allowing for the recording of eye movements even with closed eyes. This feature has made EOG particularly useful in clinical studies, such as monitoring eye movements during sleep. However, the technique requires careful electrode placement by a skilled technician and does not offer the same accuracy and precision as more modern eye-tracking systems. Additionally, because EOG

measures eye movements relative to head position, it is often not suitable for point-of-view tracking unless combined with a head tracker [4].

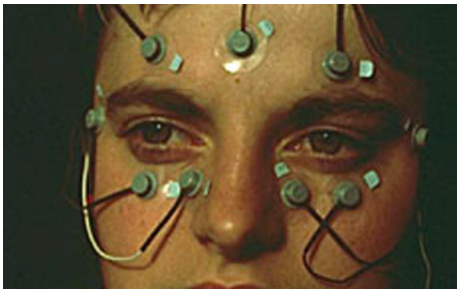


Figure 3.2: Example of electro-oculography. Adapted from: [18]

3.1.3 Photo-OculoGraphy or Video-OculoGraphy

This category includes many different methods for recording eye movements that involve measuring characteristics of the under rotation or translation of the eye, such as the apparent pupil shape, the location of the limbus (the iris-sclera boundary), and the corneal reflections of a directed light source that is positioned closely (typically infrared). Fig. 3.3 shows equipment used for limbus tracking and/or photo- or video-oculography where we can notice the presence of the light source close to the eyes. The ocular properties that these evaluating techniques provide can be measured either automatically or by visually inspecting recorded eye movements, which are usually recorded on videotape [4].

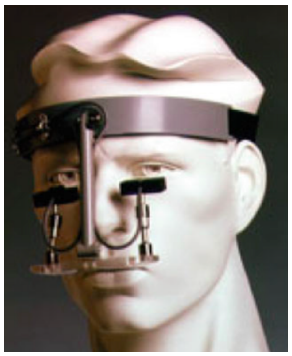


Figure 3.3: Example of infrared limbus tracker apparatus. Adapted from: [19]

3.1.4 Video-Based Combined Pupil/Corneal Reflection

Video-based trackers compute the point of regard in real time using relatively affordable cameras and image-processing gear. The device can be worn on the head or placed on a table, as illustrated in Fig. 3.4 . The location of the pupil's center is used to quantify the corneal reflection of the light source, which is usually infrared. These eye trackers are able to measure a viewer's point of attention on a correctly positioned (perpendicularly flat) surface on which calibration points are presented, provided that the proper calibration

processes are followed. To distinguish between head and eye motions, two points of reference on the eye are required. Minor head motions do not significantly alter the positional difference between the pupil center and corneal reflection, but pure eye rotation does [4].

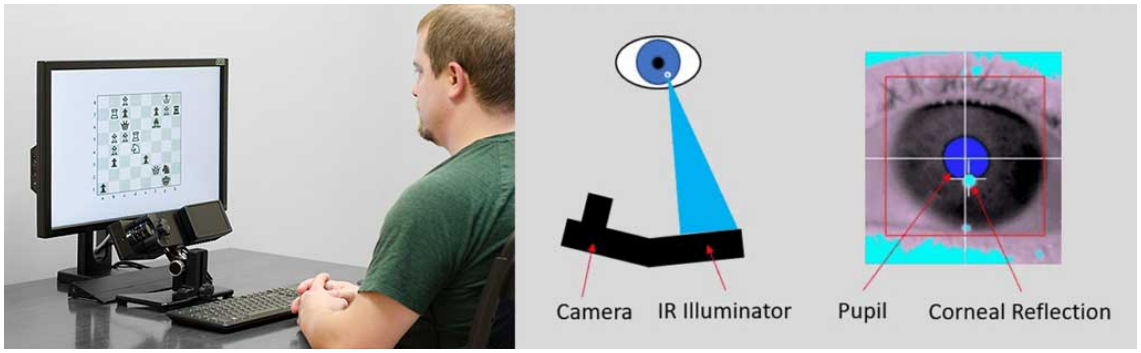


Figure 3.4: Example of Video-Based Combined Pupil/Corneal Reflection. Adapted from: [20]

3.1.4.1 Eye tracking programs

The eye tracker developed in this dissertation is part of this group of devices that, based on images and pattern detection, already have very varied programs. Among them is *Opened Eye*, which uses Matlab to provide eye tracking from both infrared and visible spectrum illumination; *xLabs*, *MyEye*, and *GazePointer*, which work with a standard webcam but do not allow for advanced data analysis; and *Ogama* and *PyGaze*, which enable the display of stimuli and provide data analysis tools. [21] [22] Additionally, other free programs such as *WebGazer*, *GazeRecorder*, and *IntelliGaze* require only the built-in camera of a laptop, eliminating the need for infrared light sources. However, this convenience can come with trade-offs, such as potential precision errors caused by small head movements or changes in ambient lighting. Calibration methods for these tools often involve the sequential display of calibration points on the screen, with the number of points influencing the accuracy of the eye-tracker.

The *Tobii Pro Fusion* is one example of an eye-tracking gadget and is considered the gold standard in the field. With a USB connection, it can be used with a variety of standalone monitors and laptop screens up to 24". Multiple sample rates of up to 250 Hz allow for gathering data on eye movements such as fixations, saccades, and smooth pursuit. In addition, data on pupil diameter and eye-opening can be recorded. The device generates visual representations of data, computes eye-tracking metrics, and exports data for further investigation in other software [23].

Figure 3.5: Tobii Pro Fusion [23]

3.2 Developed Eye Tracker of the RehabVisual Platform

This technology was developed using Matlab R2017a software and the analysis in the current section is carried out offline, requiring prior recording of the user interact with the screen. Yet, most strategies used may be easily transformed to online.

Initially, the subject is presented with a stimulus on the computer screen and asked to follow it, using only eye movements. During the stimulation, the subject's performance is recorded using a camera, so that the results can later be analyzed using eye tracking.

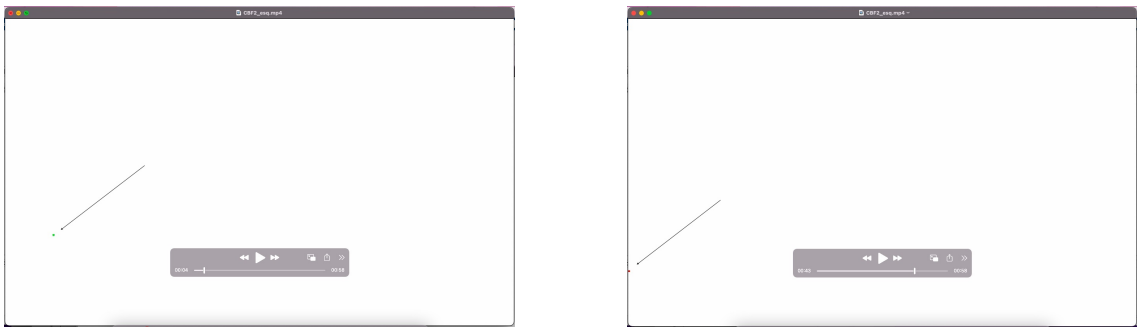
Then, for performance analysis, in the first frame of the recorded video, the analyser is asked to mark the subject's eyes, with the mouse of the computer, and a reference point, usually the nose. The process of detecting the eyes at each frame is automatic, and the user is only required to intervene if the detection fails, in which case case the reference points are lost, and they need to be marked again. Manual marking makes it possible to create two boxes of fixed dimensions, that include the eye's regions, to follow the movement of the eyes during the entire video.In order to extract the information of interest and facilitate detection, from each frame of the video, image processing is carried out on each box, making it easier to find the coordinates of the center of the pupil.

The process begins by converting the image under each box to grayscale. Subsequently, the darkest pixel is found, i.e. the one with the lowest intensity value, which should belong to the pupil. A technique for segmenting the image into regions with similar intensity values is then applied using, for example, the grayconnected() function. This method requires a manually defined tolerance factor, which controls how much the pixel intensities can vary to be considered connected. The factor ranges between 0 and 1, with higher values allowing more variation. The proper value depends on the image's lighting conditions, as higher tolerances may be necessary in low contrast situations while lower values work better in well-lit, high-contrast images.

In the next step, assuming that the iris has a circular shape, the imfindcircles() function is used, which is programmed to find circles in an image, saving the center's coordinates and the radius's value. These values are kept in arrays, along with the reference points. By making the user look towards the edges of the screen, in the first 16 seconds of the stimulus being presented, the maximum and minimum deviations between the initial

position of the eyes and their position when they follow the stimulus towards the edges are calculated. Knowing the size of the screen in pixels, a conversion factor is obtained, which is used to find the gaze positions. The end result is a matrix of coordinates, in pixels, of the position of the gaze for each eye in each frame. In order to analyze the relationship between the direction of the individual's gaze and the stimulus presented, the distances between the eyes, represented by the estimated centers of their pupils, together with the reference point, and the point on the screen where the user should be looking, are converted into pixels.

The software also makes a video of the displacement of the gaze position, characterized by a green square on a white background, Fig. 3.6a. When the gaze position does not correspond to a valid screen position, the color of the square turns red, Fig. 3.6b, and its position is displayed against the nearest valid screen edge.



(a) (b)

Figure 3.6: An illustration of a corresponding gaze position (a). An illustration of a non-corresponding gaze position (b). Adapted from: [8]

4

Methodology

4.1 Experimental Setup

One of the aims of this dissertation is to analyze and compare results obtained using the computer's built-in camera and these from an external camera in order to understand the viability of successfully collecting data on different devices, despite their different positions and internal characteristics. In this way, the experimental setup was designed so that it would be possible to record the individual's performance simultaneously on both devices.

The computer used is a 2019 MacBook Air, with an integrated Retina display monitor, which has a 13.3" screen with a resolution of 2560 x 1600 pixels. The integrated camera is the FaceTime High Definition (HD) camera, which offers 1280 x 720 (720p) HD video quality with 30 Frame Per Second (FPS) recording. The smartphone used as the external camera is an Apple iPhone 11, which has an integrated camera with 1920x1080 (1080p) resolution recording at 30 FPS.

To maximize the stability of the participant's head, the support shown in Fig. 4.1 was used. Its main components are a chin rest, built via Three-Dimensional (3D) printing, and a strap used as a forehead rest. The computer monitor was placed at a 90º angle to the table, so that the integrated camera was at a distance of approximately 66 cm from the support and at a height of 31 cm. The smartphone camera was placed at a 45º angle to the table, at a distance of 34 cm from the stand. The computer was positioned at approximately the height of the eyes. The smartphone was positioned as close as possible to it, but preventing any oclusion of the screen.

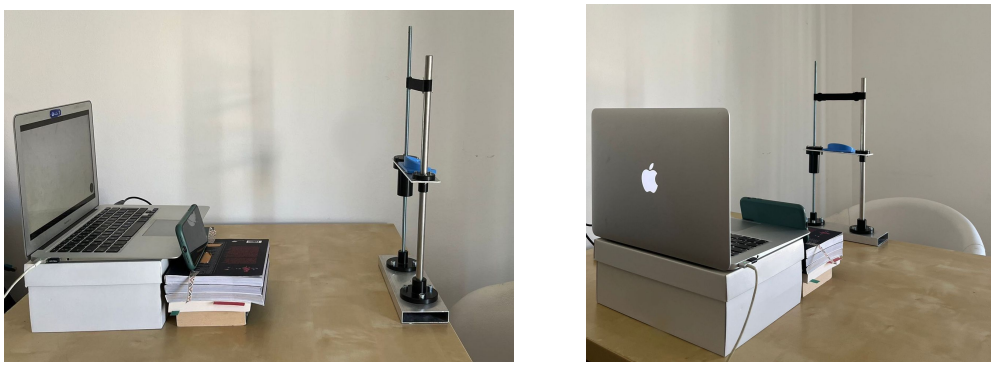
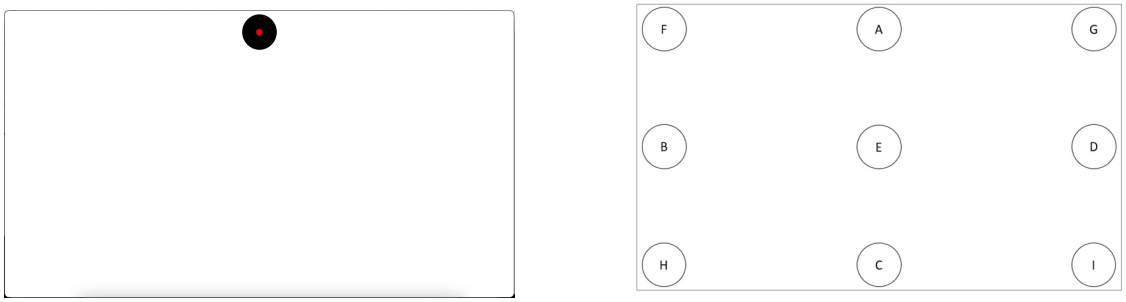


Figure 4.1: Experimental setup (a) (b).

4.2 Data Collection

Data were collected from a sample of four female volunteers, aged 22 to 28, without any known pathologies.

The experimental protocol for data collection was based on a 50-second video created in Power-Point consisting of a black circular stimulus with a red center, on a white background, as showed in Fig. 4.2a, which moves along the path A-B-C-D-E-A-C-I-F-H-G, according to the areas defined in Fig. 4.2b. Before starting the movement, the stimulus remained static for a period of 3 seconds and, at each letter of the path, it remained static for 1 second. During collection, the study participants viewed the video only once, while their gaze was recorded by both cameras. In order to ensure that the videos had maximum agreement and some initial synchronization, for both sides, each subject was asked to look at the computer camera until the stimulus started, as this was the region closest to it. The videos were also manually cut so that they ended up as synchronized as possible, since it was not feasible to do this at the same time.



(a) (b)

Figure 4.2: Stimulus used (a). Areas of interest defined for the stimulus (b). [8]

4.3 Data Processing

Below is a description of each step applied during processings:

1. Store and display horizontal displacements as a function of frame count, as it is shown in Fig. 4.3.

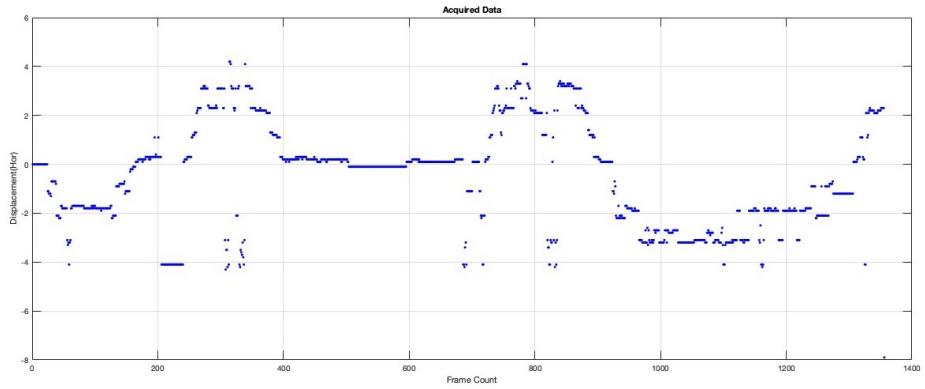


Figure 4.3: Horizontal displacements as function of frame count

2. A passband filter is then applied to remove discrepant values in the horizontal displacements, keeping only values that are between the maximums and minimums found. This is because it is assumed that the graph's maximums and minimums correspond to the maximum displacement values to the right and left, so values outside this range are not taken into account. After that data is normalized and the result is illustrated in the Fig. 4.4.

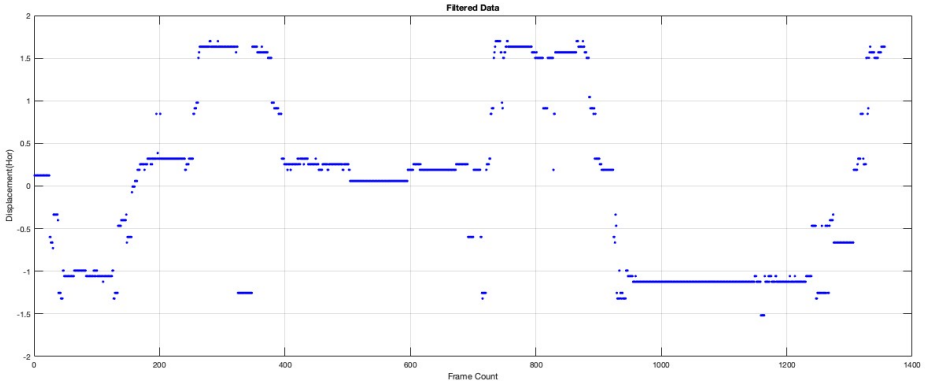


Figure 4.4: Horizontal displacements, after process 2, as function of frame count

3. A segmentation technique called Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [24] is applied. Afterwards, this technique is used to identify clusters, or groupings of data, which correspond to horizontal displacement patterns that can be grouped based on the proximity of the points. In this case, DBSCAN groups displacements are based on a maximum distance of 0.1 pixels between points and a minimum number of 20 points. The clusters are illustrated in the Fig. 4.5.

4. Displacement values that are between values belonging to a cluster but are not part of it are then "pulled" into the cluster and assigned the average value of that cluster.

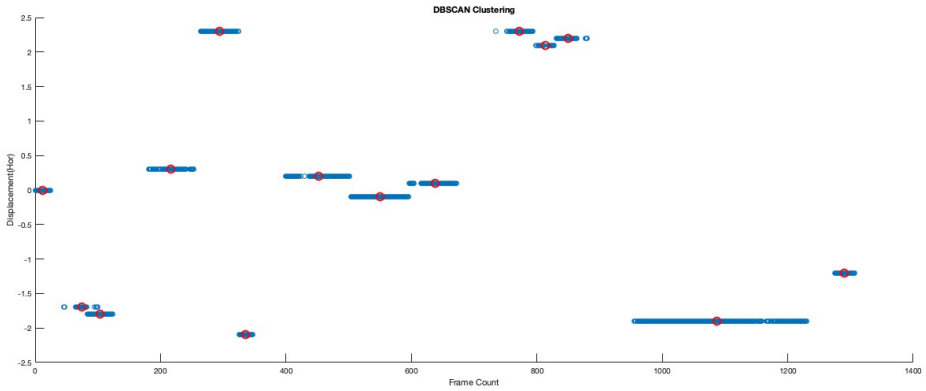


Figure 4.5: Horizontal displacements belonging to the clusters

This allows the elimination of some of the noise caused by large discrepancies between values in nearby frames. After that a moving average filter with a window of 5 is applied, in order to smooth out the remaining values. In Fig. 4.6 we can see the result.

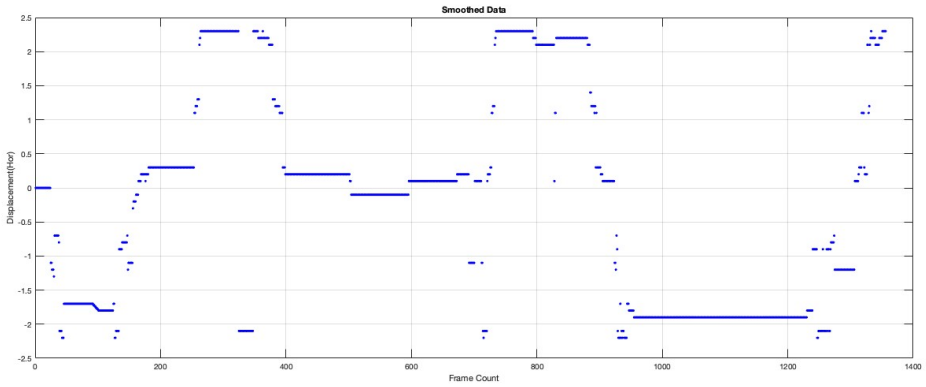


Figure 4.6: Horizontal displacements, after process 4, as function of frame count

Once the horizontal displacements have been identified, the process is repeated for the vertical displacement values.

5

Improvements To The Eye Tracker

This chapter describes the modifications made to the eye tracker system, considering the objectives specified in the introduction.

5.1 Automatic Calibration

As previously described in section 3.2, to initialize the eye tracking software, the user must manually select the eyes and reference points. To automate this calibration process, a new approach was applied, which is described in this section.

When the video is selected and therefore read, using the CascadeObjectDetector ('EyePairBig') function, which belongs to Matlab's Computer Vision toolbox, it returns possible locations of objects in the image that resemble the shape of an eye. After the initial detection, the regions that correspond to the eyes are stored and adjusted to better match their anatomy. This adjustment has been constructed to include defining a bounding box for the right eye based on the position and size of the detected left eye, ensuring that both regions are processed consistently. The Fig. 5.1a shows an example of the area of the image cut out by the bounding box for the right eye and the Fig. 5.1b for the left eye. If the left eye cannot be identified in the first frame, the following frames are analyzed consecutively until it can be identified in one of them.

Although there are already possible locations for the eyes, the direction of the gaze still needs to be found. Therefore, in the next steps, processes that help us highlight the structures that give us indications of this direction were followed.

(a) (b)

Figure 5.1: Right Eye (a). Left Eye (b).

5.2 Image Processing

Within each previously generated box, image processing is performed, consisting of channel separation, calculating the distance between color channel values for iris contrast enhancement, and applying a Gaussian filter.

5.2.1 Channel Separation

The first stage of image processing consists of preparing the color channels for more detailed analysis. In this case, the image from each eye (left and right) is separated into three channels: red, green, and blue, as seen in Fig. 5.2

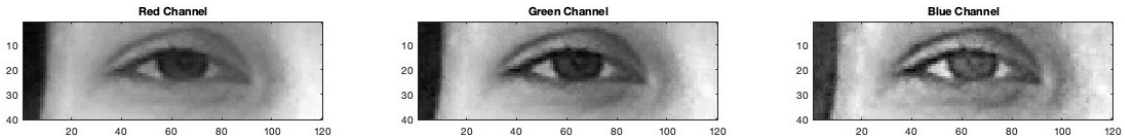


Figure 5.2: Right Eye RGB Channels: Red, Green, and Blue Components

5.2.2 Calculating Distances and Iris Enhancement

After channel separation, a "distance" measure is calculated between the values of the three color channels (red, green, and blue) in each pixel of the image. This "distance" reflects the color differences between channels, which are then used to highlight the iris.

The sum of the squared differences between the channels is calculated using the following formula:

$$\text{distance} = \sqrt{(R-G)^2 + (G-B)^2 + (B-R)^2} \tag{5.1}$$

Here, R, G, and B represent the pixel intensities of the red, green, and blue channels, respectively. The calculated distance is then normalized by dividing by the maximum possible distance, resulting in a relative measure that varies between 0 and 1. The maximum value is calculated considering that two channels have an intensity value of 0 and the other 255, giving a total of 130 050 when the formula 5.1 is applied. This normalization process is described by:

$$\text{normalized\_distance} = \frac{\text{distance}}{\sqrt{130\,050}} \tag{5.2}$$

The normalization process ensures that the distance measure is scaled appropriately. This measure is then inverted, causing areas of greater color difference (such as the iris) to be highlighted. It was observed that multiplying the image resulting from this last process by the image extracted from the blue channel, the contrast is even more pronounced,

emphasizing the iris in relation to the surrounding areas of the eye. It is possible to compare these differences in Fig. 5.3.

This highlighting is crucial to facilitate iris detection in subsequent steps.

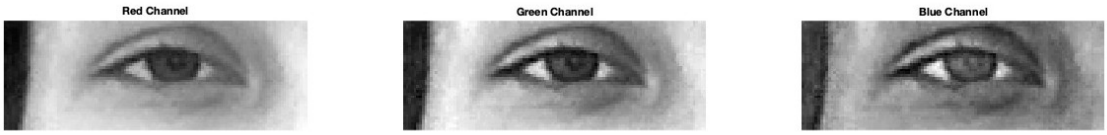


Figure 5.3: Comparison of the multiplication of the normalized channel distance by each individual color channel

5.2.3 Applying a Gaussian Filter

The Gaussian filter is widely used in image processing for smoothing and noise reduction. Its effect is based on the application of a Gaussian function, which results in a matrix (or kernel) with values that progressively decrease as they move away from the center.

For the kernel size, a value of 3 was chosen, which means that the kernel matrix is 3×3 pixels. The standard deviation σ was set to 3. These values were selected taking into account other trials with different values. The kernel size of 3×3 is relatively small, which is suitable for fine details and small areas of the image.

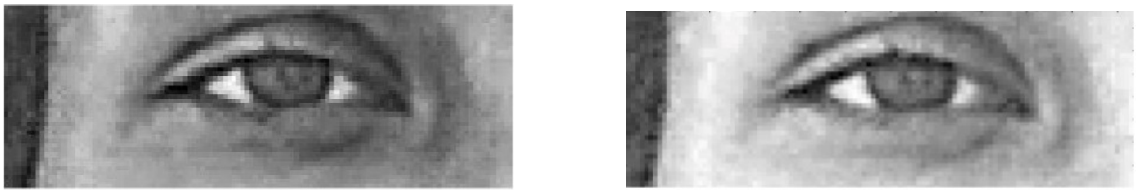
Once the Gaussian kernel has been generated, it is applied to the image using the convolution operation. The `conv2()` function performs Two-Dimensional (2D) convolution between the image and the kernel, resulting in a smoothed version of the image. The 'same' option in the convolution parameter ensures that the output of the operation is the same size as the original image, by centering the kernel on each pixel and computing the weighted average of the neighboring pixels based on the kernel values.

The Gaussian kernel $g(x, y)$ is given by:

$$g(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{5.3}$$

where σ is the standard deviation that determines the width of the Gaussian function.

It is possible to compare the image before and after the application of the Gaussian Filter, in Fig. 5.4.



(a) (b)

Figure 5.4: Before (a) and After (b) the application of the Gaussian Filter.

5.3 Eye Detection

The pupil and iris are noticeably darker than many of the other eye structures and both are roughly circular. In addition to this, they are surrounded by the sclera, which is markedly more intense than the iris and pupil. The direction of gaze is found through the complementary appearance of these structures. With this in mind, a model was built that, throughout the frames, it would allow compare the appearance of these structures, understanding of how they change throughout eye movements and, essentially, how, together, they can help detect them.

5.3.1 Eye Model

Keeping in mind the processes of the previous chapter, now, using the imfindcircles() function, which has been programmed to find circles between a given range of radius values, we obtain the centers of the two irises.

For horizontal movements, the coordinates of the center of the iris, as well as its radius, the region including the sclera and iris is cut, with a new bounding box. The length is calculated by extend the iris length left and right by its radius. Similarly, the height of the bounding box is determined by adding two different fractions of the radius to the y coordinate of the center: 1/6 and 1/2, thus defining the lower and upper limits of the box. If the calculated values fall outside the original bounding box, they are adjusted to ensure that the cut-out region belongs to it. It is possible to see an example in Fig. 5.5.

Figure 5.5: Eye Model for the Horizontal Component

For vertical movements, the length is calculated by subtracting and adding 1/6 of the radius value, defining the maximum an minimum value of the box. The height is defined between the value of the y component of the iris center and the addition of two radius below. It is possible to see an example in Fig. 5.6.

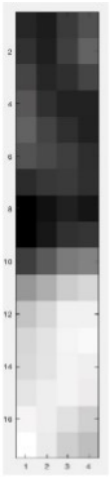


Figure 5.6: Eye Model for the Vertical Component

5.3.2 Eye Movement

In the context of this study, correlation plays a vital role in comparing the structural patterns of the eye in different frames. Therefore in section 5.3.2.1 we have an explanation of how correlation works and why it is used. Subsequently, the horizontal and vertical displacements are calculated as described in sections 5.3.2.2 and 5.3.2.3, considering that the cropped region of the first valid frame, illustrated in Fig. 5.5, is considered to be the model of the eye or "equilibrium position".

5.3.2.1 Correlation

The correlation between two variables is a measure that indicates the relationship between them. Pearson's correlation, one of the most common ways of measuring this relationship, is especially useful when there is a need to understand how two variables envolve together. [25]

Pearson's correlation coefficient (r_{xy}) is calculated using the following expression [26]:

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \tag{5.4}$$

This coefficient ranges from -1 to 1:

- $r = 1$ indicates a perfect positive correlation, where the variables envolve in a similar manner;
- $r = -1$ indicates a perfect negative correlation, where one variable increases as the other decreases;
- $r = 0$ indicates no significant linear correlation between the variables, i.e., both variables envolve in a completely non-related manner.

When working with two-dimensional matrices (2D arrays), in spatial datasets such as images, correlation measures how closely the values in one matrix correspond to the values in the other matrix at the same positions. The goal is to determine whether the patterns of intensities (or other variables) in one matrix are similar to those in the other. [26] The correlation between two matrices A and B can be expressed by the correlation coefficient r, using the following formula: [27]

$$r = \frac{\sum_m \sum_n \left(A_{mn} - \bar{A}\right)\left(B_{mn} - \bar{B}\right)}{\sqrt{\left(\sum_m \sum_n \left(A_{mn} - \bar{A}\right)^2\right)\left(\sum_m \sum_n \left(B_{mn} - \bar{B}\right)^2\right)}} \tag{5.5}$$

Where:

- A_{mn} and B_{mn}: Represent the elements of matrices A and B at the m-th row and n-th column, respectively.
- $\bar{A}$ and $\bar{B}$: These are the mean values of the elements in matrices A and B, calculated as the average over all elements in each matrix.
- $\sum_m \sum_n$: Indicates a double summation, meaning the sum over all rows m and columns n of the matrix.
- $(A_{mn} - \bar{A})$ and $(B_{mn} - \bar{B})$: Represent the deviations of each matrix element from their respective means.
- r: The resulting correlation coefficient, which quantifies the strength and direction of the linear relationship between the two matrices.

5.3.2.2 Horizontal Displacements

Horizontal displacements are defined with a step of 0.1 pixel, starting at a value of 0, i.e., when the regions are overlapping. This step size, representing 10% of the pixel intensity range, allows for the detection of subtle variations in color intensity as the region of the current frame is shifted horizontally by each of the set values. If the current region is not exactly the same size as the original, it is resized.

The correlation between the model matrix and the current matrix is calculated. A graph is obtained with the displacement values on the x-axis and the correlation on the y-axis.

The findpeaks() function identifies the displacements at which the correlation reaches its maximum.

When analyzing a second valid frame, presented with multiple peaks, the one with the smallest derivative is choosen, suggesting that there is a less abrupt difference in the correlation value and the presence of the iris at that location.

For the analysis of the remaining frames, given that the movement of the iris from one frame to the next is considered to be quite small, or often zero, we first check whether there

is a maximum in the displacement previously considered. If there is, it is compared with the peaks immediately before and after. Similarly, the one with the smallest derivative is chosen. If there is not one, simply the one with the least abrupt variation is considered.

5.3.2.3 Vertical Displacements

After finding the horizontal displacement, it is necessary to look for the vertical displacement. As the amplitude of the vertical movement is smaller than that of the horizontal movement, a step of 0.05 was considered in an attempt to accurately detect slight changes, starting at displacement 0, i.e. at the overlap of the regions.

The correlation is calculated in the same way, as is the previous displacement but now moving up and down the equilibrium frame by the current frame. While in horizontal displacement the cropped region of interest is always in the same position, in the case of vertical displacement, for each frame, this region is cropped by centering the box on the considered horizontal displacement.

At each frame, the estimated horizontal and vertical displacement values are stored so that they can be processed and analyzed at the end of the frames.

6 Results

This chapter aims to present the results obtained according to the methodology described in the Chapter 5. Section 6.1 shows the results obtained using the computer's camera, while Section 6.2 shows the results obtained using the smartphone's camera. Each of these is further divided into two subsections, depending on the eye in study.

Since horizontal and vertical movements are stored in distinct matrices, each direction was analyzed separately. Using Microsoft Excel, it was possible to obtain graphs, with the frameCount variable on the x-axis and the displacement on the y-axis.

In the graphs for horizontal movement, the displacement value will be more negative the further to the left of the equilibrium position it is and more positive the further to the right it is.

On the other hand, in the case of vertical movement graphs, the direction of gaze will be further downwards the more positive the displacement value is. This is because the equilibrium position, referring to area of interest A, has the minimum value of the coordinate on the y-axis.

6.1 Data Acquired With The Computer's Camera

6.1.1 Left Eye

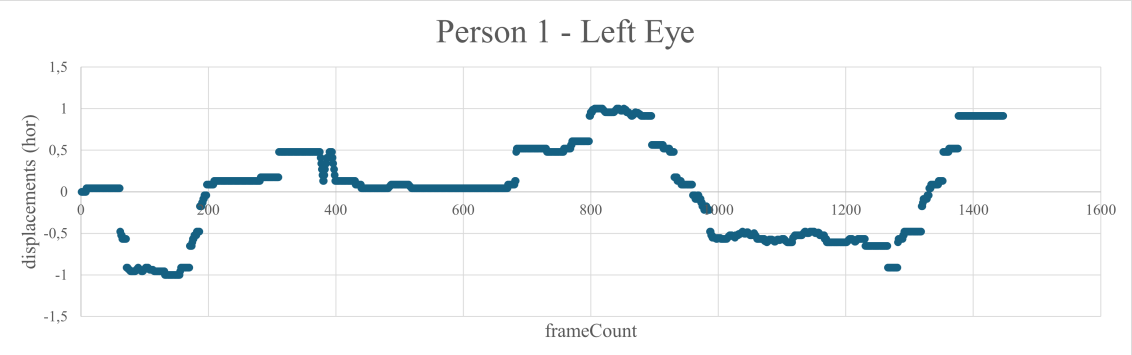


Figure 6.1: Graph of the horizontal position of person 1's left eye throughout the frames captured by the computer camera.

Fig. 6.1 shows the horizontal movement of person 1's left eye throughout the frames. It is possible to identify regions that suggest changes, and consecutive stabilizations, in the direction of the left eye's gaze.

Considering that when there is a displacement value of 0 in relation to the first frame, the eye is in its equilibrium position, it is possible to identify approximately three moments when this occurs: between frames 11 and 51, 210 and 280, 440 and 670.

Between frames 75 and 170, given the negative component of the graph, we can consider that the eye has moved to the left, as well as between frame 1000 and 1300. On the other hand, between frame 680 and 895 there is a visible shift to the right, even with some oscillations.

Figure 6.2: Graph of the vertical position of person 1's left eye throughout the frames captured by the computer camera.

Looking at Fig. 6.2, which shows the vertical displacement, it is possible to see 4 levels: 0, 0,18, 0,35, and 1. However, the values 0 and 1 stand out significantly more than all the others. This suggests greater sensitivity to detecting the equilibrium position, where the displacement is zero, and the furthest vertical position, where the displacement is 1. It is also of noted the abrupt changes between transitions of different values.

6.1.2 Right Eye

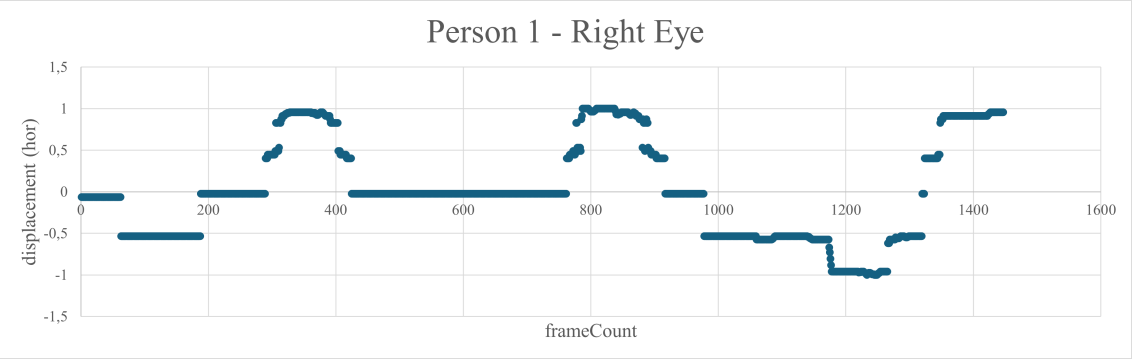


Figure 6.3: Graph of the horizontal position of person 1's right eye throughout the frames captured by the computer camera.

Fig.6.3 shows the horizontal movement of person 1's right eye throughout the frames. Although there are some differences in the values, when comparing with the left eye, it is possible to identify shifts to the right (between frames 300 and 400, 770 and 900, 1350 and 1400) to the left (between frames 70 and 200, 1000 and 1300) and eventual returns to the equilibrium position (between frames 200 and 300, 400 and 770).

Figure 6.4: Graph of the vertical position of person 1's right eye throughout the frames captured by the computer camera.

In terms of vertical movement, illustrated in Fig. 6.4, when compared to the left eye we can perceive values with considerably more noise, but with less accentuated transitions between different values. By looking at the graph, we can identify four regions that are quite clear in their difference from the equilibrium position, where we can predict the existence of a positive, vertical and accentuated movement. It can also be seen that these regions similarly show an increase in values at the beginning, followed by some stabilization and, finally, a decrease. This suggests that once there has been a shift towards the bottom of the screen, it is followed by an opposite vertical shift.

6.2 Data Acquired With The Smartphone's Camera

6.2.1 Left Eye

Figure 6.5: Graph of the horizontal position of person 1's left eye throughout the frames captured by the smartphone camera.

Fig. 6.5 shows the predicted horizontal displacement of the left eye. Compared to Fig. 6.1, we see values with similar upward and downward tendencies. Between frames 790 and 900 there is a marked discrepancy in the values. Still, the rest of the graph shows relatively smooth transitions and is even intuitive to interpret the direction of the calculated eye movement. Thus, areas of leftward displacement and consecutive return to equilibrium are identified between frames 30 and 200 and 970 and 1300, rightward displacement and consecutive return to equilibrium between frames 250 and 400 and 750 and 970, and stabilization between frames 200 and 250 and 400 and 750.

Figure 6.6: Graph of the vertical position of person 1's left eye throughout the frames captured by the smartphone camera.

For the vertical movement, we now see, in Fig. 6.6 a rather large discrepancy with the previous graphs. It suggests that, up to frame 750, the eye is directed towards the bottom of the screen, where the predicted displacement takes 1 as its value. Then, after an abrupt return to a higher vertical position, in frame 750 to frame 900, there appear to be two consecutive, but non-linear, displacements until the maximum displacement value is reached again, which is equivalent to the minimum vertical position of the screen. Subsequently, until around frame 1350, there is a displacement towards the central vertical value of the screen, followed by a return to the lower part.

6.2.2 Right Eye

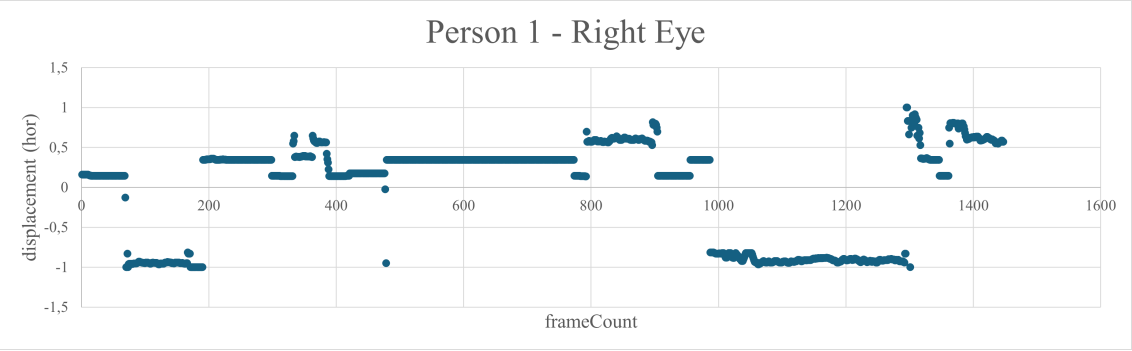


Figure 6.7: Graph of the horizontal position of person 1's right eye throughout the frames captured by the smartphone camera.

Fig. 6.7 presents the graph of the predicted horizontal displacement of the right eye. It can be observed, firstly, that although there is no stabilization at values with zero displacement, which would indicate equilibrium points in the horizontal position, we can consider that these are represented by the data sets with the shortest distance to the x-axis on the graph. A shift to the left can be seen between frames 70 and 200, and between frames 1000 and 1300. On the other hand, a shift to the right can be observed between frames 200 and 1000, except for the areas where it stabilizes again near the x-axis.

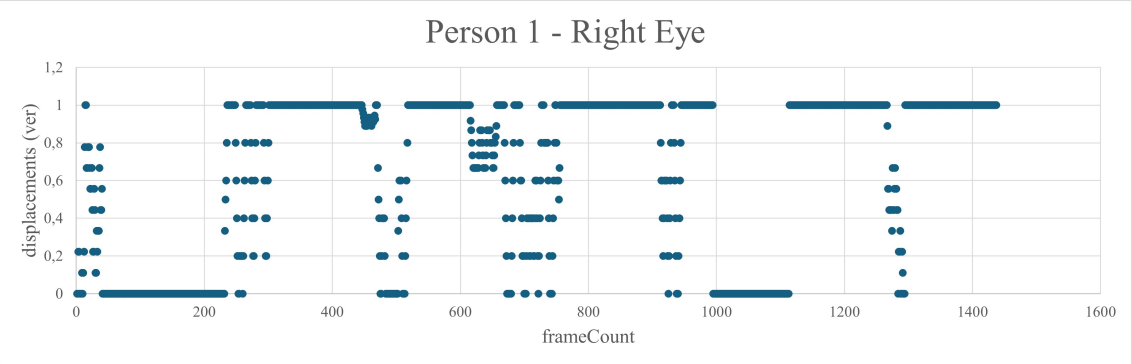


Figure 6.8: Graph of the vertical position of person 1's right eye throughout the frames captured by the smartphone camera.

On the other hand, in Fig. 6.8, which shows the vertical displacement of the right eye, we observe the presence of 'plateaus' whenever there are transitions between zero displacement and maximum displacement. This suggests that, compared to Fig. 6.4, there is greater sensitivity to intermediate values. In this context, we can identify 5 displacements towards the maximum vertical value, and corresponding opposite movements, between frames 200 and 500, 500 and 700, 700 and 900, 900 and 1100, and finally, between frames 1100 and 1300.

7

Discussion of Results

7.1 Metrics Extracted by Comparison with the Stimulus

Since the recorded video was cut manually, there is a difference when compared to the stimulus itself, so in order to synchronize the two, minimizing this difference, it was necessary to resize the matrices corresponding to their displacements, so only the time spent in the sequence of areas B, C, D, E, A, C, I, F and H was considered relevant.

7.1.1 Stimulus Processing

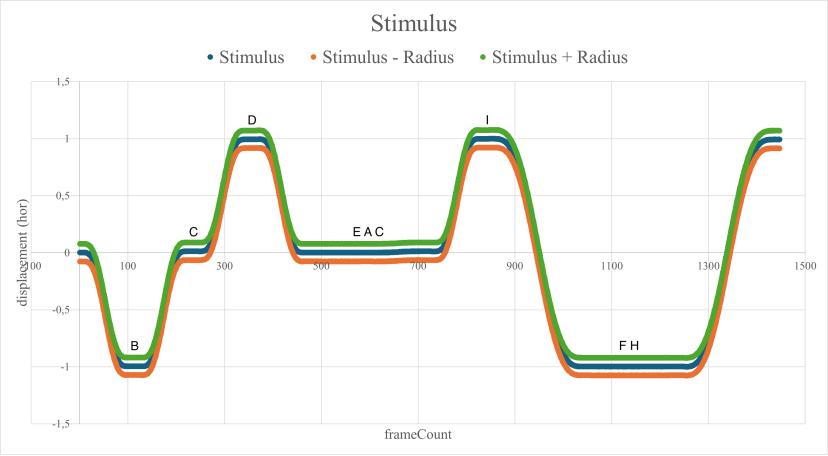


Figure 7.1: Graph of the horizontal positions of the normalized stimulus, as well as the tolerance values obtained from its radius.

In order to be able to compare the results of the eye tracker with the stimulus, the stimulus data was first normalized. Next, using the Matlab function imfindcircles(), the x and y coordinates of the stimulus along the frames were extracted, as well as its radius. In this way, it was possible to delimit a positive tolerance and a negative one associated with the stimulus area. Fig. 7.1 shows the normalized horizontal values of the stimulus, along with its upper and lower limits corresponding to its radius, while Fig. 7.2 shows the normalized vertical values with their respective limits.

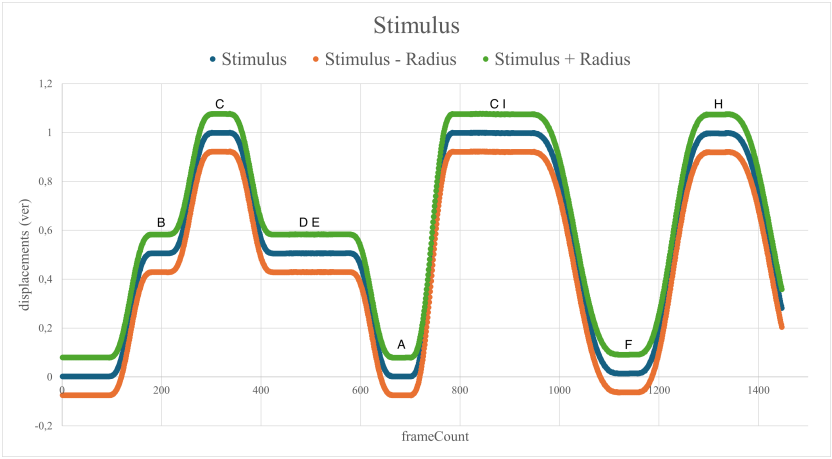


Figure 7.2: Graph of the vertical positions of the normalized stimulus, as well as the tolerance values obtained from its radius.

7.1.2 Frames in Area of Interest

Knowing the number of frames that correspond to each area of interest in the stimulus, in the next sections are presented the expected value for each sequential horizontal and vertical position, as well as the corresponding values calculated by the eye tracker, for each person. Frames are only counted when they are within the tolerance intervals shown in Fig. 7.1 for horizontal displacement and when they are within the intervals shown in Fig. 7.2 for vertical displacement.

7.1.2.1 Horizontal Displacement

Table 7.1: Comparison of frames acquired by the computer for the left and right eyes, included in the regions of interest

Area of Interest	**Left Eye**						**Right Eye**					
	B	**C**	**D**	**E-A-C**	**I**	**F-H**	**B**	**C**	**D**	**E-A-C**	**I**	**F-H**
Expected Frames	**30**	**30**	**30**	**270**	**30**	**210**	**30**	**30**	**30**	**270**	**30**	**210**
Person 1	26	0	0	180	30	0	0	30	30	270	30	71
Person 2	0	11	30	192	0	46	0	0	27	0	0	21
Person 3	0	0	0	152	0	0	0	0	16	248	5	210
Person 4	0	0	0	253	0	0	0	22	30	0	0	209

Analyzing the results of the Table 7.1 for each region individually, it can be seen that, as expected, for regions with fewer frames, correct detection is less likely to occur. The left eye shows better results in the longest-lasting region, i.e., in the E-A-C set, where the horizontal position of the eye is predicted to be the same as the initial one. On the other hand, the right eye generally has better frame detection in region D together with F-H.

Evaluating the performance of the eye tracker in relation to different individuals, person 1 is the one who shows horizontal results closest to those expected, albeit with considerable flaws.

Table 7.2: Comparison of frames acquired by the smartphone for the left and right eyes, included in the regions of interest

Area of Interest	Left Eye						Right Eye					
	B	**C**	**D**	**E-A-C**	**I**	**F-H**	**B**	**C**	**D**	**E-A-C**	**I**	**F-H**
Expected Frames	**30**	**30**	**30**	**270**	**30**	**210**	**30**	**30**	**30**	**270**	**30**	**210**
Person 1	0	0	0	0	0	190	28	0	0	0	0	7
Person 2	0	0	21	72	12	112	0	0	0	0	0	200
Person 3	0	1	0	11	0	210	0	20	0	240	0	10
Person 4	0	0	0	0	8	0	0	10	25	20	0	40

On the other hand, Table 7.2 includes the frames obtained by the smartphone. The values associated with the F-H region for both eyes stand out, as they are noticeably higher than the others. In this region, as in B, the direction of gaze is shifted to the left edge of the screen, but the detection performance is considerably different. It is also possible to see that the values change very significantly from one eye to the other, making it difficult to find a pattern in the detection of the regions.

7.1.2.2 Vertical Displacement

Table 7.3: Comparison of frames acquired by the computer for the left and right eyes, included in the regions of interest

Area of Interest	Left Eye							Right Eye						
	B	**C**	**D-E**	**A**	**C-I**	**F**	**H**	**B**	**C**	**D-E**	**A**	**C-I**	**F**	**H**
Expected Frames	30	30	150	30	150	30	30	30	30	150	30	150	30	30
Person 1	0	4	0	5	0	30	0	0	0	2	30	0	30	0
Person 2	0	0	5	0	0	0	0	0	0	0	30	0	0	0
Person 3	0	0	0	0	5	0	0	0	0	2	30	1	0	30
Person 4	0	0	0	30	0	21	0	0	0	0	30	80	0	0

Looking now at the vertical results in Table 7.3, regions B and H do not include any frames for the left eye, a situation similar to the right eye results for regions B and C. On the other hand, the detection of the right eye in region A was completely successful. In this region, the eye returns to its starting position, which contributes to better recognition when compared with the first frame of the video.

Table 7.4: Comparison of frames acquired by the smartphone for the left and right eyes, included in regions of interest

Area of Interest	Left Eye							Right Eye						
	B	**C**	**D-E**	**A**	**C-I**	**F**	**H**	**B**	**C**	**D-E**	**A**	**C-I**	**F**	**H**
Expected Frames	30	30	150	30	150	30	30	30	30	150	30	150	30	30
Person 1	0	30	2	0	0	30	30	0	0	30	2	140	10	30
Person 2	0	30	10	0	150	0	30	0	10	0	0	70	20	20
Person 3	0	30	0	0	55	0	16	0	30	0	30	140	30	16
Person 4	0	0	0	30	10	30	30	0	30	0	30	0	0	30

In Table 7.4, with regard to the left eye, in region H, the eye tracker correctly detected almost all of the frames, with the exception of some frames from person 3. In region C, only the tracking of person 4 showed problems, since it correctly detected all of the frames from the other three. For the right eye, the eye tracker again performed well in region H. In regions C, A, and C-I, despite performing very well in 3 of the 4 individuals, there was always incorrect detection in the other.

7.1.3 Distance to the Stimulus

As seen in the previous section, the detection made by the eye tracker is not always precise. However, by visually analyzing the displacement graphs in their entirety, it is possible to find detection regions with characteristics similar to those of the stimulus, even if they aren't included in the zones of detection.

An example in Fig. 7.3 that includes the predicted horizontal displacement graph of the left eye of person 1 compared to the total displacement of the stimulus represented with its tolerance intervals. It is possible to note that, approximately, between frames 300 and 400, the values calculated by the eye tracker follow a pattern similar to that of the stimulus. The value of the upper limit suggests a failure in following the movement to the right until its end. Similarly, between frame 1000 and frame 1300, there is a break in the trend to the left, resulting in a level that is higher than expected.

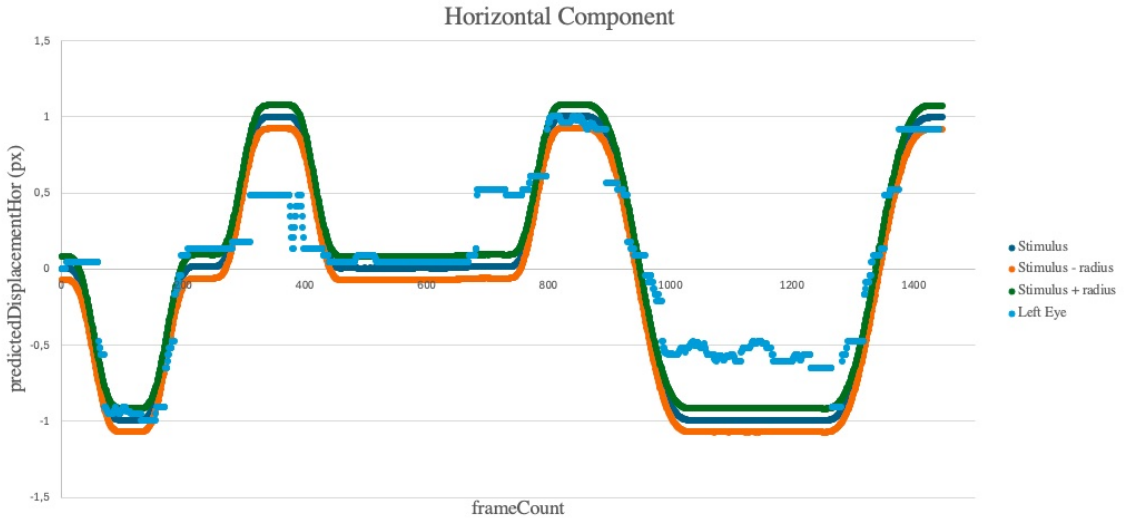


Figure 7.3: Horizontal component of the stimulus, as well as its tolerance values, along with the result of the detection from the left eye of individual 1, through the computer camera.

Regarding vertical displacement, in the Fig. 7.4 it is an example where, between frames 200 and 400 and between frames 750 and 900, there is an increasing trend in the values obtained through the eye tracker, even though these are not included in the valid values counted. The same is observed near frame 1200 and 1400, where there is an increase and

a decrease in values, respectively, following the trend of the stimulus in their respective areas.

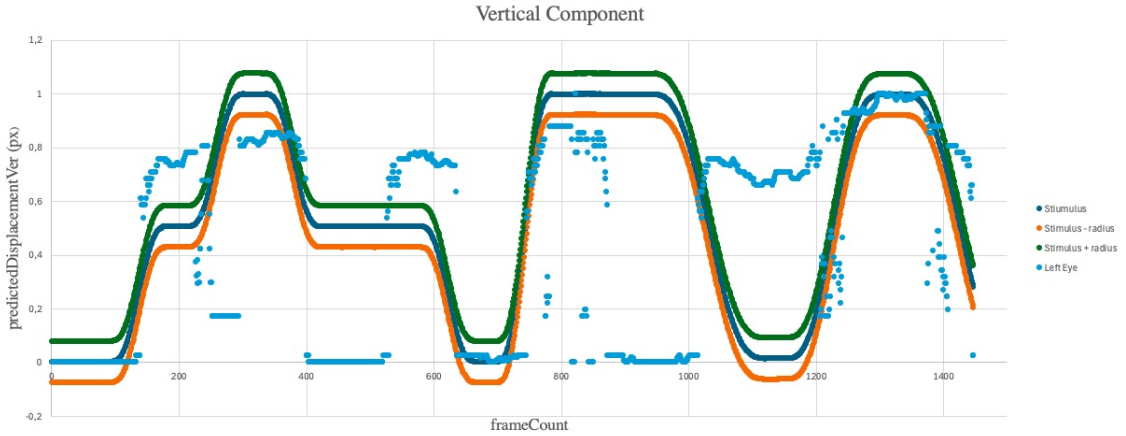


Figure 7.4: Vertical component of the stimulus, as well as its tolerance values, along with the result of the detection from the left eye of person 1, through the computer camera.

In order to quantify the differences observed in the graphs, the average distances of the horizontal and vertical components were calculated for each participant, as well as the average Euclidean Distance between the two signals. For this purpose, Microsoft Excel was used, employing equations 7.1, 7.2 and 7.3, where S_h, S_v, E_h, and E_v correspond to the horizontal and vertical components, in pixels, of the stimulus and eyes, respectively, while n corresponds to the number of frames.

$$\text{Average Horizontal Distance} = \frac{\sum |S_h - E_h|}{n} \tag{7.1}$$

$$\text{Average Vertical Distance} = \frac{\sum |S_v - E_v|}{n} \tag{7.2}$$

$$\text{Average Euclidean Distance} = \frac{1}{n} \sum_{i=1}^{n} \sqrt{(S_h - E_h)^2 + (S_v - E_v)^2} \tag{7.3}$$

Person	**Left Eye**			**Right Eye**		
	Horizontal	**Vertical**	**Euclidean**	**Horizontal**	**Vertical**	**Euclidean**
1	127	332	634	102	275	326
2	184	43	201	174	241	323
3	361	182	445	103	193	253
4	277	141	356	277	216	404

Table 7.5: Average Distances between the stimulus and the eyes recorded by the computer

Table 7.5 shows the values obtained for the data collected by the computer camera, transposed to the actual size of the stimulus, in pixels, knowing that the total size of the screen is 1440x900 pixels.

For the left eye, the values for horizontal displacement are between 127 and 361 pixels, with an average of 234 pixels and a standard deviation of 89 pixels. For vertical displacement, the values range from 43 to 332 pixels, with an average of 175 pixels and a standard deviation of 104 pixels. For Euclidean distance, the values range from 201 to 634 pixels, with an average of 409 pixels and a standard deviation of 157 pixels. This last error, compared to the screen resolution, is equivalent to 16% for the horizontal component and 19% for the vertical. On the other hand, the values for the right eye, for horizontal displacement ranged from 102 to 277 pixels, with an average of 164 pixels and a standard deviation of 71 pixels. For vertical displacement, the values vary between 193 and 275 pixels, with an average of 231 pixels and a standard deviation of 30 pixels. Finally, for the Euclidean distance, the values are between 253 and 404 pixels, with an average of 326 pixels and a standard deviation of 53. This results in an error equivalent to 11% for the horizontal resolution and 26% for the vertical one.

Person	**Left Eye**			**Right Eye**		
	Horizontal	**Vertical**	**Euclidean**	**Horizontal**	**Vertical**	**Euclidean**
1	169	252	352	174	189	301
2	95	190	250	313	239	449
3	203	388	395	136	259	277
4	325	211	524	201	180	366

Table 7.6: Average Distances, in pixels, between the stimulus and the eyes recorded by smartphone

Table 7.5 shows the values for data acquisition with the smartphone. For the left eye, the values associated with horizontal displacement are between 95 and 325 pixels, with an average of 198 pixels and a standard deviation of 83 pixels. In the case of vertical displacement, the values are between 190 and 388 pixels, with an average of 260 pixels and a standard deviation of 77. Finally, the Euclidean distance values are between 250 and 524 pixels, with an average of 380 pixels and a standard deviation of 98 pixels. By calculating the error, it is obtained an error equivalent to 14% for the horizontal component and 29% for the vertical component. Looking at the values for the right eye, the average horizontal distances range from 136 to 313 pixels, with an average of 204 pixels and a standard deviation of 66 pixels. Vertical distances range from 180 to 259 pixels, with an average of 217 pixels and a standard deviation of 33 pixels. Finally, the Euclidean distance values are between 277 and 449 pixels, with an average of 348 pixels and a standard deviation of 66 pixels. The horizontal component is therefore considered to have an error of 14% and the vertical component 24%.

7.2 Previous vs Current Eye Tracker

As seen in chapter 1, this dissertation was designed to try to make the existing eye tracker more robust, essentially in situations where the iris was not so easily perceived as circular. In this sense, taking the person 2 as a case study, as it was found that when analyzed by the previous eye tracker, there was a great loss of information. Fig. 7.5 shows one of the frames from the video, where it can be empirically seen that, for the right eye, the coordinate values returned are not recognizing the correct structure, and, for the left eye, there is no structure recognized as being similar to the iris.

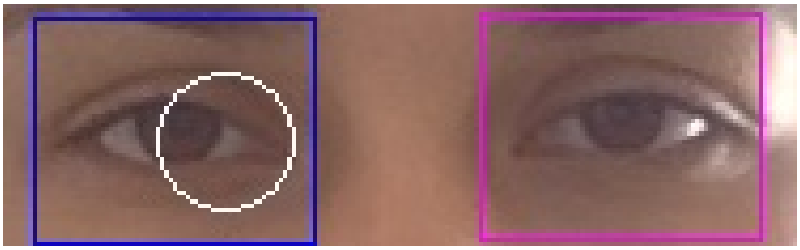


Figure 7.5: Frame taken from person 2's video, acquired by the computer's built-in camera, cropped manually.

As it was necessary more information to draw any conclusions about what was observed, Fig. 7.6 and Fig. 7.7 show the values of the horizontal and vertical displacement, respectively.

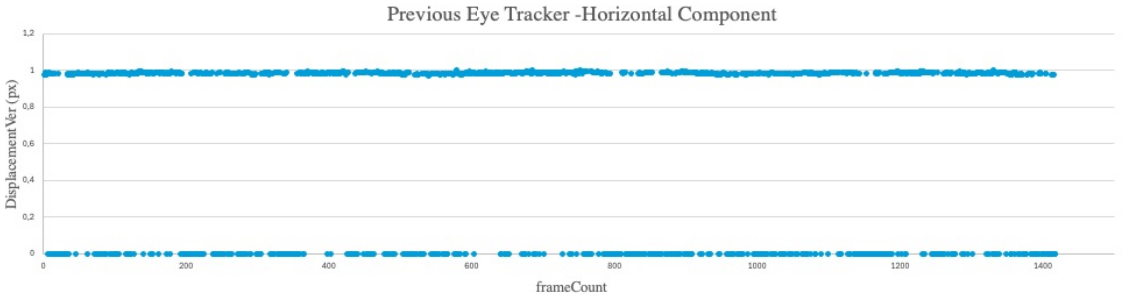


Figure 7.6: Graph of the horizontal position of person 2's left eye throughout the frames captured by the computer camera.

In both graphs, it is possbile that there are two groups of very discrepant values, which suggests that, on the one hand, when there is a possible detection of the iris, it always takes values close to the same position and, on the other hand, that it often takes the value 0, as it does not find any structure that resembles what it is looking for.

The average horizontal distance, average vertical distance and Euclidean distance were also calculated for the results of the two analysis and image acquisition devices.

Table 7.7 shows that, for the previously developed eye tracker, it was obtained an average horizontal distance of 615 pixels, an average vertical distance of 294 pixels, and a Euclidean distance of 710 pixels. These values, when compared to the pixel resolution of the screen (1440x900), suggest an error of 43% associated with the horizontal component

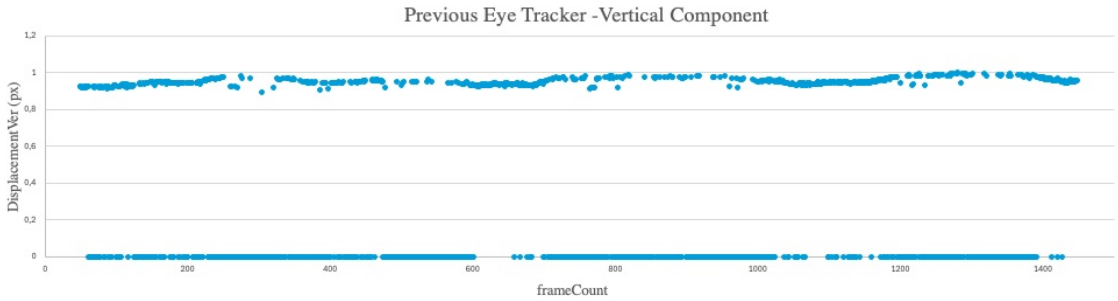


Figure 7.7: Graph of the vertical position of person 2's left eye throughout the frames captured by the computer camera.

and 33% with the vertical component. Compared to the current eye tracker, there is now a horizontal mean distance of 184 pixels, a vertical mean distance of 43 pixels, and a Euclidean mean distance of 201 pixels. These numbers suggest an error of 13% for the horizontal component and 5% for the vertical component.

Person	**Previous Eye Tracker**			**Current Eye Tracker**		
	Horizontal	**Vertical**	**Euclidean**	**Horizontal**	**Vertical**	**Euclidean**
2	615	294	710	184	43	201

Table 7.7: Average distances, in pixels, between the stimulus and the eyes recorded by the smartphone.

Table 7.8 shows the measurements taken by the smartphone's dedicated camera, with an average horizontal distance of 587 pixels, a vertical distance of 310 pixels and a Euclidean distance of 698 pixels, giving an error of 41% for the horizontal component and 34% for the vertical component. In relation to the current data, it was obtained an average horizontal distance of 95 pixels, a vertical distance of 190 pixels and a Euclidean distance of 250 pixels, with an error of 7% for the horizontal component and 21% for the vertical component.

Person	**Previous Eye Tracker**			**Current Eye Tracker**		
	Horizontal	**Vertical**	**Euclidean**	**Horizontal**	**Vertical**	**Euclidean**
2	587	310	698	95	190	250

Table 7.8: Average distances, in pixels, between the stimulus and the eyes recorded by the computer.

Although the focus was only on the results of person 2, it is possible to consider that the error given by the current eye tracker in this case is considerably lower, suggesting that the method used improves the tracking of visual attention.

7.3 Left Eye vs Right Eye

As seen in the previous sections, it is difficult to find a pattern among the results obtained for each eye individually. This is because, for example, in the case of the metric "frames in the region of interest," the left eye generally showed lower results than the right eye, especially in vertical displacement. However, on the other hand, despite its average horizontal distance to the stimulus being greater in the video recorded with the computer, in the other conditions, it exhibited a lower error than the right eye.

Some of the factors that may have contributed to these differences are:

1. The morphology of the eyes, in the same individual, is different for each of them. Possibly, the area cut out for the left eye and the right eye does not allow for eye movements to be detected in the same way, given that the horizontal distance between the center of the iris and the end of the sclera is different for the right and the left.

2. Lighting conditions. During data acquisition, the individuals were in a room with the light source to their left. Given that the same initial image preprocessing was performed for each eye, if they show somewhat discrepant values initially, it is to be expected that the results will behave differently.

7.4 Horizontal Displacement vs Vertical Displacement

Comparing now the results of the horizontal and vertical components, it was consistently found that, both in the number of frames in regions of interest and in the average distances calculated in relation to the stimulus, the horizontal component shows considerably superior results. This may be due to:

1. When the eye moves only horizontally, the change in structures in the area of study of the eye is relatively small, meaning that it is expected that the morphology of the iris and the sclera will change together. On the other hand, when it moves vertically, the change in structures such as the upper eyelid and eyelashes strongly influences detection. Given that, when comparing the frames with the initial one, the more variability in structures are presented over time, the more difficult this tracking becomes.

2. The amplitude of the horizontal movement is greater than that of the vertical movement in response to the stimulus, given the conditions of the monitors. By making movements more spaced out, it becomes easier to perceive the small differences in motion.

7.5 Computer's Camara vs Smartphone's Camara

One of the objectives mentioned from the beginning was to study the differences in the performance of the eye tracker, using two cameras positioned differently in relation to the individual. This is because, especially in vertical movements, when the eyelid covers a large part of the structures of the iris and sclera, there is a better view of them if the camera is positioned below eye level. However, since it is intended to provide image acquisition with the same device that presents the stimulus, the differences in performance between the two were studied.

Taking into account the number of valid frames calculated in the regions of interest, the horizontal displacements behave similarly. However, for vertical movement, the number of frames is considerably higher in the data obtained through the external camera.

In relation to the distances to the stimulus, they had quite similar average results, with the exception of the vertical component of the left eye, where the smartphone camera showed a greater average difference of values, 29%, compared to the 16% obtained by the computer's integrated camera.

8

Conclusion and Future Perspectives

In the course of this work, adjustments were made to the existing eye tracker, focusing on a different approach to recognizing structures, namely using correlation. To do this, a sequence of image processing methods was built to highlight the most important structures while preserving their contours.

The stimulus was applied to 4 individuals without associated pathologies to simulate a control group. Their results were represented with an illustration of the data corresponding to person 1 and an extensive analysis was carried out comparing all the results with the stimulus.

It was observed that, about the “frames in the region of interest”, the left eye generally showed lower results than the right eye, especially in vertical displacement. Also, it was noticed that the horizontal component produces far better results in terms of the number of "frames in regions of interest" and the average distances computed in respect to the stimulus.

Also noteworthy was the case of person 2, who was especially chosen because their eyelids were generally more closed, making detection difficult using the previously developed eye tracker. The reduction in error was quite noticeable, from 41 % in the horizontal direction and 34% in the vertical direction to 7% and 21% respectively.

As expected, the external camera was more sensitive to vertical displacements when counting frames in areas of interest. However, in the overall average distance to the stimulus, there was a greater error for the smartphone camera in the left eye, which suggests that the data acquired by the computer’s integrated camera is close to the expected by the stimulus, although it does not always fall within the range of interest.

It is also important to consider that, despite the attempt to immobilize individuals’ head movements through the chin strap, it does not guarantee complete immobilization, which affects the results.

Another point to consider comes from the fact that, when the stimulus is present and its movement begins, there is the possibility that the brain can, in most cases, predict how far it will move on the screen. Although these are individuals without associated pathologies, it is not possible to guarantee that certain involuntary eye movements will

not alter results that we are attributing to eye tracker errors. Even though there is still a long way to go before an irrefutable level of trust can be added to the eye tracker, it proves to be a low-cost tool that is valuable in complementing the RehabVisual platform, allowing stimulation, data collection, and analysis all in one device.

For future, it would be interesting to add a model of the eye with more structures, namely the eyelids, capable of adjusting throughout the frames, given that they have a significant influence on the results.

Also, the next step is to collect data again in a hospital setting, in order to compare real data from individuals with specific pathologies.

Bibliography

[1] J. M. Lourenço. *The NOVAthesis LaTeX Template User's Manual*. NOVA University Lisbon. 2021. URL: https://github.com/joaomlourenco/novathesis/raw/main/template.pdf (cit. on p. i).

[2] W. H. Organization. *Eye care, vision impairment and blindness*. 2023. URL: https://www.who.int/health-topics/blindness-and-vision-loss (visited on 2024-08-01) (cit. on p. 1).

[3] Y. Fandakova and C. Hartley. "Mechanisms of learning and plasticity in childhood and adolescence". In: *Developmental Cognitive Neuroscience* 42 (2020) (cit. on p. 1).

[4] A. Duchowski. *Eye Tracking Methodology: Theory and Practice*. 2007. ISBN: 978-1-84628-608-7. DOI: 10.1007/978-1-84628-609-4 (cit. on pp. 1, 3, 6–9).

[5] R. B. Machado. *Desenvolvimento de uma aplicação para estimulação das competências visuomotoras*. 2017 (cit. on p. 1).

[6] C. Santos. *Desenvolvimento e validação de uma ferramenta para estimulação das competências visuomotoras em bebés com alterações de desenvolvimento decorrentes do nascimento*. 2018 (cit. on p. 1).

[7] P. Dias. *Actualização e validação da plataforma RehabVisual: Ferramenta para estimulação das competências visuomotoras* (cit. on p. 1).

[8] P. Fonseca. *Validação da plataforma RehabVisual: Ferramenta para estimulação das competências visuomotoras - Aplicação a Doentes com AVC* (cit. on pp. 1, 11, 13).

[9] A. Russo and J. Reagen. *Seeley's Anatomy & Physiology*. Vol. 11. 2016. ISBN: 1264103883 (cit. on pp. 3, 4).

[10] G. Tortora. *Principles of Anatomy and Physiology*. Vol. 14. 2014, pp. 580–595 (cit. on pp. 3, 4).

[11] *Textbook of Medical Physiology*. Vol. 1. 2016. ISBN: 978-1-4557-7005-2 (cit. on p. 5).

[12] E. B. Huey. *Psychology and Pedagogy of Reading*. Vol. 1. 1921 (cit. on p. 6).

[13] A. Schall and J. Bergstrom. *Introduction to the eye tracking*. 2014, pp. 3–26. ISBN: 978-0-12-408138-3. DOI: 10.1016/978-0-12-408138-3 (cit. on p. 6).

[14] A. M. Barreto. "Eye Tracking: como método de investigação aplicado às ciências da comunicação". In: *Revista Comunicando* 1.1 (2012) (cit. on p. 6).

[15] A. Dobhan and D. Röhner. *Eye-Tracking and Usability in (Mobile) ERP Systems*. 2021 (cit. on p. 7).

[16] L. R. Young and D. Sheena. "Survey of Eye Movement Recording Methods". In: *Behavior Research Methods & Instrumentation* 7.5 (1975), pp. 397–429 (cit. on p. 7).

[17] P. Green. *Review of Eye Fixation Methods and Recording Equipment*. Vol. 1. 1992 (cit. on p. 7).

[18] Skalar. *Courtesy of Skalar Medical*. 2017. URL: http://www.skalar.nl (visited on 2024-01-31) (cit. on p. 8).

[19] EyeMove. *EyeMove*. 2023. URL: http://www.eyemove.io (visited on 2024-01-31) (cit. on p. 8).

[20] EyeLink. *Pupil-Corneal Reflection (P-CR) Eye Tracking*. 2023. URL: https://www.sr-research.com/about-eye-tracking/ (visited on 2024-09-20) (cit. on p. 9).

[21] Metrovision. *Eye tracking systems*. 2023. URL: https://metrovision.fr/eyetrackers-us.html (visited on 2024-01-31) (cit. on p. 9).

[22] B. Farnsworth. *10 Free Eye Tracking Software Programs [Pros and Cons]*. 2023. URL: https://imotions.com/blog/insights/trend/free-eye-tracking-software/ (visited on 2024-02-04) (cit. on p. 9).

[23] Tobii. *Tobii Pro Fusion*. 2024. URL: https://www.tobii.com/products/eye-trackers/screen-based/tobii-pro-fusion (visited on 2024-01-25) (cit. on pp. 9, 10).

[24] J. Sander and X. Xu. "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise". In: *Institute for Computer Science, University of Munich* () (cit. on p. 14).

[25] Z. M. A. *Correlation and Regression Analysis*. 2015 (cit. on p. 20).

[26] A. Asuero and A. G. González. "The Correlation Coefficient: An Overview". In: *Critical Reviews in Analytical Chemistry* 36 (2006) (cit. on pp. 20, 21).

[27] MathWorks. *corr2*. 2006. URL: https://www.mathworks.com/help/images/ref/corr2.html (visited on 2024-08-01) (cit. on p. 21).

A

Visualization of Results

This appendix shows the resulting graphs associated with each person.

A.1 Data Acquired With The Computer's Camera

A.1.1 Left Eye

A.1.1.1 Horizontal

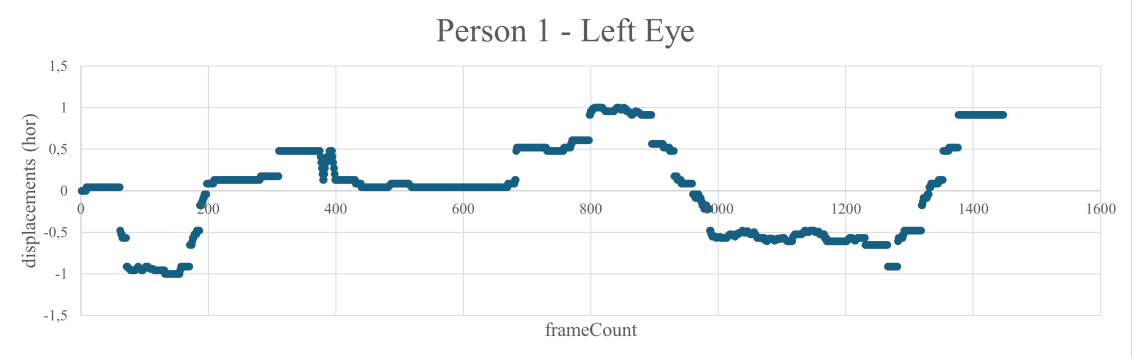


Figure A.1: Graph of the horizontal position of person 1's left eye throughout the frames captured by the computer camera.

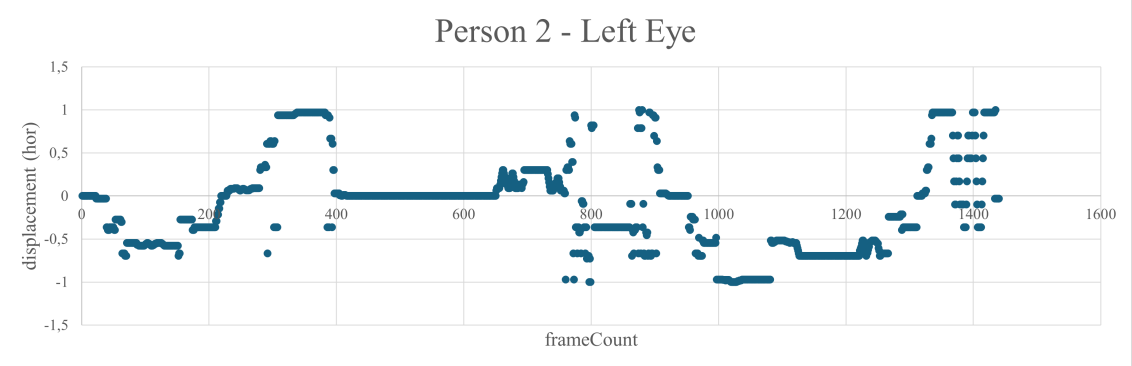


Figure A.2: Graph of the horizontal position of person 2's left eye throughout the frames captured by the computer camera.

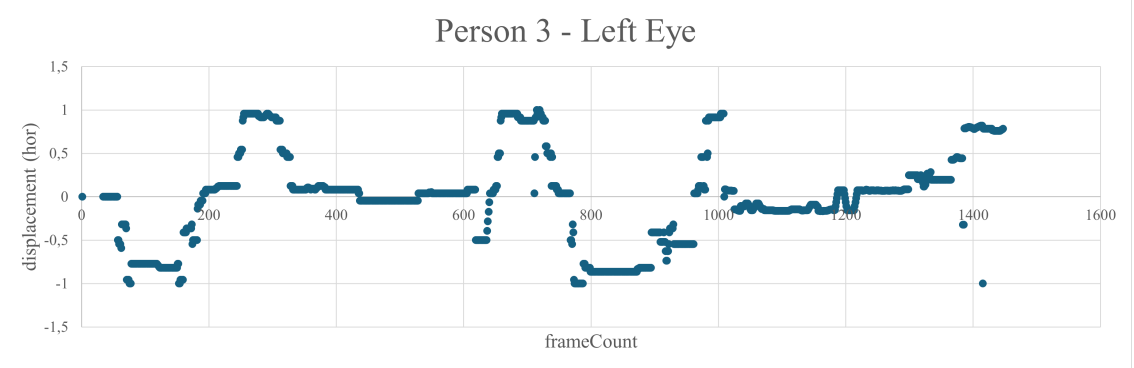


Figure A.3: Graph of the horizontal position of person 3's left eye throughout the frames captured by the computer camera.

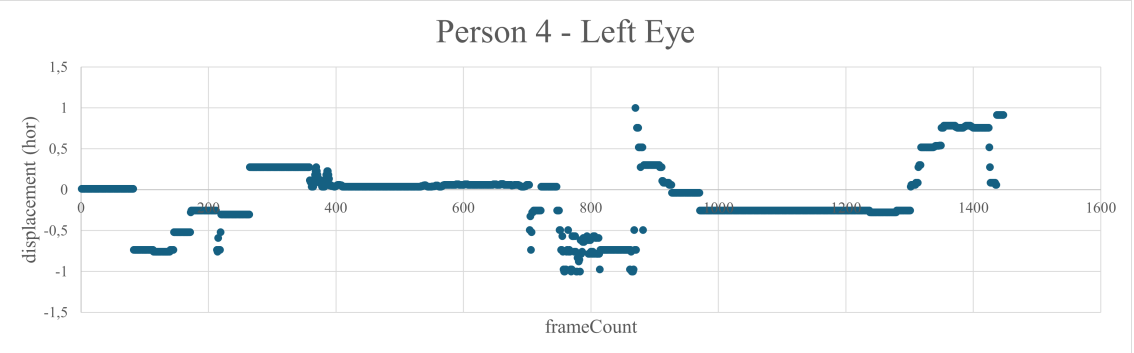


Figure A.4: Graph of the horizontal position of person 4's left eye throughout the frames captured by the computer camera.

A.1.1.2 Vertical

Figure A.5: Graph of the vertical position of person 1's left eye throughout the frames captured by the computer camera.

Figure A.6: Graph of the vertical position of person 2's left eye throughout the frames captured by the computer camera.

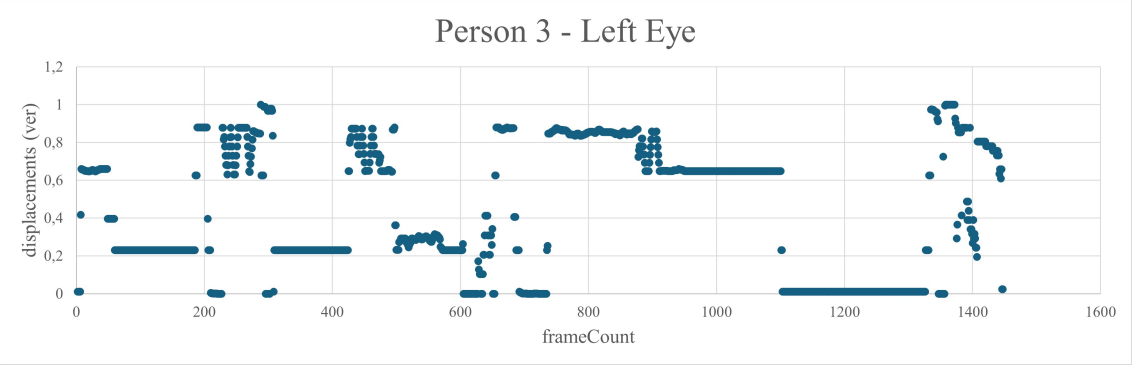


Figure A.7: Graph of the vertical position of person 3's left eye throughout the frames captured by the computer camera.

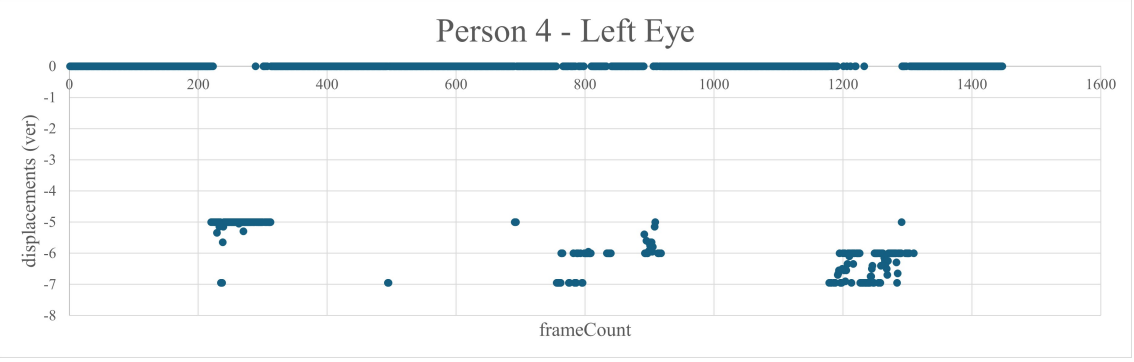


Figure A.8: Graph of the vertical position of person 4's left eye throughout the frames captured by the computer camera.

A.1.2 Right Eye

A.1.2.1 Horizontal

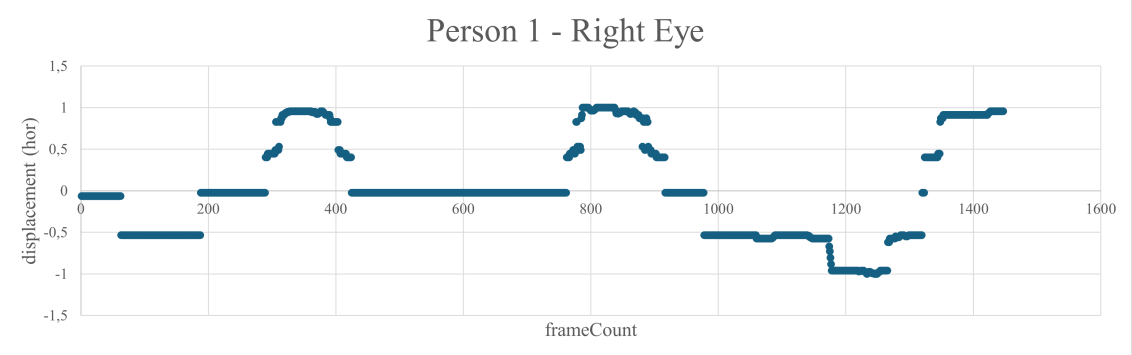


Figure A.9: Graph of the horizontal position of person 1's right eye throughout the frames captured by the computer camera.

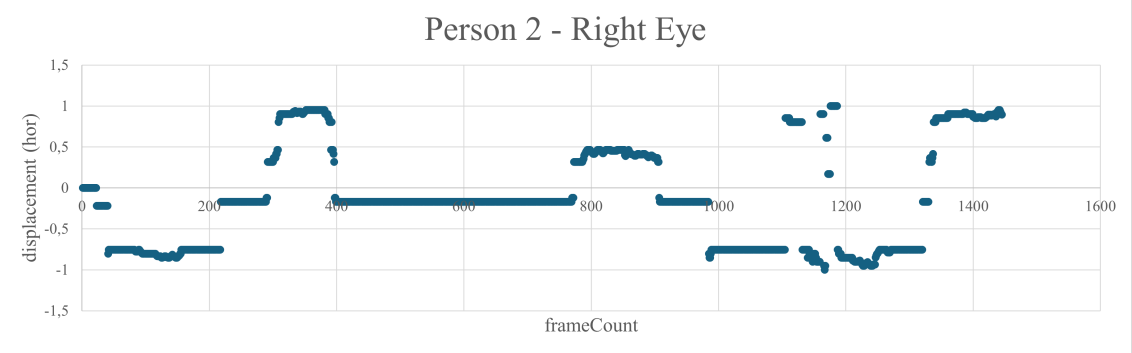


Figure A.10: Graph of the horizontal position of person 2's right eye throughout the frames captured by the computer camera.

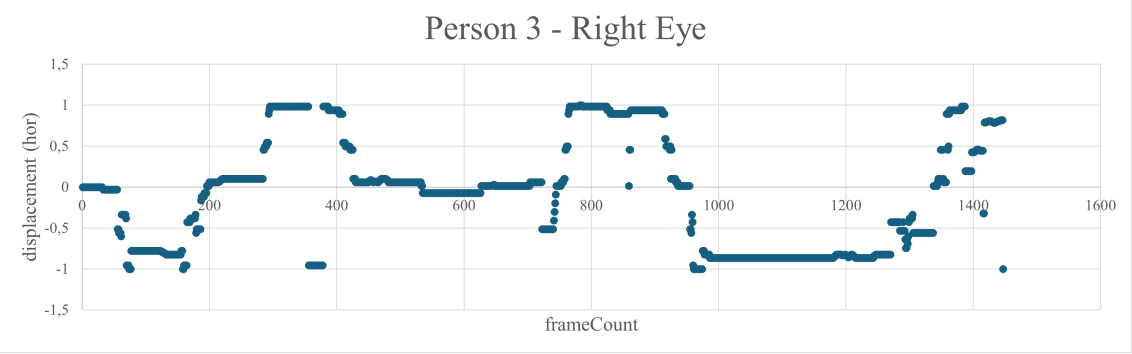


Figure A.11: Graph of the horizontal position of person 3's right eye throughout the frames captured by the computer camera.

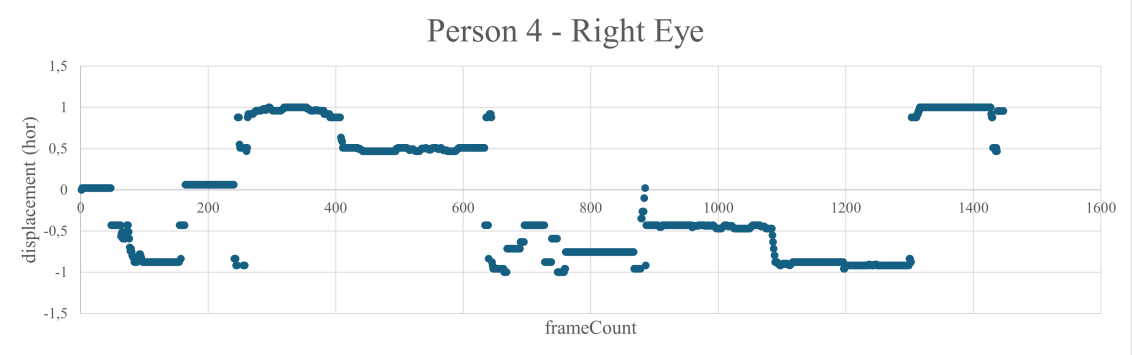


Figure A.12: Graph of the horizontal position of person 4's right eye throughout the frames captured by the computer camera.

A.1.2.2 Vertical

Figure A.13: Graph of the vertical position of person 1's right eye throughout the frames captured by the computer camera.

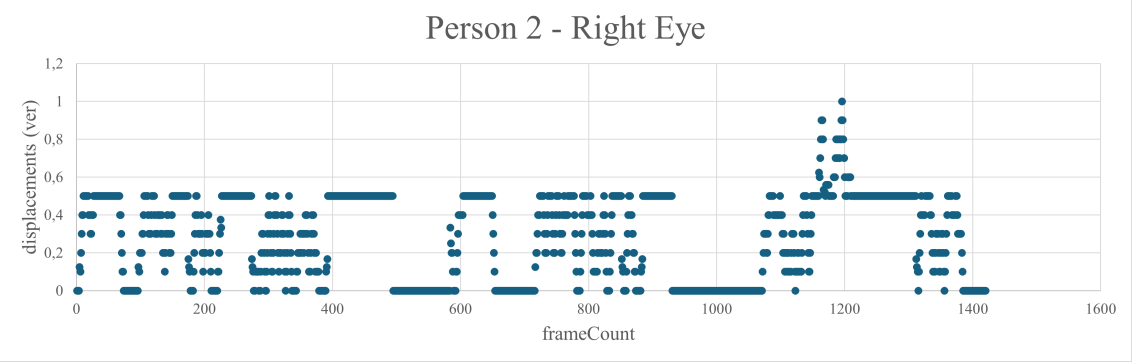


Figure A.14: Graph of the vertical position of person 2's right eye throughout the frames captured by the computer camera.

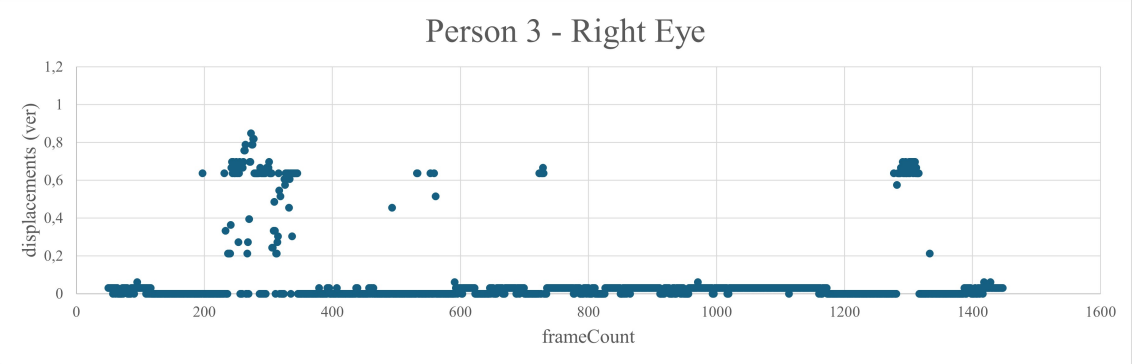


Figure A.15: Graph of the vertical position of person 3's right eye throughout the frames captured by the computer camera.

Figure A.16: Graph of the vertical position of person 4's right eye throughout the frames captured by the computer camera.

A.2 Data Acquired With The Smartphone's Camera

A.2.1 Left Eye

A.2.1.1 Horizontal

Figure A.17: Graph of the horizontal position of person 1's left eye throughout the frames captured by the smartphone camera.

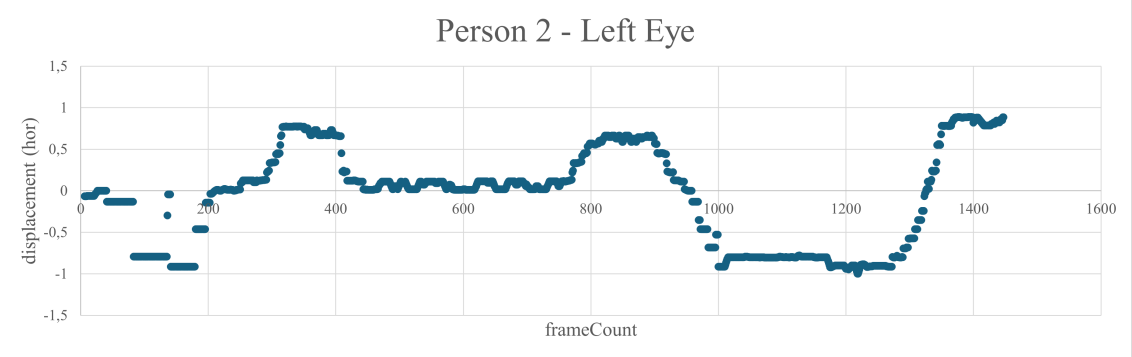


Figure A.18: Graph of the horizontal position of person 2's left eye throughout the frames captured by the smartphone camera.

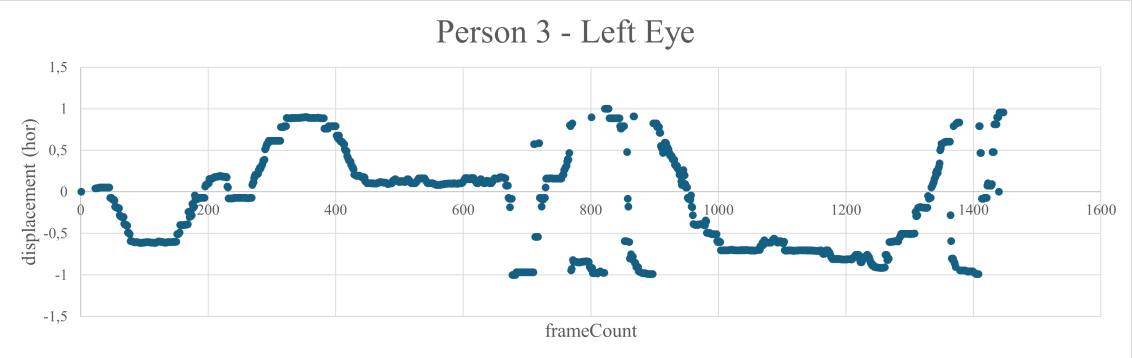


Figure A.19: Graph of the horizontal position of person 3's left eye throughout the frames captured by the smartphone camera.

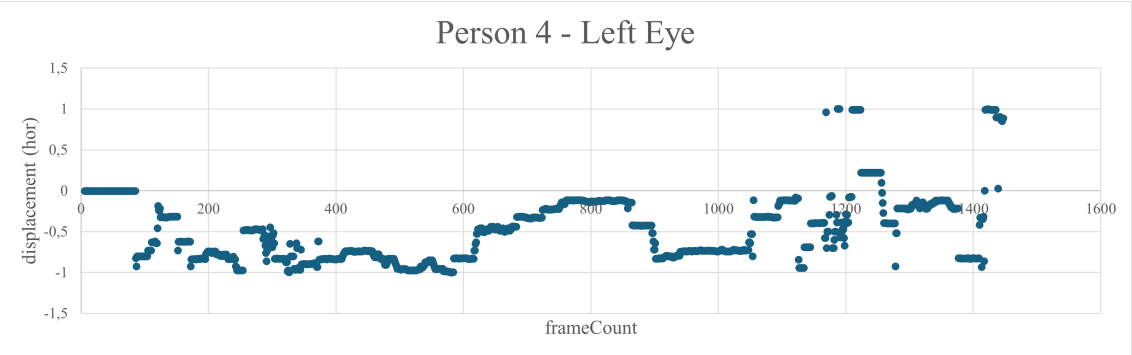


Figure A.20: Graph of the horizontal position of person 4's left eye throughout the frames captured by the smartphone camera.

A.2.1.2 Vertical

Figure A.21: Graph of the vertical position of person 1's left eye throughout the frames captured by the smartphone camera.

Figure A.22: Graph of the vertical position of person 2's left eye throughout the frames captured by the smartphone camera.

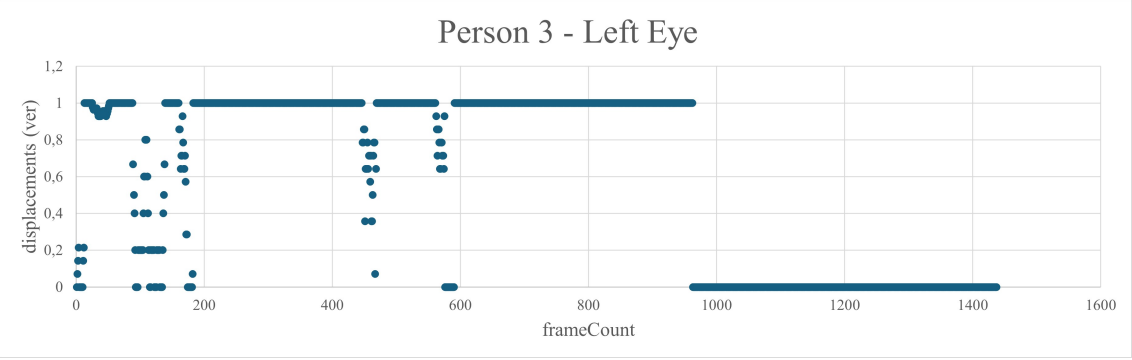


Figure A.23: Graph of the vertical position of person 3's left eye throughout the frames captured by the smartphone camera.

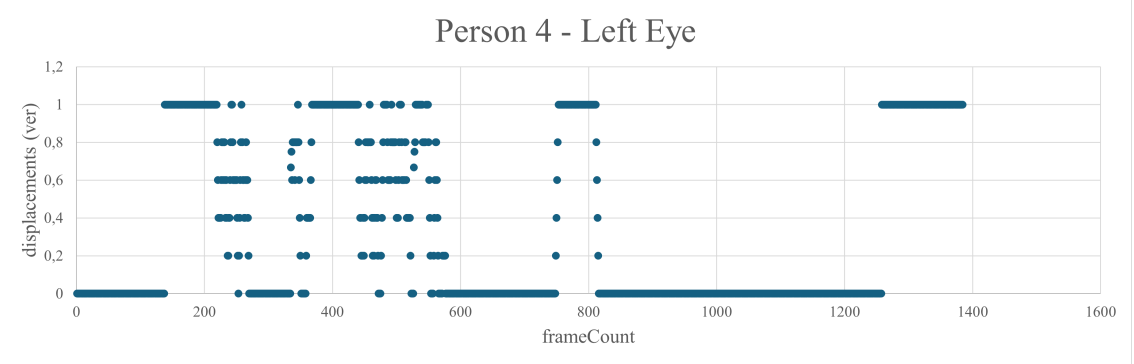


Figure A.24: Graph of the vertical position of person 4's left eye throughout the frames captured by the smartphone camera.

A.2.2 Right Eye

A.2.2.1 Horizontal

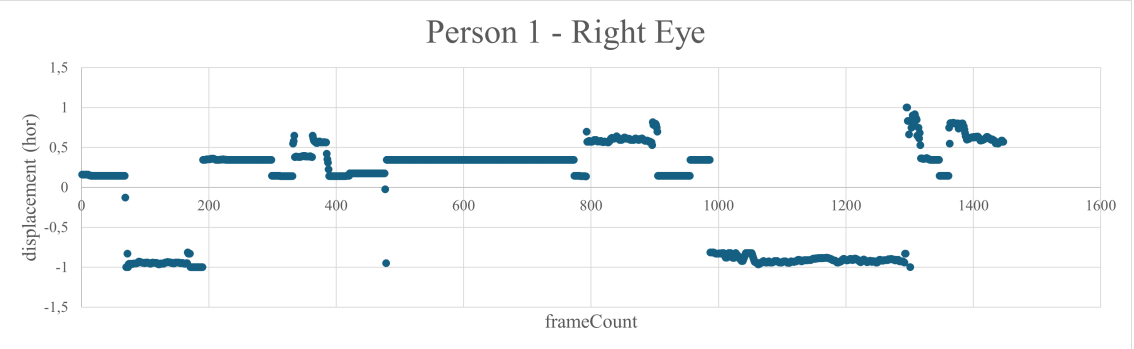


Figure A.25: Graph of the horizontal position of person 1's right eye throughout the frames captured by the smartphone camera.

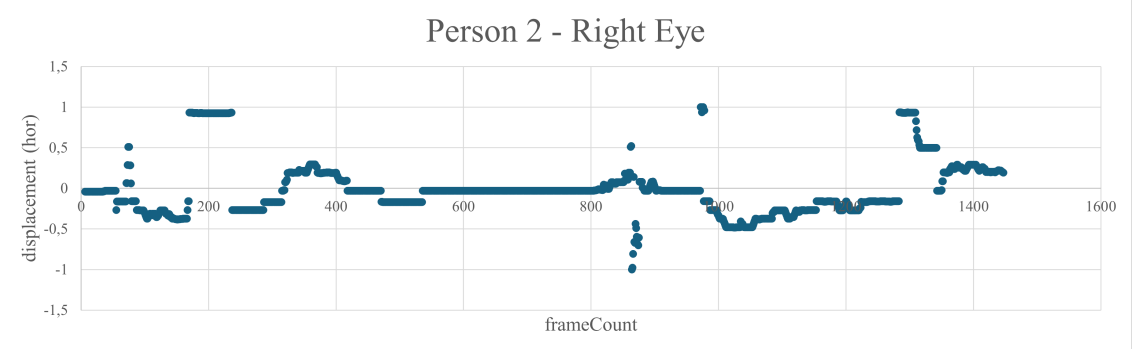


Figure A.26: Graph of the horizontal position of person 2's right eye throughout the frames captured by the smartphone camera.

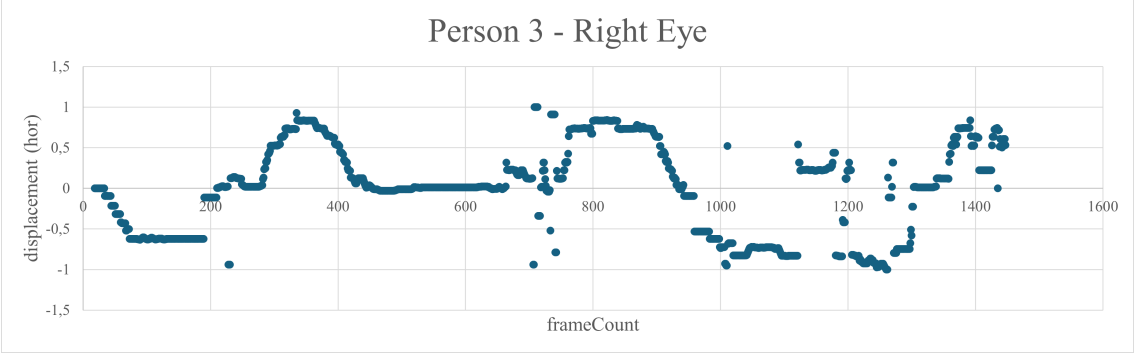


Figure A.27: Graph of the horizontal position of person 3's right eye throughout the frames captured by the smartphone camera.

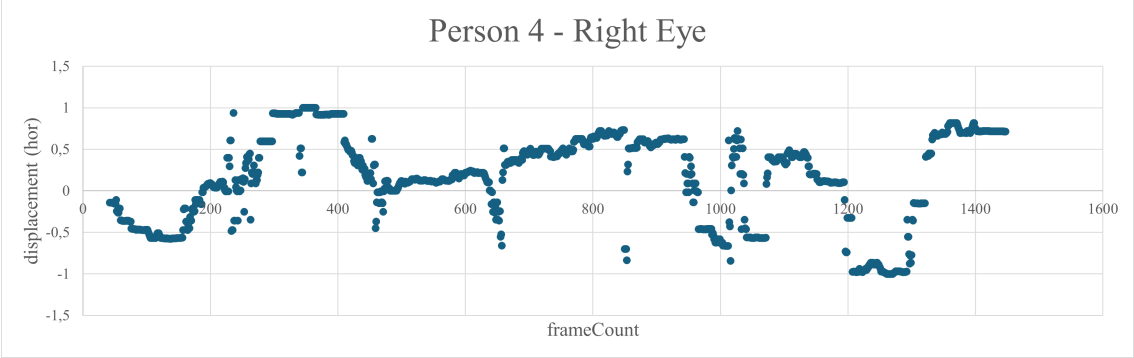


Figure A.28: Graph of the horizontal position of person 4's right eye throughout the frames captured by the smartphone camera.

A.2.2.2 Vertical

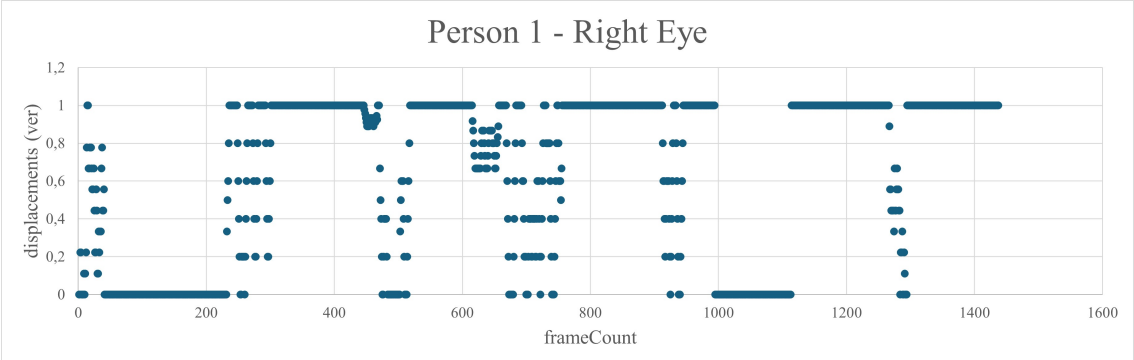


Figure A.29: Graph of the vertical position of person 1's right eye throughout the frames captured by the smartphone camera.

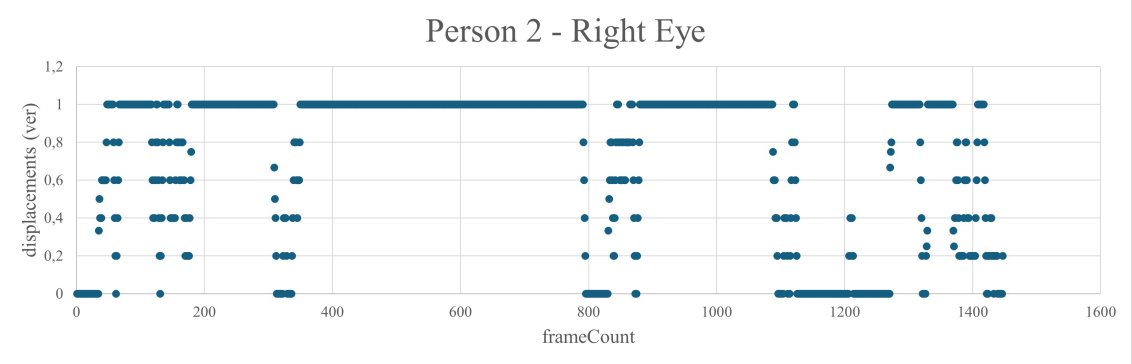


Figure A.30: Graph of the vertical position of person 2's right eye throughout the frames captured by the smartphone camera.

Figure A.31: Graph of the vertical position of person 3's right eye throughout the frames captured by the smartphone camera.

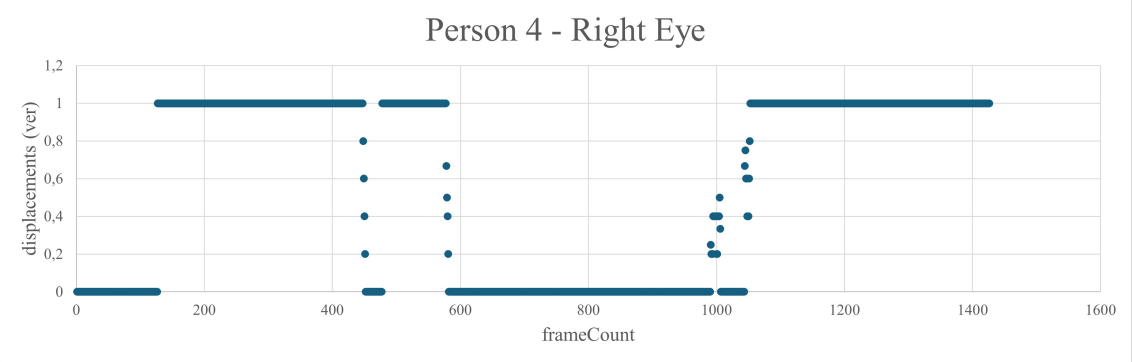


Figure A.32: Graph of the vertical position of person 4's right eye throughout the frames captured by the smartphone camera.