Video-oculography eye tracking towards clinical applications: a review

A. J. Larrazabal[a], C. E. García Cena[b], C. E. Martínez[a]

[a]Research institute for signals, systems and computational intelligence , sinc(i), FICH-UNL/CONICET, Ruta Nac. No 168, km 472.4 (3000), Santa Fe, Argentina

[b] Centre for Robotics and Automation, UPM-CSIC, José Gutiérrez Abascal Street, 28006 Madrid, Spain

Abstract

Most neurological diseases are usually accompanied by a broad spectrum of oculomotor alterations. Being able to record and analyze these different types of eye movements would be a valuable tool to understand the functional integrity of brain structures. Nowadays, video-oculography is the most widely used eye-movements assessing method. This paper presents a study of the existing eye tracking video-oculography techniques and also analyzes the importance of measuring slight head movements for diseases diagnosis. In particular, two types of methods are reviewed and compared, including appearance-based and feature-based methods which are further subdivided into 2D-mapping and 3D model-based approaches. In order to demonstrate the advantages and disadvantages of these different eye tracking methods for disease diagnosis, a series of comparisons are conducted between them, addressing the complexity of the system, the accuracy achieved, the ability to measure head movements and the external conditions for which they have been designed. Lastly, it also highlights the open challenges in this research field and discusses possible future directions.

Keywords: Eye tracking, Eye gazing, Disease diagnoses, Head movements, Saccadic movements

1. Introduction

Neurodegenerative motor disorders are usually accompanied by a broad spectrum of oculomotor alterations. Studies have shown that activity related to eye movements is observed in cortical and subcortical areas, which are directly and indirectly connected with several neural systems which interact among each other to control the suitable performance of the ocular and ocular-cephalic movements. For instance, brain basal ganglia, brainstem nuclei and the vestibular system are organized to produce different eye movements [1, 2].

Therefore, an accurate and detailed eye movements analysis would be a key experimental tool for understanding the functional integrity of brain structures involved in motor and cognitive processing. In addition, it becomes a unique opportunity to detect the presence of different injuries in the nervous center thanks to the existing connection with different oculomotor control abnormalities. These injuries involve a wide variety of neurological diseases including parkinsonian syndromes [3, 4] amyotrophic lateral sclerosis [5], Huntingtons disease [6], Alzheimers disease [7], minimal hepatic encephalopathy [8], among others.

Various types of eye movements are performed in people's daily lives, usually without being aware of them. These movements can be divided into two functional classes [2]. The first class is the one in charge of making the images remain fixed in the retina and comprises the next movements: vestibulo-ocular reflexes (VORs) which make the direction of the eyes remain constant when the head is moved; fixation system which make the gaze resting on a small predefined area; smooth pursuit which describe the eye following a moving object, and optokinetic nystagmus, which stabilize the fovea in relation to objects in the surrounding environment. Some variables related to fixation are commonly measured in different studies including total and mean fixation duration, fixation sequences and fixation rate.

In addition to this, humans use a second class of eye movement, named saccades, in order to rapidly shift the fovea in a stepwise manner onto a new target and bring its superior visual accuracy to bear on objects of interest during fixation. Several of these movements are made each second and they are an intrinsic part of the constant cycle of perception, action and cognition. Measurable saccade related parameters include saccade number, amplitude, fixation-saccade ratio and velocity peak detection. Being able to measure them is also an important contribution.

Besides, there is another kind of saccadic movement which is used alongside fixations. If the eyes remained completely fixed at one point for a long time, the retina would adapt to the constant input and induce the visual image to slowly fades away. In order to avoid the neural adaptation, short saccadic movements, known as microsaccades, shift the image on the retina back and forth in an involuntary manner in small magnitudes ranging from 3 min of arc to 1°[9].

Saccadic and microsaccadic eye movements are likely to be affected by cognitive impairments, as well as by dysfunctions related purely to oculomotor execution. Because of that, they have been extensively studied for a wide range of applications including the drowsiness detection [10], neurological disease diagnosing and sleep disorders studies [11]. On the other hand, within the field of clinical research, fixations are often analyzed in neuroscience, autism alteration studies [12], and psychological studies to determine a persons focus and level of attention [13].

Over the past 20 years, some of these measured variables of eye movements have been selected as possible markers for differential diagnosis including in presymptomatic individuals. Being furthermore eye tracking a less invasive and low cost clinical test compared with other diagnosis methods, this technique has become an invaluable tool for clinicians in diagnosis of neurodegenerative disorders.

Some eye movements abnormalities can be clinically assessed by trained doctors using, for example, Frenzel glasses or ophthalmoscopes [14]. Instead, in some cases it is necessary to make really accurate measurements, like metrics of saccadic accuracy, latencies with respect to stimulus onset, or eye velocity peak estimations which are not possible to be done by simple examination. Furthermore, for some tests with cognitive impairment, it is desirable to present stimulus under specific conditions such as defined target positions, which is also made difficult without the assistance of a synchronized device.

However, eye movements can be easily measured and recorded in the laboratory, covering the main necessities of accurate information. These recordings are highly useful for objective and precise identification of disease status and monitoring of disease progression. For example, increased error rate of antisaccades, which indicates cognitive dysfunction, can only be detected in the laboratory [2]. A number of different eye movements assessing methods have been developed and some of them are currently used. Some examples are electro-oculography [15], scleral search coil system [16], and video-oculography (VOG) [11]. In recent years, the latter has become the most widely used, as it is the only one considered non-invasive and allows for easy coordination of test design and stimuli provision that make it possible to automatically analyse the data.

VOG real-time eye detection and eye tracking is an active area of research in computer vision community since a long time. Presently, there are numerous of gaze trackers devices and software in the market, and there are more and more applications in which these techniques make an important contribution. Human-computer interaction [17, 18], driving safety applications [19], pilot training [20], market and marketing research [21], studies of perception, attention and learning disorders [22] are some examples of these. But, while the idea of eye tracking exists for a long time, recent technological advances enable more precise quantification and automated evaluation making this technology broadly available to disease diagnosis [23, 24].

Naturally, since there are so many possible application regarding eye movements, a lot of different approaches have been proposed for VOG gaze tracking. Although there already exist some papers that review the current gaze tracker systems or methods [11, 25, 26], as far as we know, our review is the only one specifically aimed at analyzing and comparing the VOG gaze estimation methods for application to neurological disease diagnosis and research, taking into account their own requirements and use conditions. In addition, the importance of measuring slight head movement present in some neurological diseases is also analyzed.

2. Eye tracking Video Oculography Techniques

Technically, gaze tracking is the procedure of determining the point-of-gaze (POG) -where one is looking- on some monitor or screen or the visual axis of the eye in the 3D space. For this, video oculography devices are equipped at least with one or more video cameras that send images to a personal computer for image processing.

Recently, several approaches to gaze estimation have been reported in the literature. Although they have been studied focusing on different use conditions, in general, they can be classified in two categories: appearance-based and feature-based methods. Clearly, their choice depends on the application for which they have been designed. The quality of the camera and hardware, the required accuracy, the environmental conditions, the desired cost and the freedom of head movements are some factors that determine the approach to be followed. In the following, these two methods are going to be introduced, in order to compare their specifications with those required for medical applications.

2.1. *Appearance-based methods*

In recent years more and more applications have been developed to analyze human behavior in everyday situations. To do this, it is necessary to monitor the eye movements in uncontrolled scenarios where it is impossible to adjust the lighting conditions, perform a calibration or request the user's assistance. In this context, it is critical to design very robust methods for the different types and qualities of the images.

Fortunately, for most applications very high accuracy is usually not required. For instance, in environmental control or eye typing, where only a few buttons need to be activated, it may be more important to reduce costs by using web cameras, allowing easy and flexible hardware configurations, and avoiding the use of lighting systems and feature detection algorithms.

On the basis of this new approach, several papers [27, 28, 29, 30] have presented methods that work with low-resolution images in different environmental conditions in which appearance-based methods seem to be a promising option. These methods address the gaze estimation problem by learning a mapping function directly from eye images to gaze directions. As input, they use all eye regions pixel values as high-dimensional feature vectors for estimating gaze directions [31]. The output, gaze direction, can be represented as the coordinates (x, y) on the screen where the gaze falls or the rotation angles of the eye with respect to the head position. For low quality images, this is a great advantage compared to the techniques employed by feature-based methods, which have to segment and analyze geometrically derived eye features from high-resolution observations as will be seen in the next section.

The mapping function allows to relate the raw input image with the coordinates of the gaze direction. These functions do not address any particular model, but are designed ad-hoc and are trained with eye images of known gaze direction using various regression techniques, including neural networks [32, 33, 34], local interpolation [35, 36], or Gaussian process [37, 38]. Its formulation depends on the regression technique followed. Figure 1 shows an example where a convolutional neural network (CNN) is used as a mapping function.

These approaches make the system less restrictive, and even though the precision is not good enough for certain applications, they are very robust even when they are applied to relatively low-resolution cameras or

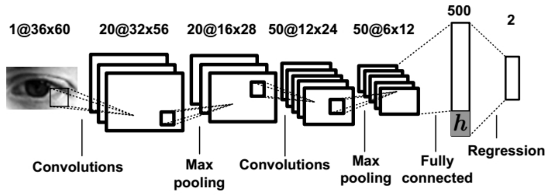

Figure 1: Architecture of the CNN used as mapping function to predict gaze direction. (From Park et al. [27]).

under natural illumination, such as with a phone or computer applications or human computer interaction.

The main problem of these methods is that the appearance of an eye depends not only upon gaze direction but also upon the head poses, imaging conditions and even on the identities of subjects, making it necessary to generate a person-specific training. In addition, due to the high dimensional feature vectors that must be mapped into the gaze directions, thousands of individual training samples are required to calculate the mapping coefficients.

To overcome these limitations, Sugano et al. [39] propose a learning-by-synthesis approach to appearance-based gaze estimation using a large dataset that contains diverse people, head poses, and gaze directions. Also to avoid the need of a person-specific training, Lu et al. [40] extract more advanced eye features, which help to learn a person-independent relationship between eye gaze change and eye appearance variation. On the other hand, Schneider et al. [41] perform embedding for each person in the training set and then learn a linear transformation that maps out the individual, subject-dependent manifolds avoiding the need of individual calibration.

Despite the recent research progress in the field of computer vision, estimating human gaze directions from only eye appearance is still an open challenge. The performance of appearance-based methods generally depends on the quality and diversity of the training data and generalization ability of the regression algorithm. Moreover, their accuracy is not high enough for clinical uses. For these reasons, appearance-based methods can be ruled out for devices designed for this purpose.

2.2. *Feature-based methods*

Methods using extracted local features such as contours, eye corners, and eye reflections, called feature-based methods, are the most popular approach for gaze estimation. These methods use geometrically derived eye features from high-resolution eye-images captured by zooming in the user's eyes (See Fig 2). Once the

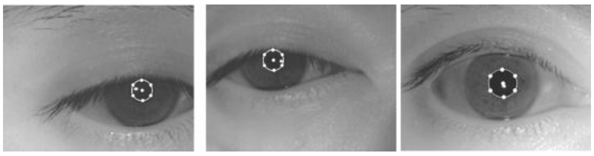

Figure 2: Features from high-resolution eye-images (from Park et al. [43]).

features are extracted, the connection between the gaze directions and them can be modeled in various ways. Besides, depending on whether they are based on eye geometry or not, these methods can be divided into two main groups: 2D mapping-based gaze estimation methods and 3D model-based gaze estimation methods.

The 3D model-based methods [42, 43], directly compute the 3D gaze direction vector from the eye features based on a geometric model of the eye. Then, the point of gaze is estimated by intersecting the gaze direction with the object being viewed, i.e a computer monitor. In order to calculate the center of the cornea and the eye vector, these models require accurate estimation of many user-dependent parameters such as cornea radii, angles between visual and optical axes, the distance between the cornea center and pupil center, among others. To understand why these parameters should be estimated, and which complex hardware calibration should be made during initial setup, the model proposed by Guestrin et al. [44] will be developed. This example is also a good basis for understanding model-based methods. The model and their parameters are shown in Figure 3.

Considering a ray that comes from the light source I_i, reflects at a point $\mathbf{q}_{i,j}$ on the corneal surface, which is modeled as a convex spherical mirror of radius R, passes through the nodal point of the camera $\mathbf{o}_j$, and intersects the camera image plane at a point $\mathbf{u}_{i,j}$, the next two equations can be formulated:

$$\mathbf{q}_{ij} = \mathbf{o}_j + k_{q,ij}(\mathbf{o}_j - \mathbf{u}_{ij}) \ for \ some \ k_{q,ij} \tag{1}$$

$$||\mathbf{q}_{ij} - \mathbf{c}|| = R \tag{2}$$

In addition, based on the beam reflection laws, two more equations can be raised for these points.

$$(\mathbf{i}_i - \mathbf{o}_j) \times (\mathbf{q}_{ij} - \mathbf{o}_j) \bullet (\mathbf{c} - \mathbf{o}_j) = 0 \tag{3}$$

$$\begin{aligned} &(\mathbf{i}_i - \mathbf{q}_{ij}) \bullet (\mathbf{q}_{ij} - \mathbf{c}) \cdot ||\mathbf{o}_j - \mathbf{q}_{ij}|| \\ &= (\mathbf{o}_j - \mathbf{q}_{ij}) \bullet (\mathbf{q}_{ij} - \mathbf{c}) \cdot ||\mathbf{i}_i - \mathbf{q}_{ij}|| \end{aligned} \tag{4}$$

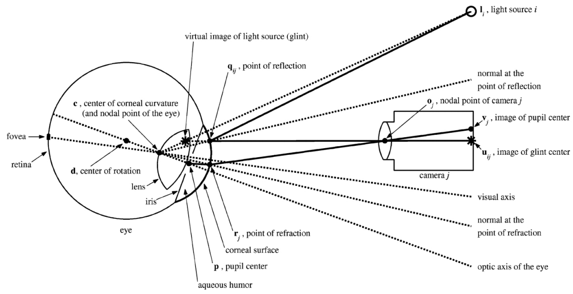

Figure 3: Schematic representations of the eye, a camera, and a light source (from Guestrin et al. [44]).

In the same way, considering a ray that comes from pupil center $\mathbf{p}$, refract at the point $\mathbf{r}_j$ on the corneal surface, passes through the nodal point of camera $\mathbf{o}_j$, and intersects the camera image plane at a point $\mathbf{v}_{ij}$, two more equations can be obtained.

$$\mathbf{r}_{ij} = \mathbf{o}_j + k_{r,j}(\mathbf{o}_j - \mathbf{v}_{ij}) \ for \ some \ k_{r,j} \tag{5}$$

$$||\mathbf{r}_{ij} - \mathbf{c}|| = R \tag{6}$$

Then, applying beam refraction laws, the following equations are derived where n_1 and n_2 are the refraction index of the aqueous humor and cornea combined and of air respectively.

$$(\mathbf{r}_j - \mathbf{o}_j) \times (\mathbf{c} - \mathbf{o}_j) \bullet (\mathbf{p} - \mathbf{o}_j) = 0 \tag{7}$$

$$\begin{aligned} &n_1||(\mathbf{r}_j - \mathbf{c}) \times (\mathbf{p} - \mathbf{r}_j)|| \cdot ||\mathbf{o}_j - \mathbf{r}_j|| \\ &= n_2||(\mathbf{r}_j - \mathbf{c}) \times (\mathbf{o}_j - \mathbf{r}_j)|| \bullet ||\mathbf{p} - \mathbf{r}_j|| \end{aligned} \tag{8}$$

Finally, considering K as the distance between the pupil center and the center of corneal curvature leads to:

$$||\mathbf{p} - \mathbf{c}|| = K \tag{9}$$

By means of solving the proposed system of equations for $\mathbf{c}$ and $\mathbf{p}$, the optic axis of the eye in the space can be reconstructed as the line defined by these two points. It is important to note that to solve these equations, all the subject-specific parameters (R, K and n_1) have to be known. In general, if only one camera is available, they are obtained by the calibration process -detailed below-. Also, the angle between the optic axis and visual axis must be calculated and is usually done during the calibration procedure.

This parameters also rely on metric information requiring camera calibration and exact knowledge of the light sources and monitor position. These values may be directly measured once during the first setup but, to achieve a high accuracy, the eye parameters need to be estimated independently for each individual, making that a previous calibration step cannot be omitted.

The 3D model-based approaches can handle head movements in a robust manner with high accuracy but involving this relatively complex initial setup. They need to use at least a single camera with multiple calibrated light sources [44] or stereo cameras [45, 46, 47]. Even so, for some clinical diagnoses, it is important to be able to differentiate between oculocephalic and pure eye movements, so calculating the absolute position of the gaze is not always useful. Furthermore, regardless of the model complexity, the calibration might be only simplified, but not avoided at all. In some works, to avoid the calibration process, a very simplified eye model is used. While it reduces calibration times and complexity, the accuracy obtained also greatly decreases.

On the other hand, the 2D mapping approaches [48, 49, 23] are based in finding a mapping function from 2D feature space like Pupil-Center-Corneal-Reflections (PCCR), contours, etc. to gaze point such the computer screen coordinates. That function avoids the need for the direct measurement or estimation of the eye model parameters throughout the system setup. Instead, they are implicitly included in the learning of the mapping function simplifying the setup process itself. The same happens with the camera calibration process and the system geometry determination.

Different features are used as inputs to the mapping function depending on the application and the image conditions. Mostly, they can be further divided into active light techniques such as PCCR or passive light techniques such as shape-based methods, depending on whether they require external light sources to detect eye features.

In the recent years, eye tracking applications using webcams under natural illumination have gained highly relevance in the community. In particular, passive image-based algorithms for eye localizing and tracking in the visible spectrum have been researched over the last years [50, 27, 51]. These algorithms propose the search for some features like iris or pupil center. For the purpose of iris tracking, the limbus, which is the boundary between the sclera (normally white) and iris (comparatively dark) is optically detected and tracked. Pupil tracking is similar to iris tracking except that a smaller boundary between iris and pupil is used for relative measurement.

Although without active illumination it is easier to segment the limbus due to the higher contrast between the iris and the sclera compared to the contrast between the pupil and the iris, pupil tracking has a lot of advantages. The pupil, which is much less covered by the eyelids than the limbus, enables vertical tracking. In addition, the sharper edge between the pupil and the iris provides a higher resolution.

Various iris and pupil center localization methods have been reported in the literature [52, 53]. Several treat iris or pupil center localization as a circle detection or ellipse fitting problem [54, 55, 56, 57]. Depending on the viewing angle, both iris and pupil appear elliptical and consequently can be modeled by different shape parameters. Simple ellipse models consist of voting-based methods [58, 59] and model fitting methods [60, 61]. Once the iris center has been successfully localized, regression-based methods can be used for finding the corresponding gaze points on the screen.

Since these methods directly map the eyes iris center or pupil center location to a target plane such as the monitor screen, the accuracy and robustness of the center localization significantly affect the performance of gaze tracking. For example, detection has some problems when the iris moves toward the corners or when the upper and lower boundaries of the iris are occluded by the eyelids and eyelashes, leading to gaze estimation errors.

On the other hand, for applications like clinical research, where experiments are performed in a doctor's office, it is not a problem to have infrared lighting, and thus active methods would be a better option. PCCR is the most common approach for feature-based gaze estimation methods.

When a light source (usually infrared) illuminates the eyes at different layers, the boundaries between the lens and the cornea act as convex mirrors and produces some reflections or virtual images, which are called corneal reflections or Purkinje images. In particular, the Purkinje image formed by the reflection of the outer surface of the cornea, called the first Purkinje image, is known as glint. The glint is the brightest and easiest reflection to detect and track. The PCCR technique uses the vector formed by the subtraction between the estimated center of the pupil and one or more near infrared (NIR) corneal reflections to estimate the gaze direction [62, 49].

To compute the pupil-glint vector, the pupil center must also be extracted from the image. As it was already mentioned, different techniques are avaible for doing this but, with active illumination, the bright pupil-dark pupil method (BP-DP) is one of the most widely used for determining the accurate location of the pupil. When a light source is placed collinearly to the optical axis of the camera, most of the light is reflected back to the camera and the eye image shows a bright pupil. Conversely, when a light source is located away from

sinc(*i*) Research Institute for Signals, Systems and Computational Intelligence (fich.unl.edu.ar/sinc)
A. Larrazabal, C. García Cena & C. E. Martínez; "Video-oculography eye tracking towards clinical applications: a review"
Computers in Biology and Medicine, Vol. 108, pp. 57-66, 2019.

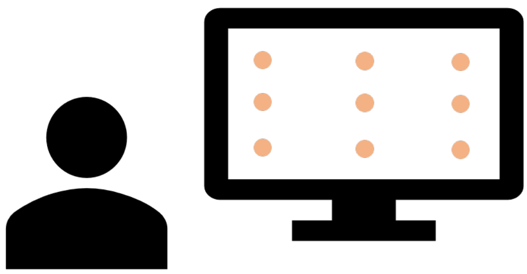

Figure 4: Example of the calibration points.

the camera's optical axis, the image shows a dark pupil. Therefore, eye trackers with active IR illumination can use the difference between dark and bright pupil images by synchronously switching between the two light sources. This technique is very simple and robust in controlled conditions [11].

Besides that, the detection of the corneal reflections requires a narrow field of view (FOV) camera (long focal length) since the reflections are in general very small. Therefore, these systems work with high-resolution eye images captured by zooming in on movement-restricted users. Under these conditions, eye features can be easily and robustly extracted, this being an advantage over other methods.

These reported techniques are widely used and achieve really good results, but they have two major issues. First, because the mapping function is different for each person and for each system configuration, it is necessary to perform a tedious calibration procedure before each test to obtain the necessary parameters. In a typical calibration procedure, a set of visual targets such as those shown in Figure 4, is presented to the user who normally has to stare to the computer for a period while the corresponding measurement is being done. Afterwords, from these correspondences, a mapping function is calculated.

The second drawback is that once the calibration has been performed, the person's head must remain motionless. Otherwise, there will be large errors between the actual and estimated directions. To avoid these errors head restraint systems are often used, and the calibration process is repeated every time movements are observed in the patient. Figure 5 shows an active feature-based system, and the restraint system it uses to prevent head movement. With this device, Hernandez et al. [23] achieve an accuracy of less than 0.4°, reported as one of the minimum reaches in the literature.

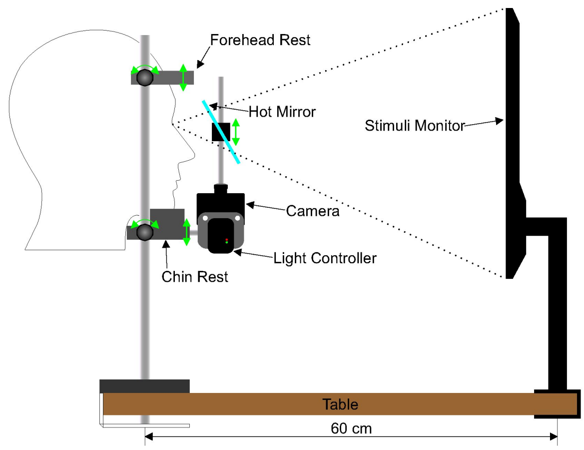

Figure 5: Example of a gaze tracker with a restraint system (from Hernandez et al. [23])

A number of efforts are being made to minimize these shortcomings. 2D mapping methods assume that the mapping function have a particular parametric form such as a polynomial or a non-parametric form such as a neural networks whose coefficients have no physiological or physical meaning.

Polynomial interpolation is one of the main tools for parametric mapping functions, mainly due to its simplicity of execution and the good quality of the result obtained from it. In this case, the x and y gaze coordinates are estimated by means of a polynomial function. For example, a second order polynomial transformation is defined as:

$$\begin{aligned} x_c &= a_1 + a_2 x_e + a_3 y_e + a_4 x_e y_e + a_5 x_e^2 + a_6 y_e^2 \\ y_c &= b_1 + b_2 x_e + b_3 y_e + b_4 x_e y_e + b_5 x_e^2 + b_6 y_e^2, \end{aligned} \tag{10}$$

where (x_c, y_c) is the coordinate of the point on the screen where the gaze falls, (x_e, y_e) is the coordinate of the pupil-glint vector and a_i; b_i are the polynomial coefficients. These coefficients are calculated in the calibration procedure. During this procedure, the patient is asked to stare at a set of known targets, while a set of corresponding points are obtained. For example, for a 6-point calibration procedure, 6 corresponding points are obtained (x_{ci}, y_{ci}); (x_{ei}, y_{ei}) with $i = 1, 2, ...6$ and a system of 12 equations is generated to calculate the polynomial coefficients a_i; b_i, by applying the least squares estimation procedure, that is, minimizing the quadratic error E^2 between the estimations and the calibration points coordinates. The higher is the order of the polynomial mapping function, the greater will be the number of calibration points needed to calculate all coefficients. In Equation 11 the quadratic error function for N calibration points is displayed.

$$E_x^2 = \sum_{i=1}^{N}[x_{ci} - (a_1 + a_2x_{ei} + a_3y_{ei} + a_4x_{ei}y_{ei} + ..)]^2$$
$$E_y^2 = \sum_{i=1}^{N}[y_{ci} - (b_1 + b_2x_{ei} + b_3y_{ei} + b_4x_{ei}y_{ei} + ..)]^2 \tag{11}$$

Mimica et al. [63] use a second order polynomial to minimize the number of calibration points required comparing to those required by a higher order polynomial. Cerrolaza et al. [64, 65] carried out a study on the potential effect of the order and systematic inclusion of all polynomial terms, on the accuracy and robustness of the gaze tracker. For this, a real VOG system with different configurations was used. The authors point out that the gaze estimation accuracy of a gaze tracking system is not noticeably increased with the enhancement of polynomial order or with more complete mathematical expressions due to the factors of head motion, and calculation method of the pupil-glint vector.

The choice of the mapping function determines not only the accuracy of the system but also the head movement tolerance and the calibration time. Therefore, when linear regression solution methods are applied to solve the mapping function, a second-order linear polynomial is the most used due to its advantages of less calibration markers and better approximation effect.

Alternatively, Baluja et al. [32] first proposed a method using a simple artificial neural network (ANN) to calculate a non-linear mapping function. First, they mapped images of only the pupil and cornea as the inputs to ANN to the coordinates of the gaze point as the outputs. Then, they included the total eye socket as an input to improve the system accuracy (about 1.5°). In addition, Zhu and Ji [66] utilize generalized regression neural networks to estimate the gaze direction. For this purpose, 6 pupil and glint parameters were used as inputs to the calibration procedure. The parameters were chosen in such a way that they represent eye and head movements and remain relatively unchanged for different people. Therefore, even though the accuracy acquired is not good enough (about 5°), it is a free calibration process and head movements are allowed.

In a similar way, Gneo et al. [67] utilize multilayer neural feedforward networks to calculate gaze point coordinates based on pupil-glint vectors. In order to minimize the number of output neurons, they use one separate network with the same input for each gaze coordinate (x, y). The reported results were competitive with high accuracy (about 0.6°). More recently, Wang et al. propose in [49] an improved ANN based on direct

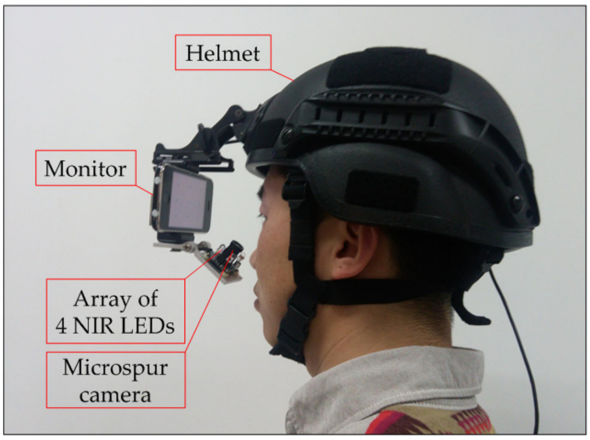

Figure 6: Example of a head-mounted gaze tracker system (from Wang et al. [49]).

least squares regression to calculate the mapping function between pupil-glint vectors and actual gaze points. They combine the advantages of both methods: the high speed of direct least squares regression and the high accuracy of ANN. They achieved a good accuracy (about 0.4°) in a head-mounted device which can be seen in Figure 6.

Thus, as it was pointed out before, the choice of the model depends on multiple factors: required accuracy, hardware cost, image quality/eye region resolution, available information in the image (e.g., glints), and configuration flexibility. For instance, feature-based methods accuracy may decrease when model assumptions are violated. In some applications such as clinical research or disease diagnosis, where it is possible to control the illumination conditions, the camera's quality, and the system settings, these methods achieve a really high accuracy which is critical to investigate imperfections in the oculomotor system [68].

Moreover, despite the fact that mapping methods provide little information about the intrinsic behavior of the system, they are much simpler to construct than the model-based methods and do not require additional hardware calibration, which makes setup much faster for the system user. That is why, most commercial gaze tracking systems use 2D mapping features-based methods with IR camera and active IR illumination, as it is shown in Figure 7, to achieve the highly accurate performance of gaze estimation.

2.3. Head Movements

So far we have talked about gaze tracking as a process that belongs exclusively to the eyes, but it is known that the gaze is a product of two contributing factors, the

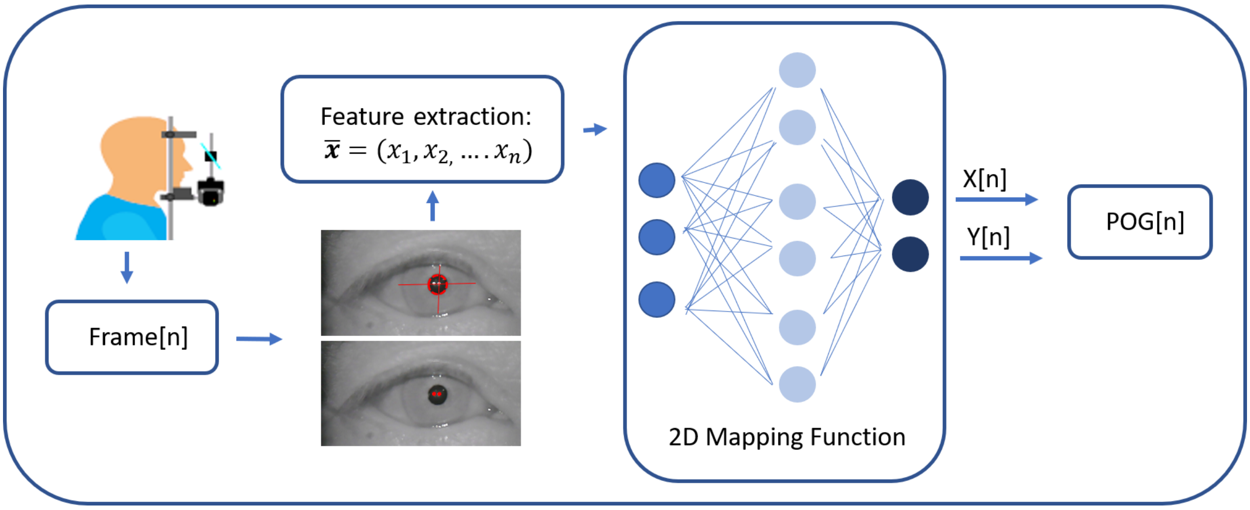

Figure 7: Diagram of a standard eye tracker with 2D mapping method

head pose (position and orientation) and the eyeball orientation. A person can change gaze direction by rotating the eyeball while keeping the head stationary; similarly, a person can change gaze direction by moving the head while keeping the eye stationary relative to the head.

Usually, a person moves the head to a comfortable position before orienting the eye. Head pose, therefore, determines the coarse-scale gaze direction while the eyeball orientation determines the local and detailed gaze direction. For example, when the target is located over 20°of the field of view, it is much more comfortable to rotate the head than to rotate the eyeball.

While in the previous section we reviewed the state-of-the-art in gaze estimation systems, focusing on tracking eye movement, the problem of ensuring the invariance to head movements is also an important and a challenging research topic. In almost all applications, people move their heads while they are using a gaze tracker. For this reason, for an accurate gaze estimation, it is necessary to (either directly or implicitly) model both head pose and eye rotation.

As we pointed out above, appearance-based methods have been developed in order to be applied in environmental conditions without the possibility of monitoring user conditions. This user freedom movement situation requires the method to be robust to changes in head position. For solving this problem two possibilities can be held: learning generic gaze estimators from large amounts of head pose-independent training data or adding head pose information to the eye images.

In particular, Lu et al. in [69] address the head motion problem by synthesizing new training images for different head poses from those already seen in estimation, while in [70] they perform the gaze estimation by assuming a fixed head pose and then compensating for the estimation biases caused by the head pose using a head pose tracker. Conversely, Zhang et al. in [27] used a multimodal convolutional neural network to learn a mapping function from both the head poses and eye images to the unique gaze directions. Lai et al. in [71] also combine the eye image information with head pose tracking selecting features by means of the neighborhood-based regression algorithm. Although these results outperform the appearance-based methods state of the art, the accuracy achieved is still quite poor what that required for clinical applications.

Regardless of the eye tracker method applied, other researchers have attempted to measure the head movements or head position directly, and to use that information to correct the gaze measurements. These systems usually include two or more cameras and use a complex facial model to track the movement of the face [72, 73]. In order to track the face from these models, the FOV of the tracking camera has to be large enough to cover the entire user's head. This is not a problem for daily applications which do not need such a high accuracy like in [74], but, when feature-based methods are applied, these restrictions make it difficult to locate the small eye features and result in less accurate gaze tracking.

To overcome this drawback, some works have proposed the use of two cameras in combination with pan and tilt mechanisms that allow freedom of person motion while maintaining the feature method accuracy. In [75], Hennessey et al. proposed a system that rotates an

sinc(i) Research Institute for Signals, Systems and Computational Intelligence (fich.unl.edu.ar/sinc)
A. Larrazabal, C. García Cena & C. E. Martínez; "Video-oculography eye tracking towards clinical applications: a review"
Computers in Biology and Medicine, Vol. 108, pp. 57-66, 2019.

eye tracker with a narrow-angle camera using pan and tilt servo motors, while in [76] Cho et al. propose a binocular eye gaze tracking system that, using pan, tilt and zoom movements, continue to track the eyes with a narrow camera while the user moves his head freely in depth. Then both estimate the POG by a 2D mapping function modified with the depth of the eyes. All these methods work relatively well but are very complex, expensive and, most notably, slow. These limitations restrict its use so that in practice, head pose information is rarely used directly in the gaze models. It is more common to incorporate this information implicitly either through the mapping function (regression-based method) or through the use of reflections on the cornea (3D model-based approaches).

Apart from that, the 3D model-based methods are the most robust to head pose changes and they can obtain the head pose invariance through various hardware configurations and prior knowledge of the geometry and cameras. Guestrin et. al [44] presented a general study for PCCR covering all the possible system configurations in terms of number and positioning of IR light sources and cameras. In that work the authors claimed that using only one camera with two light sources is the simplest configuration that allows for both the estimation of POG and free head user movements. To estimate the POG with this system configuration, it is necessary to make a subject-specific calibration procedure that requires the subject to fixate on multiple points. To avoid the need of calibration, an additional camera is also necessary.

Using one camera and one light source the POG can be estimated only if the head is completely stationary. This restriction is shared with the 2D regression methods, which assume static head conditions. In general, gaze estimation systems that use one camera and one light source assume that the head movements are negligible. Therefore, it should be noted that the only video oculography method that allows large head movements while maintaining good accuracy are some 3D model-based methods. But as a drawback, they require a really complex and calibrated system setting, difficulting their use in common applications. In addition, the movement of the head is taken into account implicitly, and it is not possible to differentiate between oculocephalic and purely ocular movements.

This is another reason why the eye location algorithms found in commercially available eye trackers use the 2D regression techniques where the 3D eye location is usually unknown and only the relative orientation of the user's eye with respect to the user's head is measured. Particularly, gaze estimation is based on the relative position between pupil and glint. Assuming a static head, methods based on this idea use the glint as a reference point, thus the vector from the glint to the center of the pupil will describe the gaze direction. While contact-free and non-intrusive, these methods work well only for a static head, but even minor head can fail these techniques.

In addition, since the pupil and glints are very small, the FOV of the eye tracking camera has to be confined to obtain a high definition eye image. This aspect also limits the head so that the eye does not disappear from the FOV and emphasizes the problem of sensitivity to head pose variations, requiring the user to be either equipped with a head-mounted device or to use a high-resolution camera combined with a chin rest to limit the allowed head movements.

In many clinic applications where tests are conducted for only a few minutes and it is not easy to perform complex system calibrations, those restraint systems are suitable and work very well. Even so, despite the fact that the head movements are restricted, in people with certain neurological diseases, it is possible to observe some involuntary movements and it is very important be able to measure them not only to correct the gaze estimation errors but also because these measurements are indicators of the presence or progress of certain diseases.

As it was pointed out before, there are a lot of research works that deal with the problem of obtaining enhanced gaze estimation in presence of large movements and head pose variations for daily applications [74]. But, despite their importance for clinic diagnosis, there are not many studies performing the feasibility of a gaze estimator that considers both the head and eye movements in a really zoomed and high-resolution images with the objective to detect short involuntary head movements.

Some works study the way of achieving head pose invariance for 2D regression methods, for example [66] use Generalized Regression Neural Networks (GRNN) instead of polynomial functions to account head implicitly by the gaze mapping function. They also include more parameters as mapping function inputs like glint coordinates and pupil radio to account for the different head motions. In [77] authors also employed GRNN but with the aim of compensating the errors generated by the non-linear polynomial for different head poses obtaining a gaze estimation robust to head but with much lower spatial gaze resolution.

Despite these advances, there are no studies done to quantify these slight movements based only on the zoomed eye image which is necessary for some neurological disease diagnosis.

3. Clinical applications

One of the best-known clinical applications of high-precision eye tracking is refractive surgery assistance [78, 79]. With this type of surgery, eye refractive disorders (myopia, hypermetropia and astigmatism) are corrected by modelling the cornea by laser ablation [80]. As expected, this procedure requires great precision, and although the patient is asked to look at a fixed point during the few minutes that the surgery lasts, the patient's eye is not fixed and small movements may occur. Therefore, modern laser devices are equipped with sophisticated eye-tracking systems, which by detecting and tracking the pupil and limbus, follow the eye movements and rectify in real time the position of the mirrors that direct the laser. To achieve this adjustment, cameras with acquisition frequencies of the order of 1000 Hz and latencies of less than 3 ms are used [81]. Most of these devices perform the compensation for linear eye movements (in the x/y-axis), rolling eye movements and eye torsions around the visual axis known as cyclotorsions [82].

As discussed in the introduction, eye tracking is used not only as an assistant for clinical applications but also for the diagnosis of neurological diseases, applications on which this review focuses. In the last few years a series of tests have been developed to analyse different alterations in eye movements. In particular, it has been found that the alterations of saccadic eye movements (SEM) provide relevant information.

A widely used test to measure these movements consists of asking the patient to look at a point in the centre of the screen until a new point appears on the periphery. At this moment, the patient must look towards it (Pro-Saccade task) or the opposite side (Anti-saccade task) as rapidly and as accurately as possible. In addition, the second target can be turned on before the first target is turned off (Overlap) or a few milliseconds after the first target is turned off (Gap). By combining these paradigms, other metrics can be obtained for analysis. This type of test are very simple to perform for the patient and allow to measure different parameters to be used as markers such as those that can be observed in Figure 8.

Numerous studies based on this type of test have been carried out in patients with Alzheimer's disease (AD) [84], where it has been possible to distinguish some of the ocular movement alterations, finding significant contributions for the diagnosis of the disease, even in its early stages. In [85, 83] the authors presented a

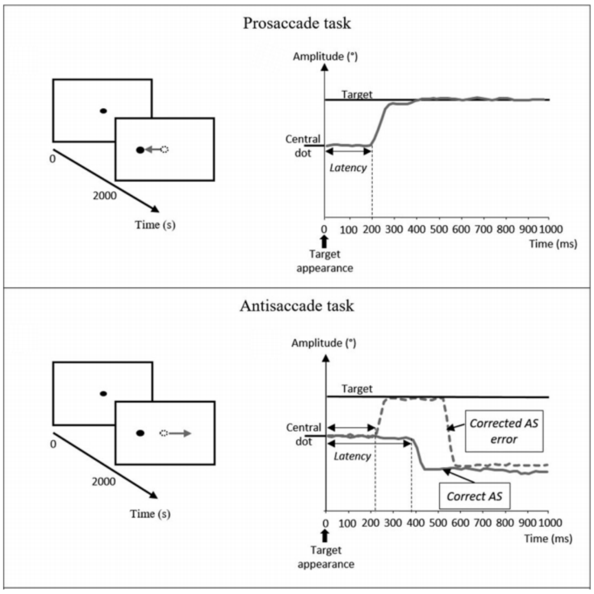

Figure 8: Dynamics of the saccade as a function of time in the Saccade tests. (from Noiret et al. [83]).

study of SEM in patients with Alzheimer and healthy control patients finding significant differences. For instance, it was found that AD patients had higher latency and latency variability regardless of the tasks. Moreover, AD patients made more uncorrected Anti-Saccade (AS) and took more time-to-correct incorrect AS. In addition, close relationships were found between the majority of SEM variables and dementia screening tests, especially the MiniMental State Examination (MMEE) and episodic memory measure. Also in [86] the AS test was performed, finding significant differences in both the number of AS errors and the number of errors that remain uncorrected.

Parkinson's disease and other parkinsonian syndromes [87] are another of the conditions in which this type of measurements provides very important information [88, 89]. One of the most prominent features of eye movement abnormality in PD are saccade hypometria, abnormally fragmented saccades, called multistep or staircase saccades, an increment of the latency of the saccades and a marked difficulty in inhibiting the saccade movement reflex during the test of the AS [87].

Measurements of eye movement abnormalities have also drawn attention as a biomarker in the diagnosis of multiple sclerosis (MS) [90]. For this task, a particular paradigm known as the endogenously generated sacadde paradigm is used. The same is similar to the pro-sacadde test but an intermediate step is added to it. After the initial target disappears from the screen, an

sinc(i) Research Institute for Signals, Systems and Computational Intelligence (fich.unl.edu.ar/sinc)
A. Larrazabal, C. García Cena & C. E. Martínez; "Video-oculography eye tracking towards clinical applications: a review"
Computers in Biology and Medicine, Vol. 108, pp. 57-66, 2019.

Table 1: Comparison of the different methods.

Method	Ref.	Accuracy (degrees)	Head	Calibration	System Setup
Appearance	[39]	6.5°	Free head pose	Free	1 Camera
	[40]	7.8°	Not considered	Free	1 Camera
	[27]	6.9°	Free head pose	Free	1 Camera
	[71]	2° 5°	Allowed (tracked)	Needed	1 Camera
	[69]	2.5°	Free head pose	Needed	1 Camera
Model	[44]	0.9°	Moderate (2-3 cm)	Multiple points	1 Camera - 2 light sources
	[45]	< 1°	Moderate (10 dm^3)	One point	2 Cameras - 2 light sources
	[46]	< 1°	Yawing and pitching allowed	One point	2 Cameras - multiple light
	[47]	> 1°	Natural head movements	4 points	2 Cameras - 2 light sources
	[47]	> 1°	Natural head movements	4 points	2 Cameras - 2 light sources
Shape	[52]	2.42°	Fix	9 points	1 Camera
	[57]	1°	Allowed (tracked)	Needed	2 Cameras
	[60]	4°	Fix	4 points	1 Camera
	[74]	2-5°	Allowed (tracked)	Needed	1 Camera
2D Mapping	[63]	0.8°	Fix	9 points	1 Camera - 2 light sources
	[65]	0.38°	Fix	16 points	1 Camera - 2 light sources
	[23]	0.4°	Fix	9 points	1 Camera - 2 light sources
	[32]	1.5°	Allowed	Moving mouse	1 Camera
	[66]	5°	Natural head movements	Free	1 Camera - 2 IR arrays
	[67]	0.6°	Fix	4 x 5 grid	1 Camera - 3 IR arrays
	[49]	0.4°	Head mounted	16 points	1 Camera - 4 light sources
	[75]	<2°	Free head pose (Tracked)	9 points	2 Cameras - 1 IR plate 1 pan-tilt mechanism
	[76]	0.69	Back and forth movements	6 points 3 depth	4 Cameras - 16 NIR leds pan-tilt mechanism
	[77]	-	Natural head movements	9 points	1 Camera - 1 light source

arrow is shown on it whose direction may be equal or opposite to that of the final target. With such studies, saccade latencies were found to increase in patients with MS even as a function of the disease duration [91]. In addition, these patients frequently exhibit fatigue symptoms at any stage of the disease, having a major impact on their quality of life. Recent reports have shown an increase in endogenous saccade latencies and a reduction in the peak velocities associated with this symptom [92].

4. Conclusions

In this article, several gaze tracking algorithms and their respective advantages and disadvantages for use in disease diagnosis were analyzed. Table 1 summarizes the state of the art of VOG methods by grouping them according to their classification. The method accuracy is calculated as the mean angular error of the gaze estimation. In addition, the system setup (number of cameras and infrared lamps), the number of calibration points used, and the allowed head movements are specified. It should be noted that it would be a significant contribution to be able to provide the resolution of the hardware used, unfortunately this information is not available in many of the original works. In addition, due to the variety of methods presented, providing only the camera resolution would not be enough to account the hardware resolution. Since in some cases the images are taken from the whole face while in others are taken with zoom on the eye, these differences make it very difficult to uniquely characterize acquisition systems.

Appearance-based and passive-shape based methods are not suitable for clinical applications because of their low accuracy. These methods have been researched for daily applications which are carried out in natural illumination conditions and/or with low-resolution cameras but not for clinical applications, where the experimental environment can be controlled with the possibility of even using infrared illumination.

3D model-based methods have an excellent head tolerance. However, the hardware requirements for their implementation is really complex as they need several light sources, multiple cameras and a perfect system calibration. Besides, it is not possible to differentiate between oculocephalic and purely ocular movements.

Thus, we conclude that active 2D regression-based methods are the best option for clinical diagnosis or research applications, since they use features coming from the human eye, such as pupil center and corneal reflections, and they can be implemented using a single camera and a few NIR LEDs. However, these techniques are very vulnerable to head movements and require users to hold their head very still using a headrest, chin rest or bite bar. This restraint system are not enough in presence of some neurological disease, where involuntary head movements occur.

In light of the aforementioned advantages and disadvantages of each methodology considered, it is clear that it is still an open issue to find an optimal way of measuring the head movements, to be able to use both to correct the gaze point estimation and as another biomarker in the disease analysis. We consider that the path to follow to reduce these shortcomings is to focalize on the application for which the device is being designed. In the case of eye trackers for neurological disease diagnosis, it would be helpful to exploit the hardware functionalities where the patient's movements are restricted to a minimum of space during the test duration. In this way, research could be focused on methods that also incorporate measurements of small displacements or rotations from the initial position.

Acknowledgements

The authors wish to thank the support of the *Universidad Nacional de Litoral*, and the *Consejo Nacional de Investigaciones Científicas y Técnicas* (CONICET) from Argentina.

[1] Martin Gorges, Elmar H Pinkhardt, and Jan Kassubek. Alterations of Eye Movement Control in Neurodegenerative Movement Disorders. *Journal of Ophthalmology*, 2014, 2014. Article ID 658243, 11 pages, doi: 10.1155/2013/453402.

[2] Tim J Anderson and Michael R MacAskill. eye movements in patients with neurodegenerative disorders. *Nature Reviews Neurology*, 9(2):74, 2013.

[3] Elmar H Pinkhardt and Jan Kassubek. Ocular motor abnormalities in parkinsonian syndromes. *Parkinsonism & related disorders*, 17(4):223–230, 2011.

[4] Elena Pretegiani and Lance M Optican. Eye Movements in Parkinson's Disease and Inherited Parkinsonian Syndromes. *Frontiers in Neurology*, 8:592, 2017.

[5] Colette Donaghy, Matthew J Thurtell, Erik P Pioro, J Mark Gibson, and R John Leigh. Eye movements in amyotrophic lateral sclerosis and its mimics: a review with illustrative cases. *Journal of Neurology, Neurosurgery & Psychiatry*, 82(1):110–116, 2011.

[6] Stephen L Hicks, Matthieu PA Robert, Charlotte VP Golding, Sarah J Tabrizi, and Christopher Kennard. Oculomotor deficits indicate the progression of Huntington's disease. In *Progress in brain research*, volume 171, pages 555–558. Elsevier, 2008.

[7] Gerardo Fernández, Pablo Mandolesi, Nora P Rotstein, Oscar Colombo, Osvaldo Agamennoni, and Luis E Politi. Eye movement alterations during reading in patients with early alzheimer disease. *Investigative ophthalmology & visual science*, 54(13):8345–8352, 2013.

[8] Sara Montagnese, Harriet M Gordon, Clive Jackson, Justine Smith, Patrizia Tognella, Nutan Jethwa, R Michael Sherratt, and Marsha Y Morgan. Disruption of smooth pursuit eye movements in cirrhosis: relationship to hepatic encephalopathy and its treatment. *Hepatology*, 42(4):772–781, 2005.

[9] Jorge Otero-Millan, Jose L Alba Castro, Stephen L Macknik, and Susana Martinez-Conde. Unsupervised clustering method to detect microsaccades. *Journal of vision*, 14(2):18–18, 2014.

[10] Leandro L Di Stasi, Michael B McCamy, Andrés Catena, Stephen L Macknik, José J Canas, and Susana Martinez-Conde. Microsaccade and drift dynamics reflect mental fatigue. *European Journal of Neuroscience*, 38(3):2389–2398, 2013.

[11] Dan Witzner Hansen and Qiang Ji. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE transactions on pattern analysis and machine intelligence*, 32(3):478–500, 2010.

[12] Kim M Dalton, Brendon M Nacewicz, Tom Johnstone, Hillary S Schaefer, Morton Ann Gernsbacher, Hill H Goldsmith, Andrew L Alexander, and Richard J Davidson. Gaze fixation and the neural circuitry of face processing in autism. *Nature neuroscience*, 8(4):519, 2005.

[13] Jacob L Orquin and Simone Mueller Loose. Attention and choice: A review on eye movements in decision making. *Acta psychologica*, 144(1):190–206, 2013.

[14] Martin Stetter, Raimund A Sendtner, and GEORGE T TIMBERLAKE. A novel method for measuring saccade profiles using the scanning laser ophthalmoscope. *Vision research*, 36(13):1987–1994, 1996.

[15] Andreas Bulling, Jamie A Ward, Hans Gellersen, and Gerhard Troster. Eye movement analysis for activity recognition using electrooculography. *IEEE transactions on pattern analysis and machine intelligence*, 33(4):741–753, 2011.

[16] Ronald S Remmel. An inexpensive eye movement monitor using the scleral search coil technique. *IEEE Transactions on Biomedical Engineering*, (4):388–390, 1984.

[17] Moritz Kassner, William Patera, and Andreas Bulling. Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication*, pages 1151–1160. ACM, 2014.

[18] Enrique Cáceres, Miguel Carrasco, and Sebastián Ríos. Evalu-

ation of an eye-pointer interaction device for human-computer interaction. *Heliyon*, 4(3):e00574, 2018.

[19] Amna Rahman, Mehreen Sirshar, and Aliya Khan. Real time drowsiness detection using eye blink monitoring. In *Software Engineering Conference (NSEC), 2015 National*, pages 1–7. IEEE, 2015.

[20] Emilien Dubois, Colin Blättler, Cyril Camachon, and Christophe Hurter. Eye movements data processing for ab initio military pilot training. In *Intelligent Decision Technologies*, pages 125–135. Springer, 2015.

[21] RG Vishnu Menon, Valdimar Sigurdsson, Nils Magne Larsen, Asle Fagerstrøm, and Gordon R Foxall. Consumer attention to price in social commerce: Eye tracking patterns in retail clothing. *Journal of Business Research*, 69(11):5008–5013, 2016.

[22] Luz Rello and Miguel Ballesteros. Detecting readers with dyslexia using machine learning with eye tracking measures. In *Proceedings of the 12th Web for All Conference*, page 16. ACM, 2015.

[23] Erik Hernández, Santiago Hernández, David Molina, Rafael Acebrón, and Cecilia E García Cena. Oscann: Technical characterization of a novel gaze tracking analyzer. *Sensors*, 18(2):522, 2018.

[24] Juan Biondi, Gerardo Fernandez, Silvia Castro, and Osvaldo Agamennoni. Eye-movement behavior identification for ad diagnosis. *arXiv preprint arXiv:1702.00837*, 2017.

[25] Anuradha Kar and Peter Corcoran. A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms. *IEEE Access*, 5:16495–16519, 2017.

[26] Stefania Cristina and Kenneth P Camilleri. Unobtrusive and pervasive video-based eye-gaze tracking. *Image and Vision Computing*, 2018.

[27] Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling. Appearance-based gaze estimation in the wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4511–4520, 2015.

[28] Kar-Han Tan, David J Kriegman, and Narendra Ahuja. Appearance-based eye gaze estimation. In *Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on*, pages 191–195. IEEE, 2002.

[29] Yasuhiro Ono, Takahiro Okabe, and Yoichi Sato. Gaze estimation from low resolution images. In *Pacific-Rim Symposium on Image and Video Technology*, pages 178–188. Springer, 2006.

[30] Onur Ferhat and Fernando Vilariño. Low cost eye tracking. *Computational intelligence and neuroscience*, 2016:17, 2016.

[31] Zhizhi Guo, Qianxiang Zhou, and Zhongqi Liu. Appearance-based gaze estimation under slight head motion. *Multimedia Tools and Applications*, 76(2):2203–2222, 2017.

[32] Shumeet Baluja and Dean Pomerleau. Non-intrusive gaze tracking using artificial neural networks. In *Advances in Neural Information Processing Systems*, pages 753–760, 1994.

[33] Li-Qun Xu, Dave Machin, and Phil Sheppard. A novel approach to real-time non-intrusive gaze finding. In *BMVC*, pages 1–10, 1998.

[34] Weston Sewell and Oleg Komogortsev. Real-time eye gaze tracking with an unmodified commodity webcam employing a neural network. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, pages 3739–3744. ACM, 2010.

[35] Feng Lu, Yusuke Sugano, Takahiro Okabe, and Yoichi Sato. Inferring human gaze from appearance via adaptive linear regression. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 153–160. IEEE, 2011.

[36] Kar-Han Tan, David J Kriegman, and Narendra Ahuja. Appearance-based eye gaze estimation. In *Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on*, pages 191–195. IEEE, 2002.

[37] Yusuke Sugano, Yasuyuki Matsushita, and Yoichi Sato. Appearance-based gaze estimation using visual saliency. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(2):329–341, 2013.

[38] Ke Liang, Youssef Chahir, Michèle Molina, Charles Tijus, and François Jouen. Appearance-based gaze tracking with spectral clustering and semi-supervised gaussian process regression. In *Proceedings of the 2013 Conference on Eye Tracking South Africa*, pages 17–23. ACM, 2013.

[39] Yusuke Sugano, Yasuyuki Matsushita, and Yoichi Sato. Learning-by-synthesis for appearance-based 3d gaze estimation. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 1821–1828. IEEE, 2014.

[40] Feng Lu and Xiaowu Chen. Person-independent eye gaze prediction from eye images using patch-based features. *Neurocomputing*, 182:10–17, 2016.

[41] Timo Schneider, Boris Schauerte, and Rainer Stiefelhagen. Manifold alignment for person independent appearance-based gaze estimation. In *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pages 1167–1172. IEEE, 2014.

[42] Jixu Chen and Qiang Ji. A probabilistic approach to online eye gaze tracking without explicit personal calibration. *IEEE Transactions on Image Processing*, 24(3):1076–1086, 2015.

[43] Kang Ryoung Park. A real-time gaze position estimation method based on a 3-d eye model. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(1):199–212, 2007.

[44] Elias Daniel Guestrin and Moshe Eizenman. General theory of remote gaze estimation using the pupil center and corneal reflections. *IEEE Transactions on biomedical engineering*, 53(6):1124–1133, 2006.

[45] Elias Daniel Guestrin and Moshe Eizenman. Remote point-of-gaze estimation requiring a single-point calibration for applications with infants. In *Proceedings of the 2008 symposium on Eye tracking research & applications*, pages 267–274. ACM, 2008.

[46] Sheng-Wen Shih and Jin Liu. A novel approach to 3-D gaze tracking using stereo cameras. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(1):234–245, 2004.

[47] Jixu Chen, Yan Tong, Wayne Gray, and Qiang Ji. A robust 3D eye gaze tracking system using noise reduction. In *Proceedings of the 2008 symposium on Eye tracking research & applications*, pages 189–196. ACM, 2008.

[48] Carlos H Morimoto and Marcio RM Mimica. Eye gaze tracking techniques for interactive applications. *Computer vision and image understanding*, 98(1):4–24, 2005.

[49] Jianzhong Wang, Guangyue Zhang, and Jiadong Shi. 2d gaze estimation based on pupil-glint vector using an artificial neural network. *Applied Sciences*, 6(6):174, 2016.

[50] Kyle Krafka, Aditya Khosla, Petr Kellnhofer, Harini Kannan, Suchendra Bhandarkar, Wojciech Matusik, and Antonio Torralba. Eye tracking for everyone. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2176–2184, 2016.

[51] Onur Ferhat and Fernando Vilariño. Low cost eye tracking. *Computational intelligence and neuroscience*, 2016:17, 2016.

[52] Seung-Jin Baek, Kang-A Choi, Chunfei Ma, Young-Hyun Kim, and Sung-Jea Ko. Eyeball model-based iris center localization for visible image-based eye-gaze tracking systems. *IEEE Transactions on Consumer Electronics*, 59(2):415–421, 2013.

[53] Ronda Venkateswarlu et al. Eye gaze estimation from a single image of one eye. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 136–143. IEEE, 2003.

[54] David Young, Hilary Tunley, and Richard Samuels. *Specialised*

sinc(*i*) Research Institute for Signals, Systems and Computational Intelligence (fich.unl.edu.ar/sinc)
A. Larrazabal, C. García Cena & C. E. Martínez; "Video-oculography eye tracking towards clinical applications: a review"
Computers in Biology and Medicine, Vol. 108, pp. 57-66, 2019.

hough transform and active contour methods for real-time eye tracking. School of Cognitive and Computing Science, University of Sussex, Cognitive & Computing Science, 1995. CSRP no. 386.

[55] Peng Yang, Bo Du, Shiguang Shan, and Wen Gao. A novel pupil localization method based on GaborEye model and radial symmetry operator. In *Image Processing, 2004. ICIP'04. 2004 International Conference on*, volume 1, pages 67–70. IEEE, 2004.

[56] Anjith George and Aurobinda Routray. Fast and accurate algorithm for eye localization for gaze tracking in low-resolution images. *ET Computer Vision*, 10(7):660–669, 2016.

[57] Jian-Gang Wang, Eric Sung, and Ronda Venkateswarlu. Estimating the eye gaze from one eye. *Computer Vision and Image Understanding*, 98(1):83–103, 2005.

[58] Kyung-Nam Kim and RS Ramakrishna. Vision-based eye-gaze tracking for human computer interface. In *Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on*, volume 2, pages 324–329. IEEE, 1999.

[59] Roberto Valenti and Theo Gevers. Accurate eye center location and tracking using isophote curvature. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.

[60] Dan Witzner Hansen and Arthur EC Pece. Eye tracking in the wild. *Computer Vision and Image Understanding*, 98(1):155–181, 2005.

[61] Dongheng Li, David Winfield, and Derrick J Parkhurst. Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. In *Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on*, pages 79–79. IEEE, 2005.

[62] Thomas E Hutchinson, K Preston White, Worthy N Martin, Kelly C Reichert, and Lisa A Frey. Human-computer interaction using eye-gaze input. *IEEE Transactions on systems, man, and cybernetics*, 19(6):1527–1534, 1989.

[63] Marcio RM Mimica and Carlos Hitoshi Morimoto. A computer vision framework for eye gaze tracking. In *Computer Graphics and Image Processing, 2003. SIBGRAPI 2003. XVI Brazilian Symposium on*, pages 406–412. IEEE, 2003.

[64] Juan J Cerrolaza, Arantxa Villanueva, and Rafael Cabeza. Taxonomic study of polynomial regressions applied to the calibration of video-oculographic systems. In *Proceedings of the 2008 symposium on Eye tracking research & applications*, pages 259–266. ACM, 2008.

[65] Juan J Cerrolaza, Arantxa Villanueva, and Rafael Cabeza. Study of polynomial mapping functions in video-oculography eye trackers. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 19(2):10, 2012.

[66] Zhiwei Zhu and Qiang Ji. Eye and gaze tracking for interactive graphic display. *Machine Vision and Applications*, 15(3):139–148, 2004.

[67] Massimo Gneo, Maurizio Schmid, Silvia Conforto, and Tommaso DAlessio. A free geometry model-independent neural eye-gaze tracking system. *Journal of neuroengineering and rehabilitation*, 9(1):82, 2012.

[68] Marcus Nyström, Richard Andersson, Kenneth Holmqvist, and Joost Van De Weijer. The influence of calibration method and eye physiology on eyetracking data quality. *Behavior research methods*, 45(1):272–288, 2013.

[69] Feng Lu, Yusuke Sugano, Takahiro Okabe, and Yoichi Sato. Gaze estimation from eye appearance: a head pose-free method via eye image synthesis. *IEEE Transactions on Image Processing*, 24(11):3680–3693, 2015.

[70] Feng Lu, Takahiro Okabe, Yusuke Sugano, and Yoichi Sato. Learning gaze biases with head motion for head pose-free gaze estimation. *Image and Vision Computing*, 32(3):169–179, 2014.

[71] Chih-Chuan Lai, Yu-Ting Chen, Kuan-Wen Chen, Shen-Chi Chen, Sheng-Wen Shih, and Yi-Ping Hung. Appearance-based gaze tracking with free head movement. In *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pages 1869–1873. IEEE, 2014.

[72] Kwang Ho An and Myung Jin Chung. 3d head tracking and pose-robust 2d texture map-based face recognition using a simple ellipsoid model. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 307–312. IEEE, 2008.

[73] Jilin Tu, Thomas Huang, and Hai Tao. Accurate head pose tracking in low resolution video. In *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*, pages 573–578. IEEE, 2006.

[74] Roberto Valenti, Nicu Sebe, and Theo Gevers. Combining head pose and eye location information for gaze estimation. *IEEE Transactions on Image Processing*, 21(2):802–815, 2012.

[75] Craig Hennessey and Jacob Fiset. Long range eye tracking: bringing eye tracking into the living room. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, pages 249–252. ACM, 2012.

[76] Dong-Chan Cho, Wah-Seng Yap, HeeKyung Lee, Injae Lee, and Whoi-Yul Kim. Long range eye gaze tracking system for a large screen. *IEEE Transactions on Consumer Electronics*, 58(4), 2012.

[77] Chi Jian-nan, Zhang Chuang, Yan Yan-tao, Liu Yang, and Zhang Han. Eye gaze calculation based on nonlinear polynomial and generalized regression neural network. In *Natural Computation, 2009. ICNC'09. Fifth International Conference on*, volume 3, pages 617–623. IEEE, 2009.

[78] Michael Mrochen, Mostafa Salah Eldine, Maik Kaemmerer, Theo Seiler, and Werner Hütz. Improvement in photorefractive corneal laser surgery results using an active eye-tracking system. *Journal of Cataract & Refractive Surgery*, 27(7):1000–1006, 2001.

[79] Yuan Chieh Lee. Active eye-tracking improves lasik results. *Journal of Refractive Surgery*, 23(6):581–585, 2007.

[80] Kerry D Solomon, Luis E Fernández de Castro, Helga P Sandoval, Joseph M Biber, Brian Groat, Kristiana D Neff, Michelle S Ying, John W French, Eric D Donnenfeld, Richard L Lindstrom, et al. Lasik world literature review: quality of life and patient satisfaction. *Ophthalmology*, 116(4):691–701, 2009.

[81] Mohamed El Bahrawy and Jorge L Alió. Excimer laser 6 th generation: state of the art and refractive surgical outcomes. *Eye and Vision*, 2(1):6, 2015.

[82] Ioannis M Aslanides, Georgia Toliou, Sara Padroni, Samuel Arba Mosquera, and Sai Kolli. The effect of static cyclotorsion compensation on refractive and visual outcomes using the schwind amaris laser platform for the correction of high astigmatism. *Contact Lens and Anterior Eye*, 34(3):114–120, 2011.

[83] Nicolas Noiret, Nicolas Carvalho, Éric Laurent, Gilles Chopard, Mickaël Binetruy, Magali Nicolier, Julie Monnin, Eloi Magnin, and Pierre Vandel. Saccadic eye movements and attentional control in alzheimer's disease. *Archives of Clinical Neuropsychology*, 33(1):1–13, 2017.

[84] Robert J Molitor, Philip C Ko, and Brandon A Ally. Eye movements in alzheimer's disease. *Journal of Alzheimer's disease*, 44(1):1–12, 2015.

[85] Qing Yang, Tao Wang, Ning Su, Shifu Xiao, and Zoi Kapoula. Specific saccade deficits in patients with alzheimers disease at mild to moderate stage and in patients with amnestic mild cognitive impairment. *Age*, 35(4):1287–1298, 2013.

[86] Liam D Kaufman, Jay Pratt, Brian Levine, and Sandra E Black. Executive deficits detected in mild alzheimer's disease using the antisaccade task. *Brain and behavior*, 2(1):15–21, 2012.

[87] Elena Pretegiani and Lance M Optican. eye movements in parkinsons disease and inherited parkinsonian syndromes. *Frontiers in neurology*, 8:592, 2017.

[88] Anshul Srivastava, Ratna Sharma, Sanjay K Sood, Garima Shukla, Vinay Goyal, and Madhuri Behari. Saccadic eye movements in parkinson's disease. *Indian journal of ophthalmology*, 62(5):538, 2014.

[89] Martin Gorges, Hans-Peter Müller, Dorothée Lulé, Elmar H Pinkhardt, Albert C Ludolph, Jan Kassubek, LANDSCAPE Consortium, et al. The association between alterations of eye movement control and cerebral intrinsic functional connectivity in parkinsons disease. *Brain imaging and behavior*, 10(1):79–91, 2016.

[90] Christy K Sheehy, Alexandra Beaudry-Richard, Ethan Bensinger, Jacqueline Theis, and Ari J Green. Methods to assess ocular motor dysfunction in multiple sclerosis. *Journal of Neuro-ophthalmology*, 38(4):488–493, 2018.

[91] Meaghan Clough, Lynette Millist, Nathaniel Lizak, Shin Beh, Teresa C Frohman, Elliot M Frohman, Owen B White, and Joanne Fielding. Ocular motor measures of cognitive dysfunction in multiple sclerosis i: inhibitory control. *Journal of neurology*, 262(5):1130–1137, 2015.

[92] Marisa Ferreira, Paulo A Pereira, Marta Parreira, Inês Sousa, José Figueiredo, João J Cerqueira, and Antonio F Macedo. Using endogenous saccades to characterize fatigue in multiple sclerosis. *Multiple sclerosis and related disorders*, 14:16–22, 2017.