Eye tracking systems:

State of the Art and Promising Applications

Diego Felipe Castro Salvador

Final Thesis Dissertation
MSc Digital and Interaction Design

Supervisor: Prof. Margherita Pillan

Politecnico di Milano
Department of Design

A.Y. 2018/2019

Page intentionally left on blank

introduction

With the rise of developments on both hardware and software in the lasts decades new ways to shape the interaction between humans and technology are necessary. Thus, there is an existing need to re-structure the paradigms of the communication channels that connect users with computer systems. One-way conversations demand to evolve into multi channel experiences considering the possibilities that additional input and output means provide. Nowadays it is possible to comprehend user's context, actions and focus of attention supported by a wide range of receptors, sensors and patterns of use that track signals that humans perform while interacting with a system.

Eye tracking is a technology that has acquired a tremendous evolution in its development, to the point that it is not more confined to laboratories or expensive marketing researches but currently it is an affordable, non-intrusive and proliferated technology that could be taken advantage of by every user in their daily lives. Being able to track users' gaze in real time introduces a massive opportunity for designers and developers to envision and generate breakthrough innovations that support the way people communicate and interact in a natural, useful and usable way.

Therefore, the major challenge involving eye tracking technology is not anymore the approach to the tools and the complexity in its operation but it consists in exploring novel interaction techniques that integrate eye movements into systems that enhance user-to-computer communication.

Nevertheless, the fact that the eye is one of the most complex human senses represents lots of considerations for incorporating gaze movements as part of an interaction frame. The particularities of the eye physiology, its movements and its possible flaws needs to be analysed and evaluated in order to extract the potential advantages from gaze interaction and certainly approach efficiently to the interaction issues. For that reason, diverse categories of gaze-based applications have been suggested by researchers according to the characteristics of the system, its level of “consciousness” and the general objective they approach.

The aim of this work is to present a state-of-the-art of gaze-based technology and the applications that have been developed. Considering both advantages and disadvantages of varied case studies, the intention of this thesis is to serve as foundation for interaction designers who are exploring the use of eye tracking technologies into their projects, so further gaze-based applications can efficiently exploit this technology’s potential and the implementation issues could be approached and contemplated in a better way.

motivation

With the rise of developments on both hardware and software in the lasts decades new ways to shape the interaction between humans and technology are necessary. Thus, there is an existing need to re-structure the paradigms of the communication channels that connect users with computer systems. One-way conversations demand to evolve into multi channel experiences considering the possibilities that additional input and output means provide. Nowadays it is possible to comprehend user's context, actions and focus of attention supported by a wide range of receptors, sensors and patterns of use that track signals that humans perform while interacting with a system.

Eye tracking is a technology that has acquired a tremendous evolution in its development, to the point that it is not more confined to laboratories or expensive marketing researches but currently it is an affordable, non-intrusive and proliferated technology that could be taken advantage of by every user in their daily lives. Being able to track users' gaze in real time introduces a massive opportunity for designers and developers to envision and generate breakthrough innovations that support the way people communicate and interact in a natural, useful and usable way.

Table of contents

CHAPTER 1

Introduction to Eye tracking

What is eye tracking, application fields and technology context.

EYE TRACKING REFERS to the process of tracking eye movements or the absolute point of gaze (POG)- referring to the point the user's gaze is focused at in the visual scene (Fairclough & Gilleade (eds.) 2014). This information is generally translated into computer coordinates (Majaranta & Donegan 2011) that can be used in several application areas such as psychology, marketing and human-computer interaction. With eye tracking, it is possible detect where users look at a point in time, how long they look at something, and the path their eyes follow (Bergstrom & Schall 2014).

In the Human-Computer interaction field, eye-tracking importance has increase in the last decades as consequence of technological developments that makes the devices for measurement of real time eye movements easier and cheaper to use (Blascheck et al. 2014). There is a considerable body of research using eye tracking, but it has concentrated on eye movement data as a tool for studying motor and cognitive processes by recording the eye movements and subsequently analysing them. Nevertheless, there is a second category of eye tracking applications that employs the user's gaze as an input medium into the user-computer dialogue. An example of a system relying solely on gaze as input has been shown to be an important communication tool for quadriplegics users, where the eyes are used for positioning a cursor over an oversized, projected keyboard (Duchowski 2017).

Eye tracker is frequently the device used for measuring eye movements. The development of such devices started in the late 1800s, implementing invasive and uncomfortable practices involving direct mechanical contact with the cornea of the observer. In the early 1900s, the creation of the first "non-invasive" eye tracking device had a dramatic impact in the measurement of eye movements (Wade & Tatler 2005) using applied motion picture photography in order to record the temporal aspects of gaze direction in two dimensions (Jacob & Karn 2003). Between 1930 and 1950 several studies related with the tracking of the movement eyes were applied, considering a wide range of activities such as reading or controlling airplane instruments.

Hartridge and Thompson invented the first head-mounted eye tracker in 1948, making possible to perform studies focusing completely on eye position without constraining on the head movement. During the next couple of decades, efforts were focused in the adaptation of the devices to include less obtrusive and more accurate equipment. However was in 1970s when the eye tracking research flourished due to both advances in eye tracking technology and psychological theory that allowed researches to link the eye tracking data collected with perceptual and cognitive processes. The discovery that multiple reflections from the eye could be used to dissociate eye rotations from head movement (Cornsweet & Crane 1973) was adopted by military/industry development teams to improve the automation of eye tracking data analysis. During this decade psychologists studied the relationship between eye movements and simple visual stimulus properties such as target movement, contrast, and location. As engineers continued to improve eye tracking technology, psychologists began to study the relationships between fixations and cognitive factors including learning, memory, workload and deployment of attention.

With the proliferation of computers in the 1980s, researchers started to investigate how the practicing of eye tracking could be possibly applied in the human-computer interaction field. Real time information provided by more advance devices permitted to solve questions regarding on how users, for example, search for commands in computer menus (Card 1984) or even more complex tasks including a flight simulator for pilots (Tong & Fisher 1984).

Most modern devices rely on a method called corneal reflection to detect and track the location of the eyes as it moves (Bergstrom & Schall 2014). As a result of this significant developments on both hardware and software in the last decades, eye tracking system applications in the Human-Computer Interaction field has shown notable growth in two modalities (Duchowski 2017): Diagnostic, described as a mean to measure the usability of computer interfaces (Benel, Ottens & Horst 1991) and Interactive, where gaze position becomes a real time input medium (Zhai, Morimoto & Ihde 1999).

There is a lot of potential using eye movements in human-computer interfaces due to gaze is an indicator of where current visual attention is directed. Humans usually point the eyes in an object before performing an action over it, which means that tracking the position of the gaze could provide information about the user's goals in a natural way (Majaranta & Bulling 2014). It is relevant to mention that nowadays eye tracking devices are easy to operate, meaning that no training or particular coordination is required of normal users for them to be able to cause their eyes to look at an object; and the control-to-display relationship for this device is already established in the brain (Jacob 2003). Furthermore, due to the particular nature of the muscles involved in control the eye movement, little detectable fatigue or in occasions no-fatigue pointing is possible to be performed (Saito 1992).

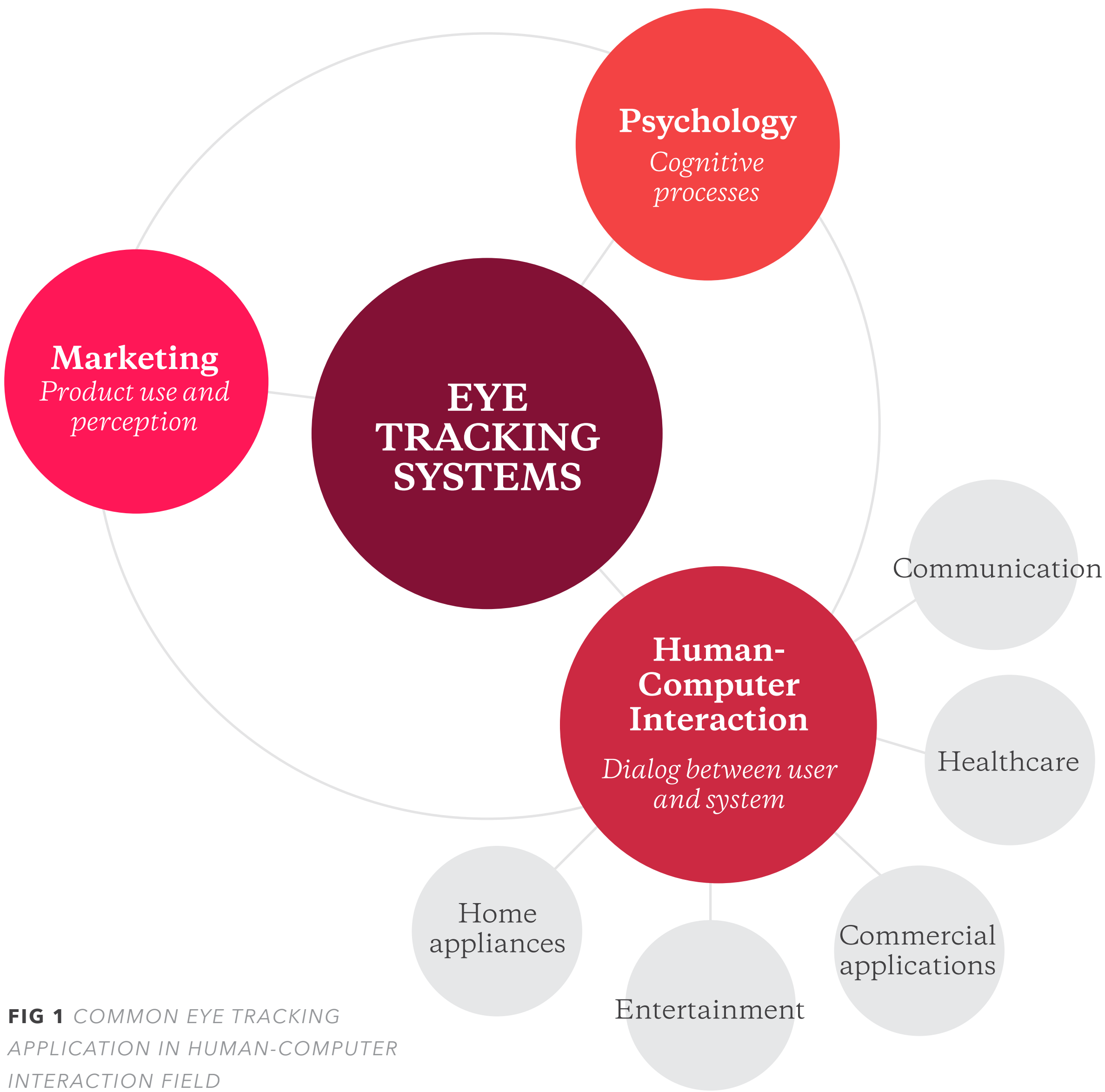

FIG 1 *COMMON EYE TRACKING APPLICATION IN HUMAN-COMPUTER INTERACTION FIELD*

CHAPTER 2

Eye physiology, brain and vision

How the eyes work and what is their relation with cognitive processes and attention.

IT IS ESSENTIAL TO UNDERSTAND the processes that the eyes and the brain perform in order to register and perceive the environment as a cogent and stable whole. For that reason it is relevant to make a review about topics such as the physiological structure of the eye, the exchange of information between the eye and the brain and the different type of eye movements and its relation with how visual attention and cognition is originated and triggered.

The path that light energy passes through the eye, to then be transformed into electrical and chemical signals, starts in the cornea. The cornea is a transparent membrane placed around the eyeball and its function is to keep the shape of the eye, protect it from dust and other possible irritants. Its surface is highly sensitive to touch and pain, so to the slightest irritation caused by any cause, its self-cleaning system is activated by blinking the lids and secreting tears from the glands. In order to stay transparent, the cornea needs to regulate its hydration level. The layer responsible of this task is called corneal endothelium and it works by regulating the flow of liquid between the cornea and the iris.

After passing though the cornea, light encounters the aqueous humour, which purpose is to keep the correct amount of pressure in the cornea to maintain its shape. Then it is placed the iris, the part of the eye that gives it colour. The central aperture of the iris is called pupil, and its function is to regulate the amount of light by contracting or expanding using two sets of muscles. A change in the size of the pupil leads to changes in the retinal luminance and the depth of focus of the optical system (Hubel 1988).

Once the light enters the pupil, it passes through the lens. The lens' composition is based on water and crystalline proteins. The eye focuses objects at varying distances by changing the shape of the lens using a set of muscles called ciliary muscles. This movement is usually known as "accommodation" and is the responsible to generate crispy and focused images. Generally in adults over 45 years old, the lens starts to loose its ability to contract or expand, causing difficulties in sight focusing. One of the main functions of the lens is to project light exactly on the back of the eye where the retina is situated.

The gel that fills the central cavity of the eyeball is called vitreous humour. It provides the eyeball with the enough amount of pressure to get its spherical shape. If the pressure is abnormal, the performance of the ciliary muscles is compromised which can cause inconvenience while trying to focus objects. So far, all the parts of the eyeball mentioned take part of the dioptric apparatus, in which relies the formation of images on the inner surface of the eye: the retina.

The human retina contains photoreceptive cells that translate light into nerve signals. These cells are known as rods and cones, and they respond to different kind of light properties. Rods are sensitive to changes in brightness and movement, while cones are sensitive to colour. There are approximately 6.4 millions of cones in the retina, and between 110 and 125 millions of rods. The retina is formed by three layers of nerve-cell bodies separated by two layers; a photosensitive area that faces the vascular membrane, and a neutral area that faces the vitreous humour. The part of the retina with the highest resolving power is called fovea centralis. It has a diameter of 0.4mm, and it lies at the macula lutea, a yellowish layer occupied mainly by cones.

The first layer the light encounters at the retina contains nerve fibres. The middle layer of the retina contains three types of nerve cells: bipolar cells, horizontal cells and amacrine cells. The bipolar cells are positioned in a key place within the retina so every signal originated in the receptors must pass through them in order to arrive to the ganglion cells. Horizontal cells links receptors and bipolar cells by relatively long connections that run parallel to the retinal layers; similarly, amacrine cells link bipolar cells and retinal ganglion cells (Hubel 1988).

Morphologists have distinguished 10 layers in the retina: 1, The pigmented epithelium; 2, the layer of outer and inner segments or rods and cones; 3, the outer limiting membrane, intersected by the rods and cones; 4, the outer nuclear layer containing the nuclei and fibres of the rods and cones; 5, the outer plexiform layer, formed by a plexus of endings of the photoreceptors with the fibres of neurons of the next layer; 6, the inner nuclear layer of bipolar cells, horizontal and amacrine; 7, the inner plexiform layer, consisting of a plexus of endings of the neurons of

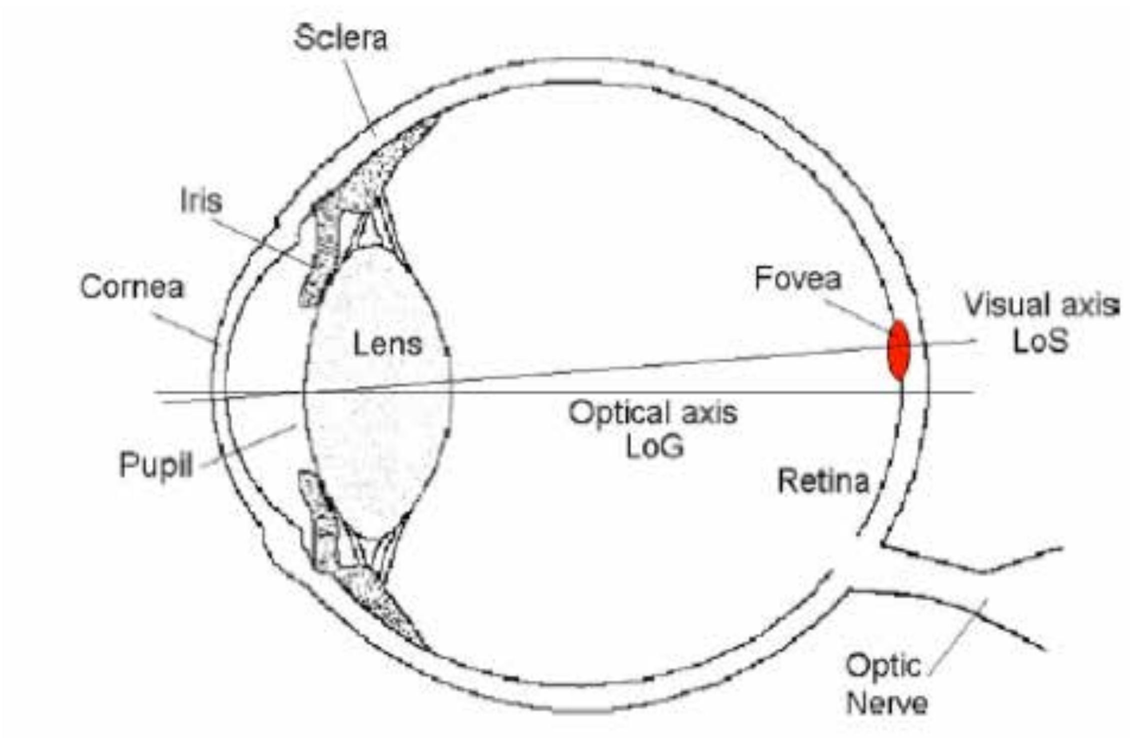

FIG 2 *STRUCTURE OF THE HUMAN EYE. TAKEN FROM MORIMOTO AND MIMICA*, 2005.

layer 6 with endings of the ganglion cells; 8, the layer of ganglion cells; 9, the layer of fibres of the optic nerve; and 10, the inner limiting membrane. The bipolar cells are of several types, differing in morphological structure and in mode of communication with other neurons. The same is true of the ganglion cells. The final centers of vision are the occipital lobes of the cerebral cortex, on both lips of the calcimine fissure (Yarbus 1967).

After light has passed through the different layers of the retina and it has been transformed into electrical signals, it encounters a part of the optic pathway called the optic nerve. This nerve contains approximately 1 million nerve fibres and its length is about 5 centimetres. Then, light signals are transported to the lateral geniculate nucleous (LGN), which is found in the thalamus, the part of the brain responsible to rely sensory information taken from the environment, send movement signals on to the cerebral cortex and to regulate consciousness and sleep. From the primary visual cortex the signal passes through two pathways of vision in humans, the dorsal and ventral visual streams (Mulvey 2012). The dorsal stream is known to be involved with movement and spatial relationships, while the ventral stream is associated with object identification (Mulvey 2012). A wider explanation about these streams will be done when attentive and cognitive processes will be reviewed.

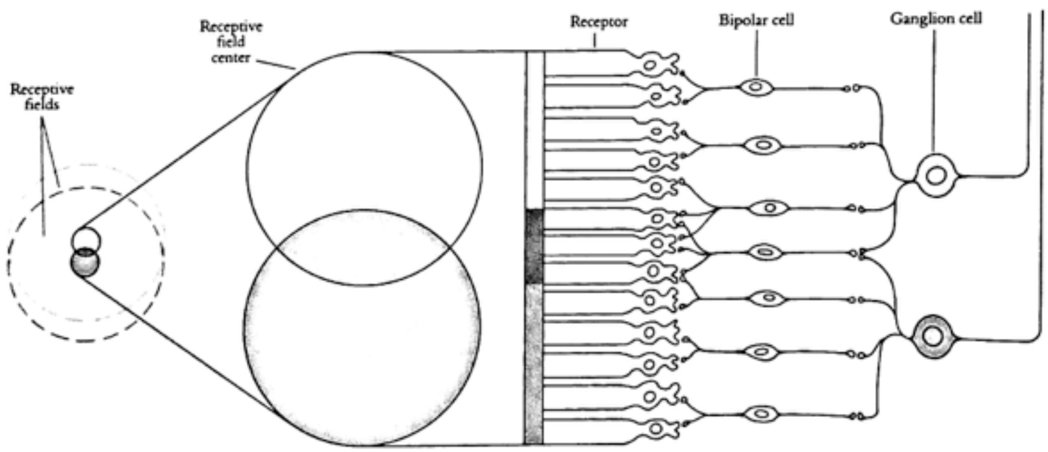

FIG 3 *RELATION BETWEEN RECEPTIVE FIELDS, RECEPTORS, BIPOLAR CELLS AND GAGLION CELLS.* TAKEN FROM HUBEL, 1988.

2.1 Eye movements

Due to the distribution of photoreceptors in the inner structure of the human vision system, people must perform a series of continuous eye movements in order to align the high-resolution area of the perceived field with the objects that are attempted to be recognized. Depending on current goals and current environmental events, the eyes will perform movements in response to either "bottom-up" (stimulus driven) or "top-down" (goal driven) (Mulvey & Heubner 2012). These movements are possible thought the contraction and relaxation of six extraocular muscles that are attached around the eyeball. The medial and lateral recti are responsible for sideways movements, the superior and inferior recti make possible to perform up/down movements and finally the twist movements are directly linked with the superior and inferior obliques. Any eye movement is directed by the brain and they are made by contracting one muscle and relaxing the opponent by just the same amount (Hubel 1988).

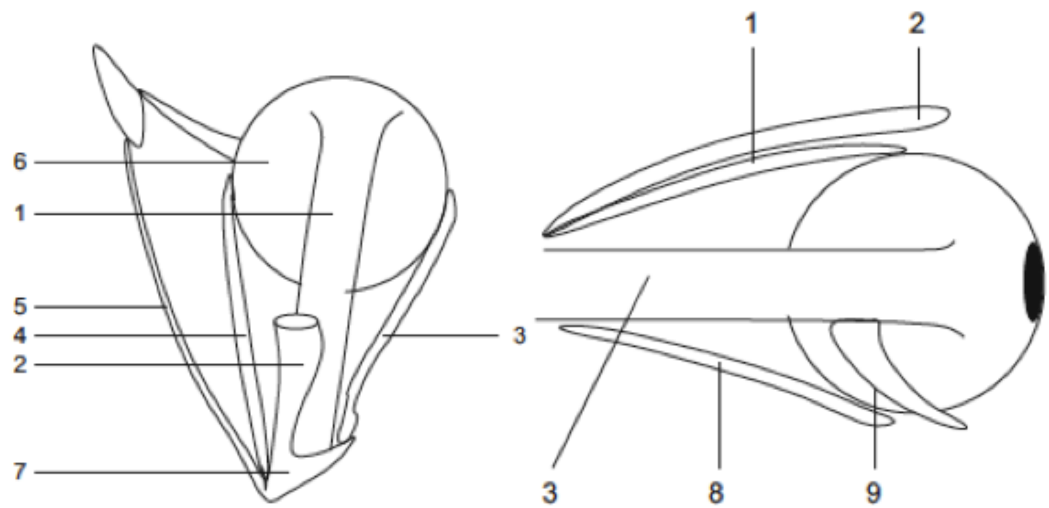

FIG 4 *EXTRINSTIC MUSCLES OF THE EYE. LEFT: 1, SUPERIOR RECTUS; 2, LEVATOR PALBEBRAC SUPERIORIS; 3, LATERAL RECTUS; 4, MEDIAL RECTUS; 5, SUPERIOR OBLIQUE; 6, REFLECTED TENDON OF THE SUPERIOR OBLIQUE; 7, ANNULUS OF ZINN. RIGHT: 8, INFERIOR RECTUS; 9, INFERIOR OBLIQUE.* TAKEN FROM DUCHOWSKI, 2017

Inside the human vision system, there are around 120 million rods and cones versus 1 million ganglion cells. This unbalance is one the cause of the relative small size of the high-resolution visual area, better known as the fovea. The connections between photoreceptive cells and ganglion cells are stronger in the centre of the fovea, but they become exponentially weaker in the periphery. In the fovea area, there are around 147,000 cones/mm2 and a slightly number of rods. At about 10º the number of cones drops sharply to less than 20,000 cones/mm2 and at 30º the number of rods in the periphery drops to about 100,000 rods/mm2. That is the reason why the visual field as a very poor resolution at its periphery. To provide or example, at 5º the acuity of the vision is only 50% (Duchowski 2017).

There are several differences of the information perceived by the different areas of the perceptive field. The fovea and the periphery have diverse properties that establish important distinctions between the signals produced by them and the translation that the brain does about these signals. Duchowski (Duchowski 2017) makes a relevant observation about this topic, addressing three human visual system aspects.

On the first, spatial vision, he explains that the entire visual field corresponds to approximately 23,400 square degree area defined by an ellipsoid with the horizontal major axis subtending 180º visual angle, and the minor vertical axis subtending 130º. The diameter of the highest resolution area is 2º, the parafovea zone reaches between 4º or 5º and the commonly called "useful" visual field extends to about 30º. The rest of the visual field has very low acuity and is principally used for perception of the environment.

> **"Although visual acuity correlates fairly well with cone distribution density, it is important to note that synaptic organization and later neural elements (e.g., ganglion cells concentrated in the central retina) are also contributing factors in determining visual acuity"**
>
> - DUCHOWSKI, 2017

The second aspect about the human visual system Duchowsky analyses is temporal vision. In this section he point out two distinct facts that response to motion: persistence of vision and the phi phenomenon and how these two facts are used in television, cinema and other visual media to produce perception of motion from successively displayed still images. Persistence of vision elucidates the inability of the retina to recognize fast changes in intensities while phi phenomenon (also known as stroboscopic motion) illustrates the apparent motion concept when the brain fills the missing information that does not exist between the perception of stationary objects. That is the reason why is impossible for the human eye to recognize the flicking in still images during the projection of a video when frames are shown at 24 frames per second or faster depending on the quality.

There is a relevant and even surprising concept when comparing motion sensitivity between the foveal and the peripheral region of the visual system. Despite the speed of a moving object seems to be slower in the periphery due to its lower quality perception capacity "...motion detection is the periphery's major task; it is a kind of early warning system for moving targets entering the visual field".

The third aspect analysed is the colour vision. The human retina's capability to recognize colours is given by the three types of cone photoreceptors that react to different peaks of light wavelengths, each one corresponding to one colour: blue, green and red. This is how the human retina efficiently recognizes around 7 million colours, and on contrary ignores, for example infrared and ultraviolet signals. Deficiency on the performance of wavelengths' recognition is the common cause for colour-blindness. Regarding the differences in this aspect between foveal and peripheral regions, colour perception is stronger in the former one as consequence of cones distribution inside the retina, especially with red and green wavelengths.

2.2. What do eye trackers measure?

In order to perform usability evaluations or develop interactive gaze-based systems, it is important for the researchers to be familiar with basic knowledge about eye trackers and eye movements. Here below, a list with the most common eye movements tracked in usability evaluations are descripted.

Fixations: To see and object in the real world, humans have to fixate their gaze at it long enough for the brain's visual system to perceived it. Fixations are often defined as pauses of at least 100ms, typically between 200 and 600ms. During any fixation, it's only possible to see a fairly narrow area of the visual scene with high acuity. The size of the high-acuity field of vision, the fovea, subtends at and angle for about one degree from the eye, but it decreases rapidly towards the periphery of the eye (Majaranta & Bulling 2014).

Fixations can be interpreted in different ways depending on the context. In an encoding task, higher fixation frequency on a particular area could possibly mean great interest of the observer in the target but also can be indicative of difficulty in recognising it (Jacob & Karn 2003; Just & Carpenter 1976). Common metrics for fixations are the fixation count (number of fixations), the fixation duration in milliseconds, and the fixation position given as x- and y-coordinates in pixel (Blascheck et al. 2014).

Saccades: The principal method for moving the fovea to view a different portion of the visual scene is a sudden and rapid motion called saccade. These movements take approximately 30-120 milliseconds and traverse a range between 1 and 40 degrees of visual angle (15-20 degrees being most typical). Besides being consider as the fastest movement the human body can perform, saccades are ballistic, which means that once started, their trajectory and destination cannot be altered (Jacob 2003). Backtracking the eye movements can act as a measure of processing difficulty during reading tasks, as well as they could demonstrate confusion in higher-level processing of the text (Rayner & Pollatsek 1989).

Typical metrics are the saccading amplitude (i.e. the distance the saccade travelled), the saccadic duration in milliseconds, the saccadic velocity in degrees per second (Blascheck et al. 2014) and saccades revealing marked directional shifts (Cowen, Ball & Delin 2002).

Smooth pursuit: This kind of movements are performed by the eyes when they follow a moving object by matching the speed and direction of the muscle movement to the speed and direction of the moving stimulus (Mulvey 2012). Smoot pursuit movements are generated involuntary, and may be independently superimposed upon saccades to stabilize moving targets on the retina (Goldberg & Wichansky 2002). The velocity of the eye while performing smooth pursuits is about 10 to 30 degrees per second (Holmqvist et al. 2011).

Scanpath: It is the sequence of alternating fixations and saccades, which can give information about the search behaviour of a participant. An ideal scanpath is viewed as being a straight line to a desired target, with relatively short fixation duration at the target would be a straight line to a specific target (Goldberg & Kotval 1999). Its metrics include the convex hull (i.e. which area is covered by the scanpath), scanpath length in pixels, or scanpath duration in milliseconds (Blascheck et al. 2014).

Area of Interest (AOI): Also known as Region of Interest (ROI) are those parts of an interface in which a defined class of information is placed. Thus, several fixation points can be found in one area of interest (Zülch & Stowasser 2003). Typical metrics for AOIs are the transition count, the dwell time within an AOI in milliseconds, and the AOI hit which defines if a fixation is within a AOI or not (Blascheck et al. 2014).

2.3. Visual attention

Everyday, a humongous amount of information is presented in the environment and this information is perceived by the sensorial systems and analysed by the brain. Stimulus such as images, smells, textures, among others, are being constantly processed in order to create a logical and stable representation of the outer world. However, due to the immense quantity of information available and the limitation of the human sensory channels, selective-filter tasks must be performed so the brain can use its mental capacities to process the stimulus of interest and examine the details in the environment. The faculty implied as the filter is attention (Duchowski 2017).

William James provided a very well accepted definition of attention: *"Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others…when the things are apprehended by the senses, the number of them that can be attended to at once is small, 'Pluribus intentus, minor est ad singula sensus.' (Many filtered into few for perception.)"*

Given that outer stimulus are in largely part visual information, the human visual system approaches the filter task by focusing the attention and performing saccades. Intuitively it can be thought that these two aspects are related, which is partly true. However, some studies demonstrate that there are situations where attention may move while the eyes remain fixated. An overt attention shift is performed when the observer moves the eyes, the head or the complete body in order to align the fovea with a new object of interest. A covert visual attention shift is performed when the focus of attention changes its position inside the peripheral visual field, but the eye is not align with it. Posner (1980) proposed that before a saccade can be programmed, attention must be oriented towards the destination. This means that in order the perform any eye movements, covert and overt attention shifts are always triggered, since the observer first needs to focus the attention in the area of interest to then execute the saccade. It is assumed that a separate go-signal is required from the programing of the saccade to its execution.

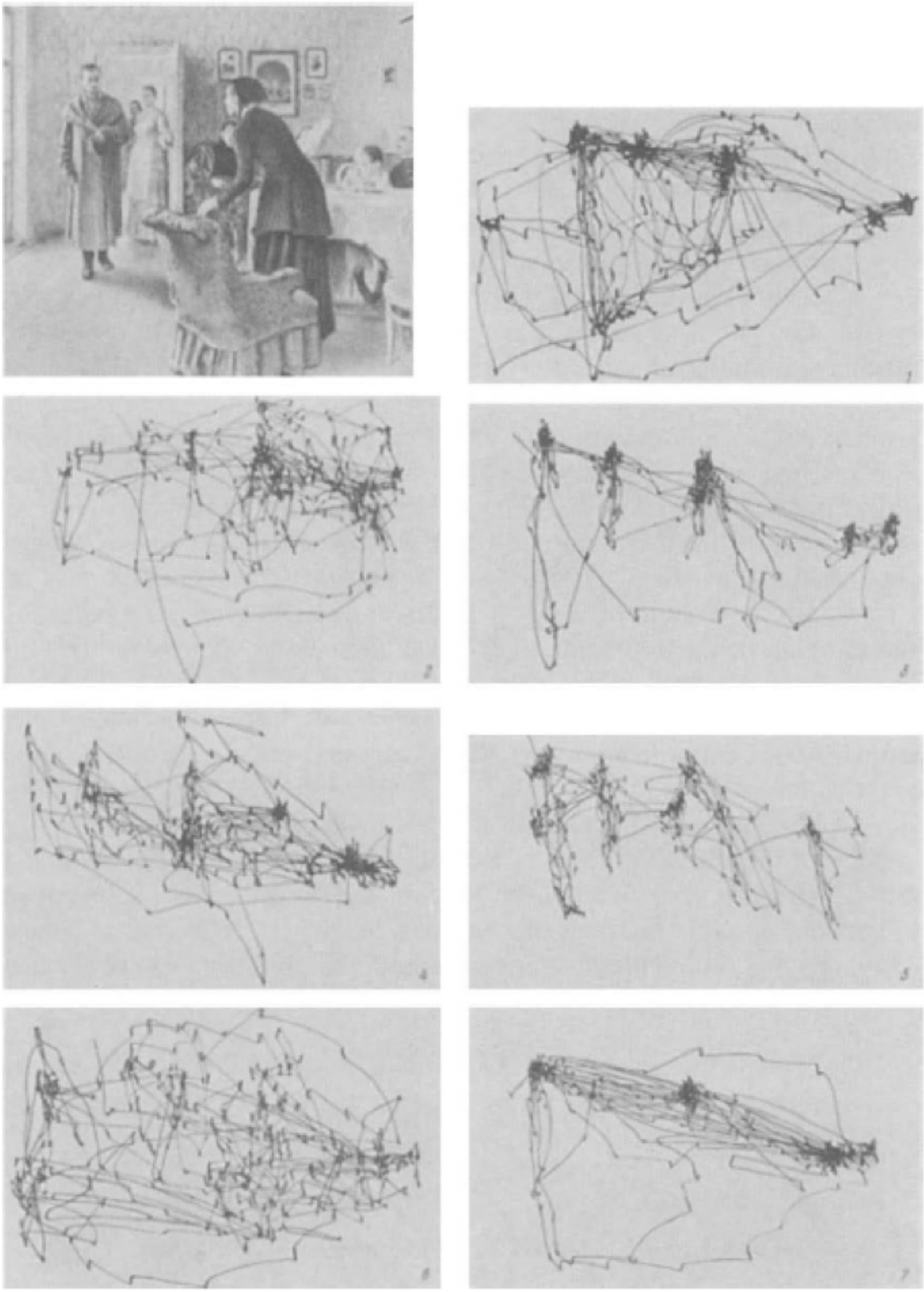

FIG 5 *EACH PATTERN REPRESENT THE MOVEMENTS PERFORMED BY OBSERVERS AFTER BEING ASKED DIVERSE QUESTIONS ABOUT THE SCENE.* TAKEN FROM YARBUS, 1976.

There are two types of eye movements that observers perform according to the origin and the relevance of the area or object of interest presented in the visual field. Exogenous eye movements can be controlled by stimulus properties, irrespective of the goals of the observer. Also can be known as stimulus-driven, involuntary or bottom-up. On the other side, Endogenous eye movements may be controlled by goals and expectations. An example could be searching for a specific target object the observer will tend to select objects that share one or more feature with this target. These movements are also known as goal-directed, voluntary or top-down.

While observing an image stimulus, the observer may consider specific areas in the image that will attract his or her attention. At first, these areas may be perceived by the peripheral vision, to then request for the alignment of the fovea to perform a more detailed inspection. When this selection task is performed over a visual field, several stimulus that present certain level of similarity with the searched target, create an activation map.

The activation map is used to guide shifts of attention and execute a selection process allowing the target to pop up. However it is important to mention that attention selects features from a master map of locations shows where all the feature boundaries are located, but not what those features are. That is, the master map specifies where things are, but not what they are (Shen et al. 2003).

The fact that eye trackers currently are able to interpret the observer's gaze position relatively in real time represents an opportunity to measure internal or subjective brain events. Therefore, analysing this information it could be possible to determine potential trace of user's cognitive states, such as level of mental load or emotional condition. Mulvey & Heubner (Mulvey & Heubner 2012) proposed: *"The aim is to identify and explain those measures from cognitive psychology which are most promising in terms of future technologies for gaze based human computer interaction"*. Nevertheless, eye tracking systems face an important challenge since, as mentioned above, the point of regard is not always aligned with the object or area that the observer is focusing the attention, meaning that eye trackers could be missing an important fragment of information regarding user's attention target.

CHAPTER 3

Eye tracking techniques

Different methods to monitor, measure and analyze users' point of regard.

SEVERAL METHODS TO TRACK EYE MOVEMENTS have been created responding to hardware and software developments during the last one hundred years. Many traditional techniques for eye gaze tracking are intrusive, which means that they required some equipment to be put in physical contact with the user. These techniques include, for example, contact lenses, electrodes, and head mounted devices. Non-intrusive techniques (or remote techniques) are mostly vision based, so they use cameras to capture images of the eye. Some camera-based techniques might be considered intrusive if they require to be head mounted (Morimoto & Mimica 2005). Intrusive eye gaze trackers usually provide higher accuracy than remote trackers, although they could produce discomfort in users specially after long periods of operation.

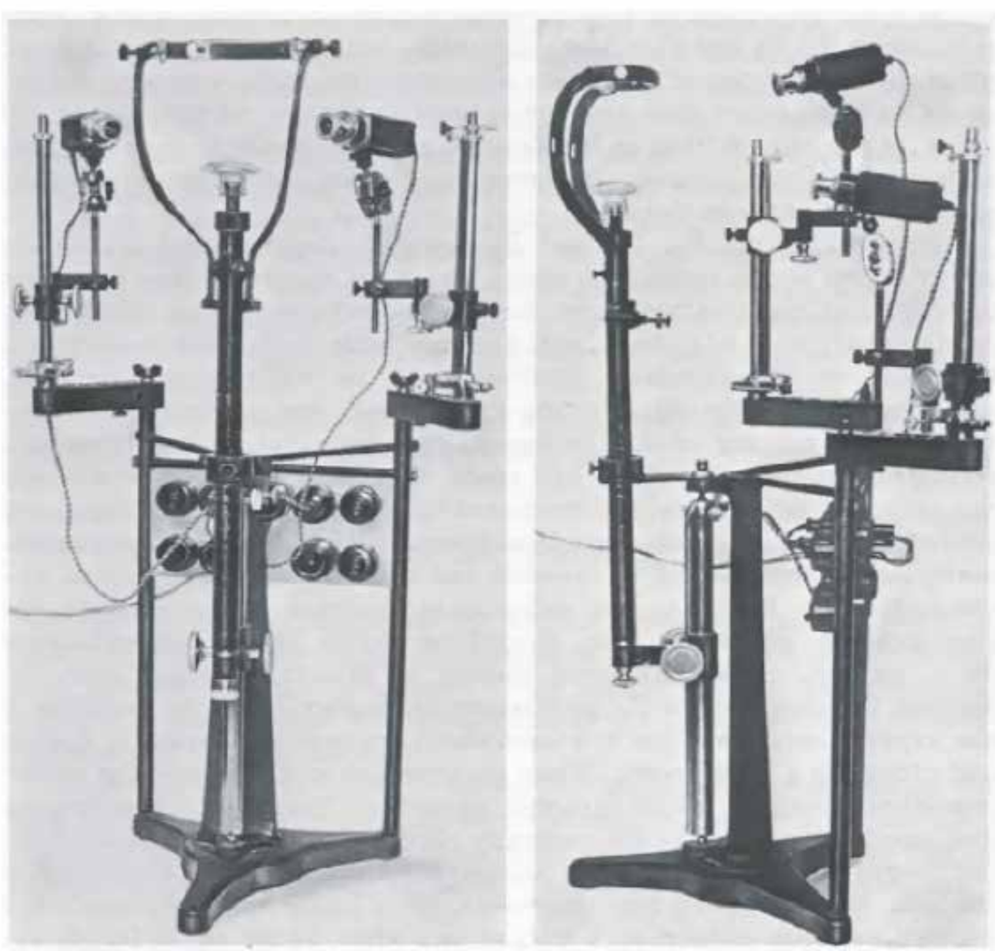

FIG 6 *EARLY APPARATUS FOR EYE TRACKING.* TAKEN FROM YARBUS. 1967.

There are three techniques that have emerged as the predominant ones and are commonly used in research and commercial applications nowadays: (1) videooculography (VOG), video based tracking using head-mounted or remote visible light video cameras, (2) video-based infrared (IR) pupil-corneal reflection (PCR), and (3) Electrooculograpy (EOG). While particularly the first two video-based techniques have a lot or properties in common, all techniques have application areas where they are most useful (Majaranta & Bulling 2014).

3.1. Electrooculography tracking

Using sensors attached at the skin around the eyes it's possible to measure the steady electric field that eyes create when rotating. This electrical signal is known as electrooculogram (EOG). The signal is measured between two pairs of surface electrodes placed around the eye with respect to a reference electrode typically placed on the forehead. Every time the eyes move from the centre position towards one of these electrodes, the retina approaches this electrode, while the cornea approaches the opposing one, this change in dipole orientation causes a change in the electric potential field, with in turn can be measured to track eye movements (Majaranta & Bulling 2014).

FIG 7 *EXAMPLE OF ELECTRO-OCOULOGRAPHY EYE MOVEMENT MEASUREMENT.* TAKEN FROM HTTP://WWW.METROVISION.FR.

This technique is not well-suited for everyday use, since it requires the close contact of electrodes to the user but is still frequently used by clinicians. However, it is a cheap, easy and invasive method of recording large eye movements. The big advantage of this method is its ability to detect eye movements even when the eye is closed (Chennamma & Yuan 2013).

3.2. Infrared Pupil-Corneal Reflection tracking

This eye tracking method uses the reflection of infrared light originated in the scelera and measures its intensity as a reference point. Given that infrared light is not visible by the human eye and does not affect pupil dilation, it is a method that does not distract the user during the observation. Gaze direction is calculated by measuring the changing relationship between the moving pupil centre of the eye and the corneal reflection. As the position of the corneal reflection remains roughly constant during eye movement, the reflection will remain static during rotation of the eye and changes in gaze direction, thus giving a basic eye and head position reference.

Within usability laboratory, diffuse lighting will generally provide fewer problems that pinpoint lighting sources that light cause additional light glints on the cornea. The ability to switch between bright and dark pupil methods can also be helpful in usability testing situations (Goldberg & Wichansky 2002).

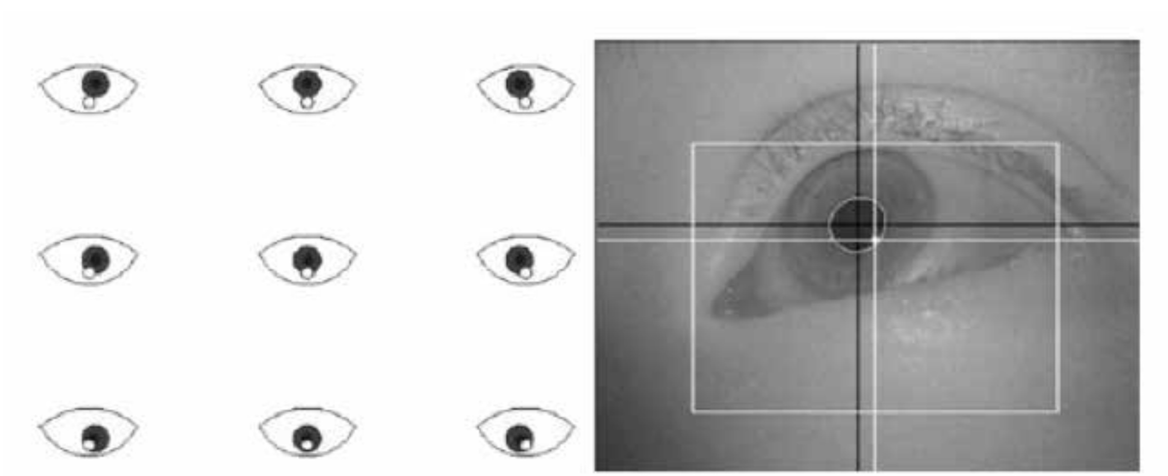

FIG 8 *RELATIONSHIP BETWEEN THE PUPIL CENTER AND THE CORNEAL REFLECTION WHEN THE USER FIXATES ON DIFFERENT LOCATIONS.* ADAPTED FROM MAJARANTA ET AT 2012.

3.3. Video-based tracking

A video-based eye tracking system setup usually consist of a video camera that records the movements of the eye(s) and a computer that saves and analyses the gaze data. In remote systems, the camera is typically based below the computer screen, while in head-mounted systems, the camera is attached either on a frame of eyeglasses or in a separate “helmet”. Head-mounted systems often also include a scene camera for recording the user’s point of view, which can then be used to map the user’s gaze to the current visual scene (Majaranta & Bulling 2014).

According to the complexity of their setup, video-based tracking system can be divided in two groups: single-camera eye tracker and multi-camera eye tracker (Chennamma & Yuan 2013). The first group of systems are relatively inexpensive and easy to install and calibrate, but they provide lower quality information compared with the second multi-camera systems, which give higher accurate data as consequence as more complex and expensive setup.

FIG 9 *PARTICIPANT USING A COMPUTER WITH EYE TRACKING SENSORS.*
TAKEN FROM BERGSTROM AND SHALL, 2014.

1880 **First naked-eye observations while reading.**

1908 **Contact lenses with pointer are created.**
Edmond Huey

1937 **Record eye movements using beam lights.**
Guy Thomas Buswell

1967 **Eye movements and vision is published.**
Alfred Yarbus

1980 **Flourishing of Eye tracking for marketing**

1981 **Gaze-orchestrated dynamic windows.**
Richard Bolt

1989 **Interactive applications for eye typing.**
ERICA

1990 **Eye tracking applied to web usability**

2000 **First applications integrated with artificial intelligence**

2011 **First commercial plug and play eye tracker.**
Tobii T60

3.4. Eye gaze tracker usability requirements

As mentioned before eye gaze tracking systems can operate in different ways, using several types of hardware and software. Nevertheless these differences, it is worth to mention the importance of the usability requirements that these systems should include. Morimoto & Mimica (Morimoto & Mimica 2005) proposed the following series of usability characteristic that gaze trackers -both intrusive and remote- should satisfy:

1. be accurate, i.e., precise to minutes of arc;
2. be reliable, i.e., have constant, repetitive behavior;
3. be robust, i.e, should work under different conditions, such as in-doors/outdoors,
for people with glasses and contact lenses, etc;
4. be non-intrusive, i.e., cause no harm or discomfort;
5. allow for free head motion;
6. not require calibration, i.e., instant setup;
7. have real-time response.

Even if the list above establishes concrete attributes for eye gaze trackers, it is difficult to find a technique that satisfies all of them. That is why it is vital to consider the needs and goals of every application/device/test based on eye tracking, so the selection of the technique can be more precise, making the general interaction more efficient and providing confidence to the users involved in the process. Donegan and Oosthuizen (Donegan et al. 2006) suggest three simple principles, framed in the KEE concept: 1) Knowledge-based: founded on what is known of the user's physical and cognitive abilities; 2) End-user focused: designed to meet the end-user's interests and needs and 3) Evolutionary: ready to change in relation to the end-users response to gaze control technology and software provided.

CHAPTER 4

Modalities for eye tracking applications in Human-Computer Interaction

Categorization of gaze-based applications, advantages, issues and case studies.

NEW INPUT METHODS provide users the possibility to broaden the channels used in the dialogue with systems and their interfaces. This means the creation of appropriate and specific interaction techniques supported by the data obtained from users such as their gaze position (Jacob & Karn 2003). For that reason several researchers have studied the value of using eye movements interactively in a user interface during the last decades, revealing promising potential for gaze-based applications but also enunciating several issues that must be considered in the development of any system that explore the human eye movement as an input.

According to the way information is utilized in a given gaze-based application and the level of awareness of the user when using his or her eyes movements to interact with the device, it can be positioned in one (or more) or the four categories that has been proposed to describe different types of interaction systems. Majaranta & Bulling (Majaranta & Bulling 2014) suggested a four-group classification continuum inspired by a scheme introduced by Fairclough (Fairclough 2011) to differentiate types of physiological computing system (Fig 10).

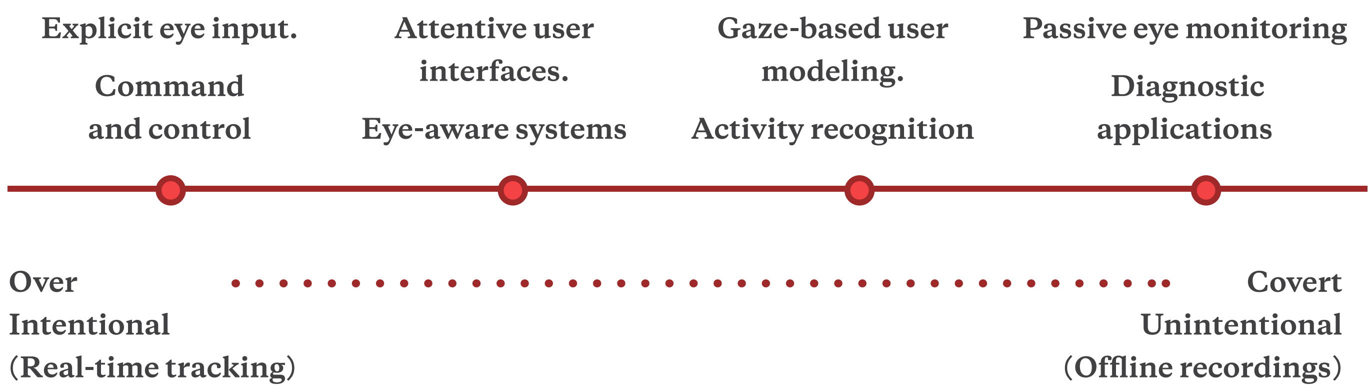

FIG 10 *CONTINUUM OF EYE TRACKING APPLICATIONS FROM INTENTIONAL TO UNINTENTIONAL SYSTEMS.* TAKEN FROM FAIRCLOUGH AND GILLEADE, 2014.

Starting from applications where the intention is explicit, and the eye movements require full, overt and conscious attention from the user; passing through attentive applications that react according to the user behaviour but not require explicit movements from the user; and finishing with systems that passively monitor eye movements, to then analyse user's behaviour. In this chapter, a summary of each modality is elaborated alongside a description of particular implementation issues and potential solutions explained through a series of case studies. However, deeper analysis of the two more overt-oriented categories will be carry out, due to their higher level of interaction with the user.

4.1. Explicit eye input modality

With today's available technology it is possible to capture eye tracking data in real time with higher level of accuracy. This fact opens a window to new developments in the Human-Computer interaction field, aiming to create interaction techniques that can measure and utilize the data provided by eye movements as a medium to improve the communication between users and computers. As Jacob and Karn explained: *"Eye trackers have existed for a number of years, but their use has largely been confined to laboratory experiments. The equipment is gradually becoming sufficiently robust and inexpensive to consider use in real user-computer interfaces. What is now needed is research into the user-computer dialogue in a convenient and natural way"* (Jacob & Karn 2003).

The research for the use of gaze for computer control has been considered since the early 1980s. The idea of *"displays that know where you are looking"* (Bolt 1982) proposed a multi-model approach for interaction with interfaces, where different human input sources such as voice, gestures and gaze position were recognised by a computer as commands. Observations about the eye's point-of-regard as an index, both from daily life and supported by controlled experiments, suggested that a computers system that can detect where on its graphics display the user is looking can use that information to infer the pattern of the user's interest in the things on display and to help formulate the elaboration of that information in both visual and synthesized speech narration (Starker & Bold 1990).

It is important to mention that a significant part of the studies related with eye movements for human-computer interaction have been focused in assistive system for disabled users, who can use only their eyes for input. In fact, COGAIN (Communication by Gaze interaction) is an European research Network of Excellence involving researchers and industries from 20 universities, institutions and companies has been established specifically to develop rapid and natural communication and control systems via eye gaze for users with high-level of motor disabilities such as Amyotrophic Lateral Sclerosis (ALS) or Cerebral Palsy (CP). The aim of COGAIN also includes the identification of user's needs and the production of new low-cost hardware and software that support technologies for motor disabled users (Bates et al. 2006). For some severely disable people, gaze-control represents the only available opportunity to communicate and perform actions, such as change the TV channels, controls the lights in the environment or manipulate the phone. For that reason, the development of accessible interactive systems based on eye tracking technology has rapidly evolved, providing huge emotional, cognitive and physical benefits (Donegan et al. 2009).

The first interactive applications were "eye typing" systems that allowed users to input text by using only the gaze direction. For example, ERICA system (Hutchinson et al. 1989), a project designed and developed by faculty and students in the School of Engineering and Applied Science at the University of Virginia, aimed to help the physically and vocally disabled to attain or regain some measure of independent communication and control. With this system, it was possible for observers to produce text using a grid of six selectable items at a time. Due to the limited accuracy of eye tracking systems in that period, the "selectable" areas of the screen were relatively big, and several keystrokes were required to enter a single letter; but even with those technical issues, ERICA became an important pioneer system for the development of gaze-based assistive applications. Significant improvements were shown in the patients after using the ERICA system, especially in the perception of their overall condition, psychological well being and physical symptoms (Donegan et al. 2009).

The “GazeTalk” system (Hansen et al. 2006), developed in collaboration with a group of Amyotrophic Lateral Scleriosis (ALS) patients, provide them with a solution for communicate with their family just gazing at a particular part of a computer screen with characters or commands. It features type-to-talk, writing, e-mail, web browsing, video playing, music (MP3) playing and smooth reading of PDF files. For text production, the system includes a character prediction function, which predicts the most likely six letters subsequent to the last typed character and predicts the most like next word.

According to Donegan (Donegan 2010) there are essentially three categories of gaze control system: (1) ‘all-in-one’, (2) modular and (3) ‘add-on’. All-in-one systems provide everything the user needs within a single device (Majaranta & Donegan 2011). It means that there are no difficulties in terms of connecting separate pieces of hardware. However, one of the disadvantages is that no updates can be done after the system gets older. Dedicated modular systems are devices that can work using a wide rage of input methods in addition to gaze control (e.g., touch, speech). These devices are worth to consider if the final user needs the possibility of using multimodal input or if gaze control will be needed at a later stage. This is particularly beneficial for users that want to continue using familiar interfaces but they need to change their input mechanism. The add-on systems will enable to transform any computer in an eye tracking system by proving their own interfaces software that offers pointer control. Although this feature appears to be ideal, it is important to consider that, in some cases, there are some inaccuracy and usability issues with pointer control using eye tracking. New devices are being developed and designed by companies as a respond to the increase on the demand for gaze-based assistive devices systems.

EASE (Eye Assisted Selection and Entry) was a system designed and implemented for Chinese text input using the implicit use of eye input to determinate potential focus of attention (Zhai 2003). Since most Chinese characters are homophonic, when users start to type pinyin (official Chinese phonetic alphabet based on Roman characters), the computer software displays many candidate characters with the same pronunciation, together with identifying numbers. The user must then select one of the candidates by pressing the key with the its associated number and depending on the quantity of candidates the user have to scroll to find the hidden options. Using EASE, the user types pinyin as usual but presses the spacebar as soon as he or she positions the gaze over the intended character. According to the studies made using EASE system, users without practice could complete their typing task as fast as the conventional method, and more importantly, users could select the intended character faster than with the traditional numeric keying method. Also, the studies showed that users' cognitive load was reduced by eliminating their need to look at the numeric keys and by employing their visual attention on the screen more optimally.

Another field where gaze-based interaction has been explored is environmental control of domestic appliances for disabled users. Some systems have been developed with the aim of enhance and extend the abilities of users with restricted motor capacities, in order to manipulate different objects presented in the environment by using the gaze orientation as a trigger. ART (Attention Responsive Technology) system was developed using a head-mounted and a remote eye movement-monitoring device (Corno et al. 2010). Using a camera, environmental devices that user may wish to operate are imaged and stored in a database through an algorithm. In this way the system is able to recognize which object is oriented with the user's gaze, and an interface is offered to the user specifically for that device alone which the user can decide to operate. Once the interface is presented, the user can interact with it by exploiting natural and intuitive gaze-based interactions. Increasing self-esteem in disabled users and reducing their reliance on continuous help is one of the most important reasons why the development of new hardware and software for assistive devices and more effective interaction techniques is crucially needed.

4.1.1 Explicit eye input implementation issues

Although gaze-based interaction with computers has huge potential to be used in a wide range of applications as it was mention above, it is also true that this kind of interaction method presents several opportunities of improvement both in technical features and in human factors.

One of the biggest challenges within eye tracking as an input method is its accuracy and responsiveness (Majaranta & Bulling 2014). This issue is not only caused by the camera resolution in video-based tracking system, but also because of the natural jitter in the eye movement, leading to space problems due to larger size of elements on interfaces (Drewes & Schmidt 2007). Additionally, since the fovea of the eye covers a visual angle of 1º arc of the retina (Bates & Istance 2002) the size of any target present in the interfaces should be adequate to avoid inaccuracy problems and error rate. As consequence, the need for a hierarchically organization of the information using menus and submenu increases, resulting in a slow down interaction (Majaranta & Bulling 2014). During a fixation, a user is generally performing unconscious small, jittery motions even if he or she thinks the gaze is steady. This means that gaze-based interaction systems should be build considering this factor, in order to differentiate the data provided by real and conscious saccades from the high-frequency jitter movements (Jacob 2003).

Another implementation issue with gaze-based interaction systems is related with whether to provide visual feedback to the user and if the case how to do it. The problem is that an eye-following cursor will tend to move around and distract user's attention, and in case of calibration inconsistency, it will cause a feedback loop when the observer try to naturally fix the delayed position of the cursor (Jacob 2003).

Probably one of the most prominent challenges related with the human-computer dialogue through eye tracking devices refers to the fact that people frequently use eye movements to scan the environment and recognize the elements around instead of operating devices using the gaze. Furthermore, eye movements cannot easily be consciously controlled or steered as they are driven by subconscious interests (Yarbus 1967). Thus increasing the error-rate associated with gaze interaction is highly probable by triggering involuntary commands all over the interface. Jacob denominated this feature as "Midas Touch" problem: *"…at first, it is empowering simply to look at what you want and have it happened. Before long, though, it becomes like the Midas Touch. Everywhere you look, another command is activated; you cannot look anywhere without issuing a command."* (Jacob 1990).

Contrarily to other input devices, such a mouse, there is no natural way for the eyes to indicate when to "click" so it is needed to create interaction techniques that allow users to perform actions only when is merely wanted. Gaze-based selections seems to be faster than selections by more commonly means (for instance mouse selections), however when additional tasks are involved such as typing or switch selection, this superiority has not been manifest, and the rate error apparently tends to increase (Hansen et al. 2004). For that reason, some of the methods implemented and tested over the years include the combination of gaze pointing with an additional input modality, such as a press a keyboard, click a mouse, use speech, touch surfaces or performing head movements. In other cases, for systems bases solely on gaze-control, the time that users spend staring to a determinate target –also called "dwell time"– is used to activate the commands in the system.

Besides the challenges described above, possible technical issues can originate in the eye tracking equipment while it is reading the gaze information from the user. Those flaws could be caused by misconnection in the camera system (in the case of video-based tracking system), light changes in the atmosphere, blinking or head movement. Even if the development of newest devices that incorporate more precise sensors and smarter calibration processes are solving these kind of technical issues, more focus in the design of the interaction between users and eye-gazed systems is demanded.

4.1.2 Potential implementation solutions and case studies

Certain techniques have been developed in order to enhance the user interaction with gaze-based systems. As mentioned before, there are several interaction and technical issues that need to be solved within the production of more efficient interactive applications based on eye tracking devices. Specifically related to the issue of small targets displayed in two-dimensional interfaces, researchers have tested different kind of methods including dwell-base selection using semantic and graphical zooming, graphical distortion and the use of experimental cursors for easier selection tasks. At the same time, the combination of gaze input with other interaction modalities, such as speech and facial muscles gestures, have been approached and tested offering interesting results and potential solutions to the 'Midas touch' issue mentioned earlier in this chapter.

Dwell-based selection

The most common method for perform actions within a gaze-control system is the use of a brief delay while the gaze position is located over a specific object or area. This so-called "dwell time", uses a specified amount of time to trigger an activation –sort of a 'click' making an analogy with mouse input– of a given element displayed in an interface. Considering the right amount of dwell time is crucial for the optimal performance of the system. Short dwell time could increase error rate or unintentional activations, while longer dwell time might produce difficulties in selection tasks due to the micro saccades performed by the eye (Jacob 1995). For this reason, one frequent solution is to leave the control of the dwell time to the user, so he or she can select the most natural threshold, usually during the calibration process. Additionally, it worth to consider a dynamic adjustment for dwell time given shorter dwell times are needed when user's familiarity and confidence with the system increases (Skovsgaard, Räihä & Tall 2012).

Silbert and Jacob made another important study regarding eye gaze interaction systems by performing two experiments to compare the object selection technique using the most common input device (mouse) vs. using eye gaze (Silbert & Jacob 2000). They found out that eye gaze interaction technique is faster than selection with a mouse, using an algorithm that made use of knowledge about how the eyes behave and utilizing dwell time (150ms) as selection method. As an indicator of user's gaze attention over an object it reacted by highlighting after the assigned dwell time. In the first experiment, users were asked to select a circle from a three by four grid of circles as quickly as possible while in the second test user's task was to select a letter from a grid of letters enclosed by circles. The subject was told which letter to select by a pre-recorded speech played through an audio speaker. Both of the experiments demonstrate that using person's natural eye gaze as a source of computer input is feasible, and even in some cases, users felt the system was anticipating their commands, almost as if it were reading their minds. Another relevant conclusion was the little conscious effort that eye movements require when working in conjunction with other activities, since people is naturally use to look at objects while performing other actions.

By conducting two experiments, Ware and Mikaelian studied object selection ad target size using an eye tracker system as a computer input (Ware & Mikaelian 1987). In the first experiment, three different selection methods were tested: (1) prolonged fixation or "dwell" time of 400 milliseconds, (2) a designated large rectangular area of the screen as a button and (3) by pressing a physical button that was placed conveniently by the user. Independently of the selection method, select times were below one second, ranking gaze selection faster than selection performed using a mouse. In the second experiment, it was studied how target size can possibly affect response speed and error rates while selection by gaze. To summarize the conclusions of both of the experiments, the authors quoted

> **"Where speed is of the essence, cost is not object, sizes are moderate, and it is important than the hands be reserved for other activities, the eye tracker may be the input device of choice".**
>
> - WARE AND MIKAELIAN, 1987

Zooming small targets

One of the major issues of the gaze-based systems is the difficulty of small targets in selection tasks. In order to improve the accessibility of this kind of targets, different techniques have been developed and tested. The change of the appearance of objects based on their scale, also known as zooming, is the most popular. Diverse methods to accomplish zooming in interfaces include the use of additional windows where elements are enhanced, or the implementation of fisheye effect in the display.

With the objective of enhance the accuracy in the selection of standard-size menu items, Spakov & Miniotas (Špakov & Miniotas 2005) developed a simple but effective technique using dynamic target expansion based on the user's gaze point position. As it can be seen in Figure X, the idea was to expand the area of the selected target of the menu after a certain dwell time, while the other menu items shifted

FIG 11 *DYNAMIC TARGET EXPANSION AND CALIBRATION.* TAKEN FROM ŠPAKOV & MINIOTAS 2005

vertically (up or down, depending on the position). They found that using this technique the selection of the items reached 91% accuracy, in spite of an increase on the selection time. Still, it can be consider as a relevant success rate for eye-based interaction. Another approach for zooming techniques is providing a zoomed image and interaction buttons in a separate window (Skovsgaard, Räitä & Tall 2012). An example of this technique was presented by Instance, Spinner & Howarth (Istance, Spinner & A. 1996). As is illustrated in Fig 11 the system displays an additional panel in the bottom part of the interface where the placement of the zoomed content is shown, while the user is looking around in the normal-size image. This solution supports the accuracy of selection especially for small targets, but due that the interaction has to be done in two instances the efficiency is negatively conditioned in terms of time. In favour of avoid the two-instance selection, direct interaction with zooming elements has been proposed. Skovsgaard, Hansen & Mateo (2008) described a continuous zoom-selection technique, which basically consist in a window centred on the user's point of regard, where the content size is gradually enlarged using a smooth animation. While the two-instance technique sacrifices time, the direct interaction technique present some disadvantages related with accuracy. For that reason, is considered that these two selection techniques could be complementary depending on the particular size of the elements present-ed in the interface and the current activity of the user.

Distortion-based interfaces

Asmore, Duchowski and Shoemaker (2005) proposed one of the most novel approaches related with the accessibility to small target within gaze-based interfaces. Their proposal was based in a gaze-contingent fisheye that allowed to magnify the content at the point of the user's gaze. One of the interesting characteristics of the

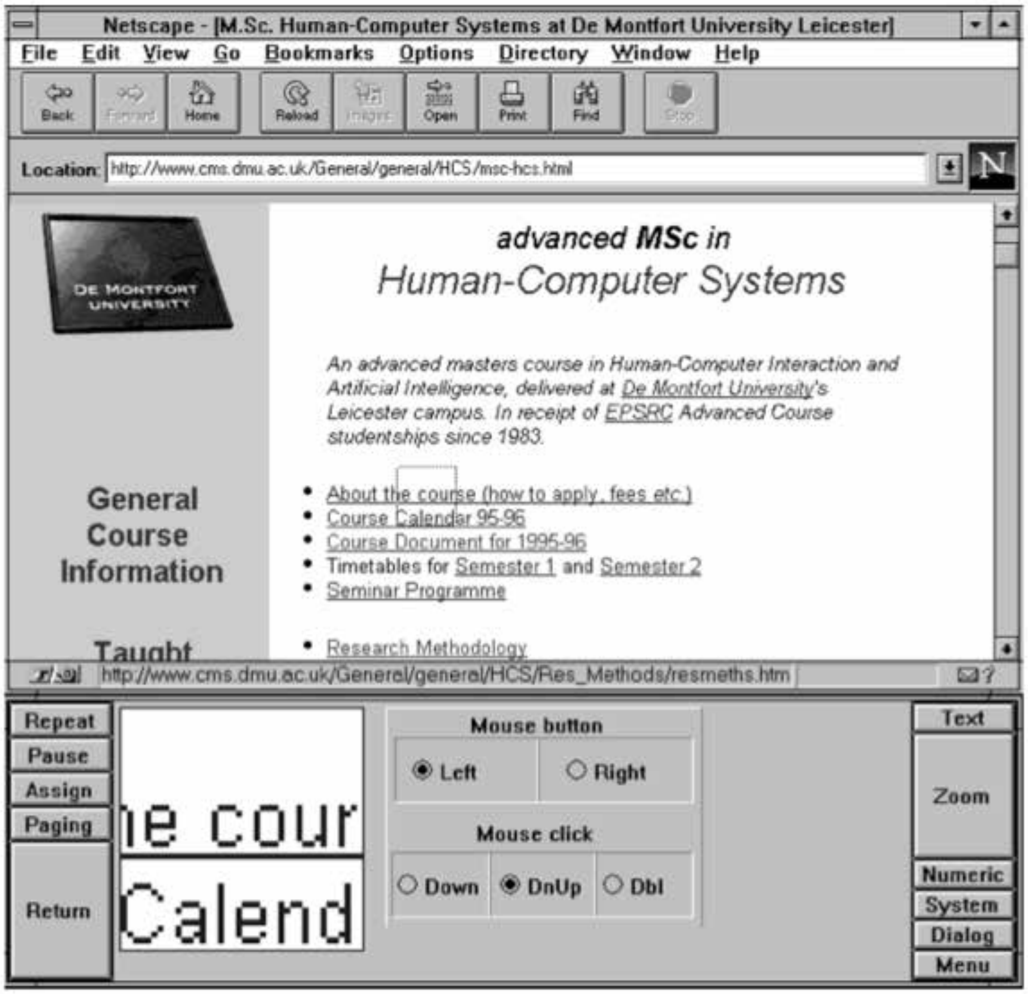

FIG 12 *ZOOMPAD USED TOGETHER WITH NETSCAPE.* TAKEN FROM INSTANCE, SPINNER & HOWARTH, 1996.

approach was to activate the fisheye distortion only when the user completed a saccadic eye movement towards a target and hiding the effect during searching and recognition tasks. This way the system allowed the user to maintain an overview of the desktop during visual search and zooming only in the region of possible interaction. They concluded that this kind of interaction technique leaded to improvements in speed and accuracy over normal pointing technique, or continued distortion technique. However, it worth to mention that due to the nature of the visual effect caused by the 'lens' effect, the special relationship of the elements presented on the interfaces could be negatively affected, making more difficult to the users the proper identification and pointing of these elements.

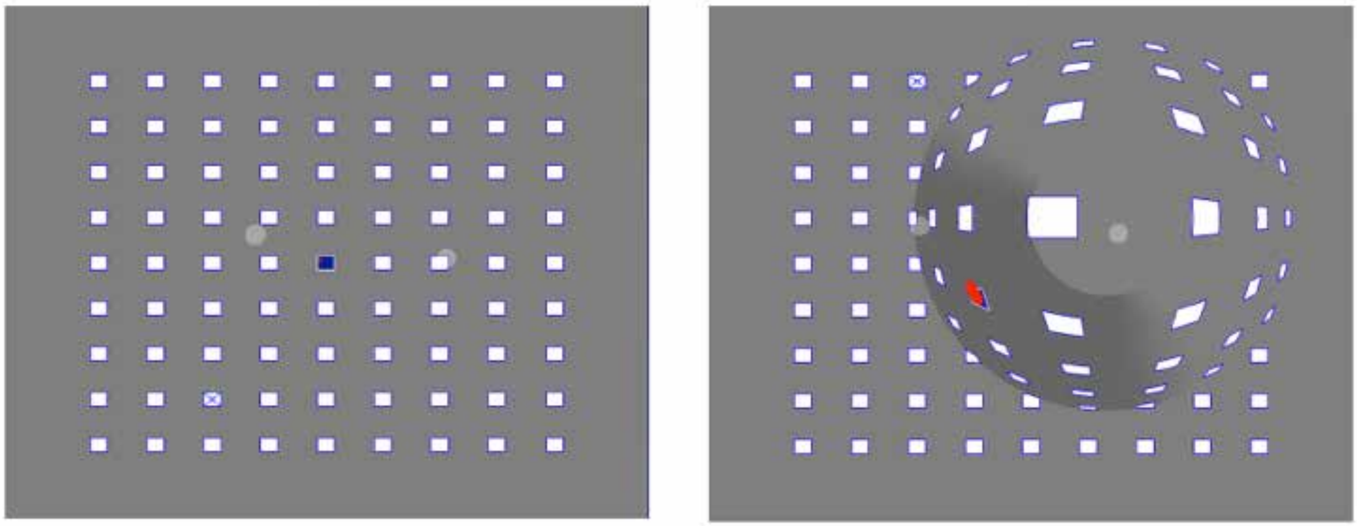

FIG 13 *COMPARISON OF THE INTERFACE DISPLAY WHEN THE FISHEYE LENS EFFECT IS ACTIVATED FOLLOWING THE OBSERVER'S POINT OF REGARD.* TAKEN FROM ASHMORE, DUCHOWSKI AND SHOEMAKER, 2005.

4.2. Attentive user interfaces

Attentive user interfaces are considered the interfaces where the information of user input -including eye natural movements- is subtly used in the background of the system, instead of change their behaviour by receiving explicit commands. Duchowski (2017) proposed two kinds of applications for interactive systems that react to user's gaze: selective, when the gaze is used as an actual control medium; and gaze-contingent, when the system is aware of the user's gaze and may adapt its behaviour based on the visual attention for the user. Vertegaal (2003) brought up a critique to current graphical user interfaces (GUI) indicating their tendency to saturate the limited attention of the user with permanent reload of information, especially in an era where several systems are always connected to wireless networks. *"Devices bombard users with requests for attention, regardless of the cost of their interruptions."* He introduced the concept of Attentive User Interfaces (AUI) referring to the displays that are created for tracking the observer's focus of attention so they structure their communication in such way that the limited resource of user attention is allocated optimally across the user's tasks and the information displayed by the interfaces is estimated by the priorities in user's activity.

Nielsen (Nielsen 1993) suggested a new generation of user interfaces beyond the WIMP (windows, icons, menus and pointing device) workstations. He envisioned systems that involved elements like virtual realities, head-mounted displays, sound and speech recognition, pen and gesture recognition, animation and multimedia, limited artificial intelligence and highly portable computers with cellular or wireless communication capabilities. He also proposed the transition from the object-oriented paradigm within graphical user interfaces, based on direct manipulation by selecting and activating commands; towards a user-oriented and task-oriented approach where communication between user and system would be characterized by noncommand-based dialogues: *"The key notion here is that the specification of both action and object are unified into a single input token rather than requiring the composition of a stream of user input"*. In the same article, Nielsen highlighted the impossibility of use eye tracker systems as a direct substitute for traditional pointing devices such as the mouse, since it is impossible (or at least difficult) to distinguish between recognition and selection with gaze. Instead, he remarked eye tracking as a potential input device for noncommand-based interfaces due to its capacity to recognise in certain level what the user means by looking at something, arguing that *"users tend to look more at things they are interested in than at things they are not interested in"*.

Studies have demonstrated that eyes' point-of-regard has the potential of provide fundamental input information for attentive interactive applications. Gaze-orchestrated dynamic windows (Bolt 1981) is an experiment aimed to introduce order and control through gaze-directed commands. It was performed in at MIT's Architecture Machine Groups-s laboratory, in a room named "Media Room". On one of the walls of this personal-office sized space a multiplicity of "windows" was projected. The size of any window reflected the density of the information and the relevance of its subject, which included topics such as conversations, groups of people, newscasts, radar maps, movies and slides. The main interaction performed by the user was the zooming-in/out of any window depending how long he or she stared to a specific point on the display. This study envisioned eyes-as-output as a promising factor in interactive graphics systems by taking advantage of the observer's attention.

The transition from Graphical User Interfaces (GUI) to Attentive User Interfaces (AUI) comprehends a meaningful change in the way the interaction between users and systems should be conceived. To visualize the change of key concepts in interface construction, Verteegal & Shell (2008) proposed a model that shows the how the GUI elements of interaction are extended and modified to follow the new paradigms of AUIs (Fig 14). As seen in the scheme, there is a substitution of windows and icons for increases and decreases of information resolution in the foreground and background of the user's attention. Also, the pointer is replaced by the information collected through the gaze sensing and the tracking of its position. Finally, menus and alerts are replaced by a turn-taking conversation, where the system is able to recognize the optimal time to activate an interruption in the system according to pending activity.

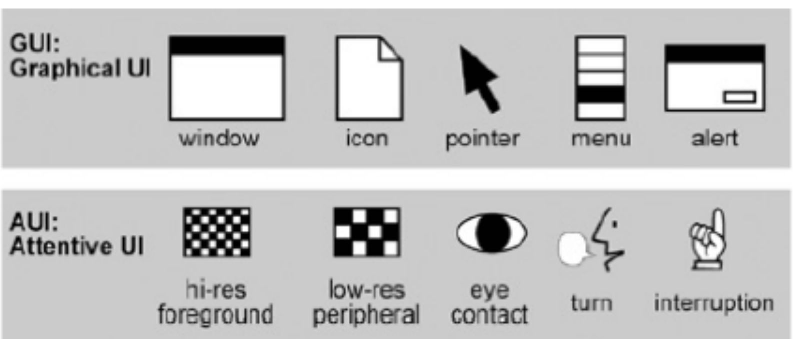

FIG 14 *EQUIVALENTS OF GUI ELEMENTS IN ATTENTIVE UI.* TAKEN FROM VERTEGAAL AND SHELL, 2008.

Zhai defined eye-tracking as one of the systems with more potential among AUI proposing the following hypothesis: "where the user looks contextualizes much of that user's action and hence can be applied to facilitate human-computer interaction" (Zhai 2003). At the same time, he refuted the potential of gaze as an input to naturally perform explicit control tasks, claiming that the eye is categorized as a perceptual organ and not as a motor-control one. *"To load the visual perception channel with a motor-control task seems fundamentally at odds with natural behaviour in which the eye searches for and takes information from the world and in which the hand manipulates objects in that world"*. To exemplify his theory he described a gaze-based cursor pointing system. MAGIC (Manual Acquisition with Gaze Initiated Cursor) is a system that tracks and analyses the user's gaze activity including fixations, saccades and interaction history, to then combine this information with the location of user interfaces objects so the system can places the cursor near to probable hot spot targets. Once the cursor's start position is defined, the user needs only to perform small adjustments of the cursor using a manual input device to reach the target. One of the advantages of this system is that, contrary to traditional eye-gaze control systems, users can select objects independently of their size, and it is well suited for small, finger-operated input devices such as touchpads. Since the cursor appears in the fovea area of user's gaze, another convenient feature of this system is the fact that user never loses the cursor and it does not impose any additional burden or cause extra fatigue while performing eye movements.

Istance & Hyrskykari (2012) mentioned the importance of eye tracking technologies within the operation of attentive interfaces, by considering the monitoring of gaze (supplemented by additional methods such as brain activity measurement) as the most reliable method to obtain information about the focus of an observer's attention. They described several terms used for referring to the interfaces that attempt to react according to user's unconscious input. One of them is "attentive systems", which relate with applications that change state based on the current activity of the user and her or his possible intentions. This means that systems pursuit to offer default commands to users corresponding to actions that they are most likely to perform or potential needs they could have. The term "adoptive systems" refers to systems that modify their state based on the information collected about the user's patterns of interaction during certain period of time. "Affective systems" make use of the user's emotional state, which can be measured through sources such as facial expression, voice or heart rate. To consolidate the terms above, a broader one is often found in researches: "context-aware systems". Every type of system mentioned before is characterized by two distinct phases 1) determining the current state of the user (attention, cognitive, affection, etc.) and 2) deciding what to do with is information; in case it worth to execute any action.

One of the most significant objectives that attentive user interfaces pursuit is the decreasing of the cognitive load of the user by optimizing the way that information is displayed across the interfaces. By doing that, the attention of the observer (in the case of gaze-based systems) is used in such efficient way that task and goals could be achieved in a faster and more natural course. Duchowski et al. (2012) described the concept of "gaze-contingent display", which refers to systems that "...*attempt to bal-*

ance the amount of information displayed against the visual information processing capacity of the observer through real-time eye movement sensing". By using different kind of techniques such as mipmapping, multi-texturing and fragmentation, Gaze-Contingent Displays (GCDs) reduce the content's resolution quality in the peripheral region of the observer's vision, making the high-resolution region moving with the user's point of regard. This method aims to match the physiological characteristics and limitations of the human retina (as explained in chapter 2), and exploit them in a way that the bandwidth of the interaction is maximized.

The range of applications using Attentive User Interfaces within gaze-based systems is considerably wide. Vesterby et al. (2005) conducted an experiment where observers could influence over the narrative branching of a two-minute video clip by only using the gaze as indicator of attention. The objective of the experiment was to discuss the possibility of merge the selection process with the viewing process, avoiding this way wilful decisions that could cause disruptions in the experience with audio visual material. Tracking the user's gaze while watching the video clip could provide information about his or her interests. This information can be potentially used to establish the narrative branch that the video clip will follow and this way fit in some way with the user's interest and expectations. Half of the participants in this experiment were unaware of the interaction with the narrative structure by using the gaze, however this fact didn't represent any issue in the experience. Actually, it seemingly means that the user's awareness of the control through gaze movements it is not necessarily required, and this fact could facilitate a more natural interaction with the narrative by not disrupting the immersion of the user in the storyline.

4.2.1 Components of an attentive system

There are different components defined as the foundation of systems that modify their behaviour based on the information collected about what user is doing, thinking of doing or just done. Istance and Hyrskykari (2012) proposed a very useful diagram that synthetize four main components and their co-relation inside a system mainly relying in visual attention (Fig 15).

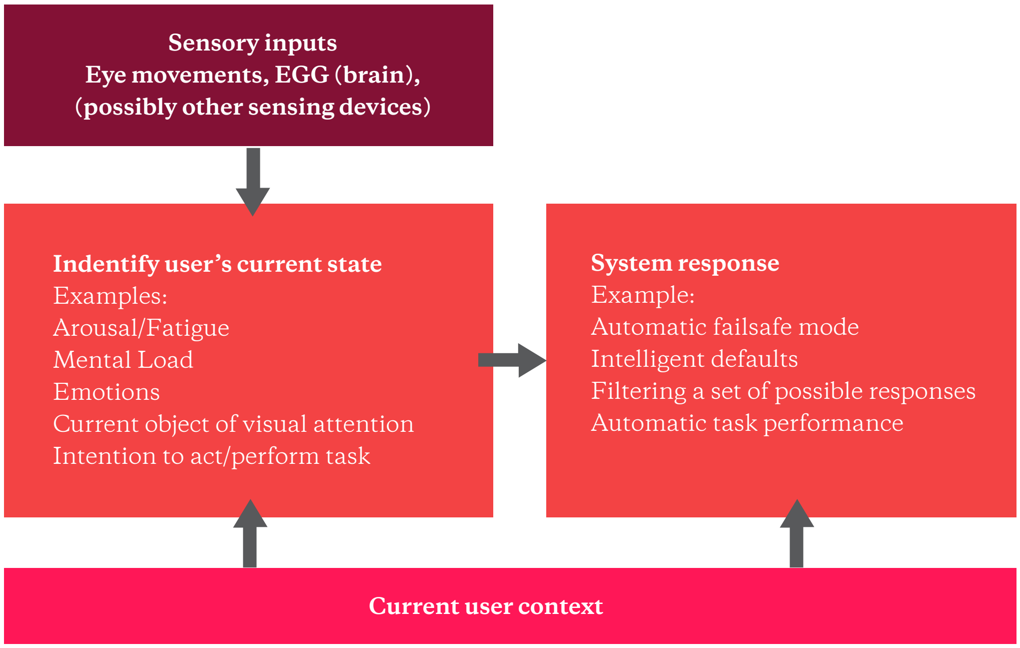

FIG 15 *MAIN COMPONENTS OF AN ATTENTIVE SYSTEM.* TAKEN FROM ISTANCE AND HYRSKYKARI, 2012

1. Identifying the current state of a user of an attentive system:

Different experiments have demonstrated how in several scenarios humans tend to align their gaze with an object prior performing an action over it (Maglio et al. 2000). This information could trigger some modifications in the system according to potential user's "next steps". By monitoring the changes in observer's focus and analysing fixations and saccades over time it is possible to create a "gaze-path" that eventually could be used by the system, along with additional information such as the application logic and the user context, to generate patterns that provide clues about user's intentions. Bee et al. (Bee et al. 2006) conducted an experiment about predictions in the selection task using neckties as variable element. They described a "cascade effect" that indicates a gradually increase in the attention over certain preferred object until it is finally selected. The success of the match between the participant's conscious choice and the prediction of the system based solely in the analysis of visual attention was close to 80%. Nevertheless, the prediction accuracy in terms of precision and timing represents some concerns related with the correct identification of the user's current state.

Information collected about mental load, fatigue and emotional status of the user could also contribute to trigger accurate responses of the system in order to enhance the experience with the user. Studies have revealed that people in positive mood are more likely to acquire new knowledge and be more creative. Thus it could be said that a frustrated or tired users possibly need some remedial actions from the system. Unfortunately, using only gaze tracking technology to recognize this kind of information from the user seems to be insufficient. That constitutes another reason why the integration of gaze-based system with additional input monitoring devices holds substantial relevance.

2. Appropriate system responses

Depending on how the attentive system is structured and positioned with regard to the user, two sub-categories are conceived. The first category involves systems that are integrated into a desktop environment and the usual input device includes a large, stationary screen. In the second category, there are the remote systems where the object that receives visual attention is away from the desktop.

a. Stationary systems:

As mentioned before, analysing the monitoring of gaze in an accurate way provides insides about the context for an action that is likely to performed. Depending on the level of influence that eye movements apply over the system, application techniques can accelerate or enhance the desktop interaction. Considering a system as MAGIC (Zhai, Morimoto & Ihde 1999) where the gaze position assists the location of the pointer on the screen when user is looking at a potential target is a good example of acceleration in desktop applications using gaze-based systems. Here, the adaptation of the system according to the gaze position is simply, and even the flaws in the measurement of the point of regard can be later corrected by moving the mouse. However, there are interactive applications where the analysis of the gaze establishes a more deep insight for the system performance, looking for enhance the experience of the user and provide meaningful responses in order to achieve it. To exemplify, iDict is aid for promoting fluent reading of foreign language documents with the help of gaze path information (Hyrskykari 2006). The system is aware of the gaze movements while the user reads an on-screen document, and detects situations where the user seems to require assistance. Long fixations over certain word, can trigger the system in order to display a visual help containing the translation of the possible word the reader is having troubles with. Due to the bigger influence that enhancing techniques apply over a system; they require a high level of user understanding, in comparison with accelerating techniques. This aspect is translated in larger efforts in the development of the intelligence of the system, which should be able to recognize separately the situations where the user require assistance and the circumstances where he or she perform normal eye movements to recognize the elements of the screen. Important principles like controllability, appropriate feedback and unobtrusive interaction must be thoughtfully considered in the design process of such systems.

b. Away from desktop

This type of attentive system usually recognize when the observer's gaze is aligned with a certain object or person, and then it triggers a direct action over the object or a change in an provided interface so the user can perform the commands related with the object previously selected by the gaze direction.

Some techniques have been developed for the use of away from desktop applications. One of them is based solely in the gaze detection of the user, with no intention of track the exact position of the gaze. Normally, a simple set up is built by using a corneal reflection sensor and a small video camera. When a person looks straight to the sensor, the system activates a respond, without getting additional information about the eye movements. Issues associated with this technique include the distance between the observer and the object, the distance between objects so the signal is not erratically read, and the amount of people's gaze that can be detected at the same time. These kinds of applications are often related with the manipulation of objects by users with movement impairments or smart home appliances.

Additionally to the gaze detection, there is a second technique that considers the position of the gaze and the different eye movements while the person is looking at some object in particular. Yonezawa et al. (2008) described a system composed by a single camera mounted on top of a board and guide toy robot. When the observer changed the position of the gaze, the robot gave more information related with the part of the board the user is looking at. Additionally, the robot's head also changed its position, giving the illusion that it was following the user's gaze path.

4.2.2 Attentive user interfaces implementation challenges

The design process of any Attentive User Interface contains specific challenges due to the key features of its operation. The complexity of a gaze-based system where the interaction is transparent for the user needs to be highly accurate so it can respond in the most efficient way to the -almost always- unconscious commands performed through unaware saccades, fixations and other eye movements. Along with this challenges, detecting and understanding the gaze patterns that users perform in order to attempt to recognize his or her specific current activity and even aim to anticipate potential and relevant needs related with the on-going context, represents a major effort within the development of AUIs and its integration with artificial intelligence.

> **The user interface to automatic behaviour requires very careful design, to ensure that it is not intrusive to the user's normal work. If it is, then there is considerable risk that the user will be irritated and disable the adaptive functionality entirely.**
>
> - ISTANCE & HYRSKYKARI, 2012

The integration of AUIs with new software and hardware developments is another essential consideration. Devices that are able to measure user's physical activity and location, detect gestures, measure eye position and other features; can be considered as a valuable input source for the system in the process to perceive the current user's context and needs. Furthermore, the recently production of autonomous robotic devices, such as cleaning assistants or smart toys, represents the need to frame interactions that make the response from the system to be more natural. However, it means a significant effort to integrate all this information received by diverse devices and sensors, and attempt to translate this data into patterns that allow the correct understanding of the user and thus provide the accurate response through the system. At the same level, the proliferation of internet access has had a major impact different aspects of daily human behaviour, and also has increase the expectations of users concerning the usability and accessibility of diverse products and technologies. For that reason, it is imperative for AUIs to be accurate in anticipating what a person is interested in doing. Any interpretation of the system that it consider useless could cause the discomfort in the user and the deactivation of the attentive features.

Finding the appropriate mechanisms to provide feedback from the system to the users is an aspect that must be treated with special care as well. The dichotomy between hiding the feedback to avoid potential interruptions and visual distractions for the user, or utilize a evident feedback so the user can be aware of the processes the system is performing in the background, opens a series of possibilities that should be measured and tested according to the level of transparency of the system and the characteristics of the activities performed by the user and the corresponding responses from the application. Hyrskykari et al. (2000) make a very interesting reference about this aspect: *"The more natural an interface gets, the more transparent it appears to the user: the user starts to interact with the task instead of interacting with the computer"*. In other article, the same author suggested that avoiding feedback in situations where the user has no conscious use would contribute to minimize the visual noise of the system: *"The feedback should be filtered in such a way that it consolidates the user's conception of background interpretation of the gaze behaviour"* (Hyrskykari 2006).

Because of the attentive user interfaces' capacity of ubiquitous sensing it is rational to think about the extensive amount of personal information that system could collect about people's habits, routines, preferences, priorities and so on. The main use of this information would be to serve as a filter within the communication between users and systems, but there is also an evident risk related with user's privacy and opening an opportunity to undesired surveillance. Vertegaal & Shell (2008) make a call for an joint effort throughout the process of designing attentive interfaces *"Designers must partner with social researchers to make use of the wealth of knowledge they possess, and social scientists must continue to take an active role in the development of new technology in order to influence and improve technologies that are changing the way we live."* They approach this issue, suggesting that information collected by AUIs should be store in local encrypted repositories. Doing so, the users could be still beneficiaries of the improved communication provided by these systems, without sacrificing informational privacy.

4.2.3 A framework for attentive user interfaces

In order to enhance the communication with users, attentive user interfaces have the need to be aware of user's past, present and future attention focus so pattern models can be created and priorities on user's tasks can be accurately established. Vertegaal and Shell (2010) defined a framework compose by 5 properties that any AUI should include:

1. Sensing attention: AUIs can determinate the device or task a user is most likely to attend by tracking the person's physical proximity, body orientation and eye fixations.

2. Reasoning about attention: By modelling interaction patterns of use, the system could be able to recognize priorities in user's tasks and thus "know" about the possible current needs to fulfil.

3. Communication of attention: The system should provide information about the users to other people and other devices. Additionally, knowing whom or what the user is focusing attention should allow AUIs to determine whether a user is available for communication.

4. Gradual negotiation of turns: Interfaces should know if a user is available to be interrupted for communication by checking

the priority of the request, then signalling this request using peripheral channels and finally sensing user acknowledge of the request before changing the resolution of foreground information.

5. Augmentation of focus: The AUIs should augment the attention of the users by performing gradually changes in the resolution of background and foreground information, only when the system recognize an optimal moment to interrupt the user.

4.3. Gaze-based user modelling applications

While the first to categories of gaze-based applications explicitly or implicitly utilize the user's point of regard as an input, gaze-based user modelling applications monitor and analyse the dynamics of visual behaviour over time. Some of these applications have a historical connection with experimental psychology, in the pursuit of understanding the cognitive processes of visual perception and the aspects related with memory, knowledge, or learning. Bulling & Roggen (2011) conducted an experiment introducing eye movement analysis as a promising method to assess the user's cognitive context. They found out that it could be possible to infer the user's visual memory status by analysing the eye movements of fourteen participants and monitoring the gaze paths, fixation and saccades while looking at familiar and unfamiliar pictures.

The models obtained through the monitoring of eye movements during certain period of time are commonly used to support computational methods from machine learning and pattern recognition. Kandemir & Kaski (2012) studied the idea of systems able to learn to use the user's natural eye movements to infer relevance of real world objects by conducting an experiment that integrates eye tracking technologies with machine learning algorithms. They choose a specific use case where users explore paintings in an art gallery. Through a wearable computer the system collected data to predict relevance in user's interest. The key goal of this kind of experiments is to obtain better understanding about user behaviour and translate those insights into automatic responses.

The detection of user's intentions provided by eye movements has also been studied by researchers. Bednarik et al. (2012) developed and tested a framework for intention prediction supported by machine learning approaches in biometric person authentication. Even if the results of the experiment were not as optimal as expected, the study opened a new discussion about the intent recognition based on eye movement data. Eye tracking technologies seem to have promising features related with the construction of prediction models, when it is combined with machine learning practices.

Further efforts are needed to efficiently build models that allow systems to approach user's cognitive status as a main component of smart applications. The information collected through user modelling applications could be considered as a foundation to be applied into attentive and explicit input interfaces. The more researchers know about people's behaviour, interests and focus of attention; more accurate automated systems can be envisioned and developed.

4.4. Eye movement in Usability Research

As mentioned before, the idea of utilize eye tracking systems to collect data related with usability issues has existed even before the creation of modern computer interfaces. Early efforts made in the 1950s by Fitts, Jones and Milton proposed some conclusions that are usefull nowadays, including the commonly cited "Fitts' law", which states that the amount of time required for a person to move a pointer (e.g., mouse cursor) to a target area is a function of the distance to the target divided by the size of the target. Therefore, the longer the distance and the smaller the target's size, the longer it takes for accurately point it (Fitts 1954).

Even if the technology used by Fitts would be appointed as obsolete today, his concepts are still applied in user experience and user interface design. With modern technology it is possible to merge gaze's real time position with data visualization in order to obtain information about observers' behaviour, such as where they looked, the length of time they spent and the gaze pattern they followed.

Tracking people's eye movements provide Human-Computer Interaction researchers the understanding of visual and display-based information processing and the factors that might possibly influence in the usability of interfaces (Ball & Poole 2006). This practice supplies insights into what the observer found interesting, that is, what drew their attention, and perhaps even provide a clue as to how that person perceived whatever scene she or he was viewing (Bergstrom & Schall 2014). For the usability specialist, eye tracking can provide information about when a participant is searching for visual targets versus reading text (Goldberg & Wichansky 2002).

In this way, diverse parameters of the elements existing in the design of an interface, such as visibility, meaningfulness and placement could be objectively evaluated in order to obtain quantitative metrics for objective design iteration (Goldberg & Kotval 1999). For instance, in a task where participants are asked to interact with buttons and icons, gaze timing and path could indicate the level of meaningfulness and accuracy of the element, and thus, present hints about its improvement.

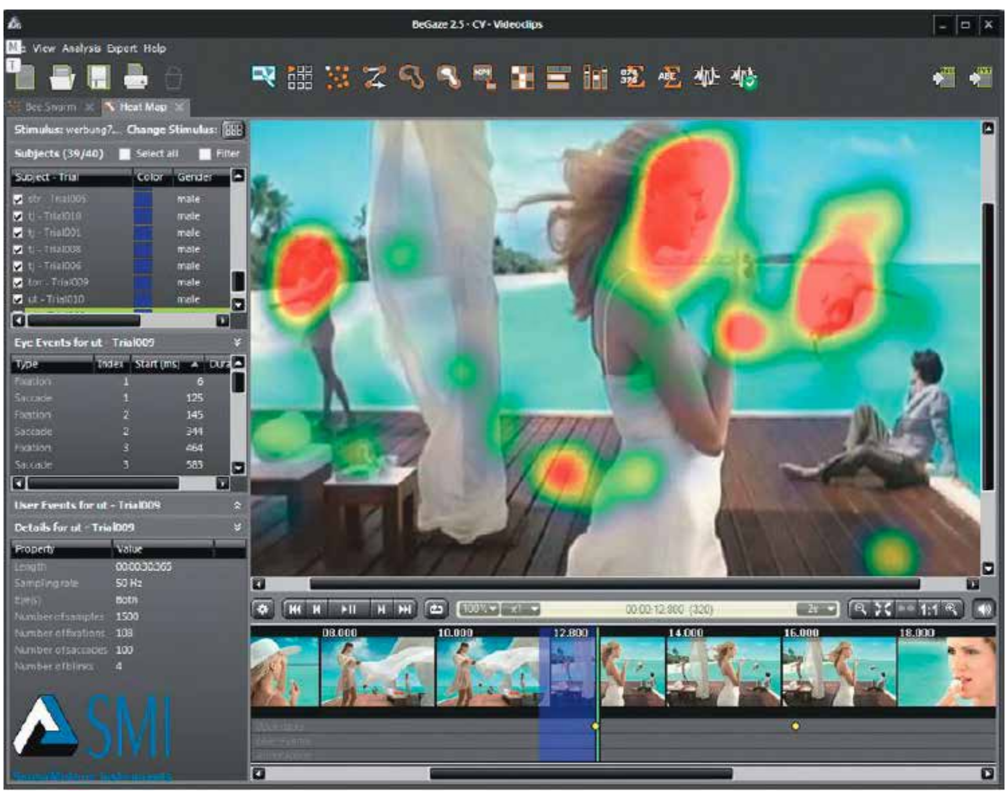

FIG 16 *EYE TRACKING ANALYSIS SOFTWARE EXAMPLE.* TAKEN FROM SHALL, 201

	Explicit input applications	Attentive User Interfaces
Definition Majaranta and Bulling (2014)	In explicit input applications people use voluntary eye movements and consciously control their gaze direction, for example, to communicate or control a computer. The most common way to implement is to use the eye's capability to point at the desired target.	AUIs can be considered as interfaces where the user is not expected to change his or her gaze behavior to give explicit commands. Instead, the information of the user's natural eye movements is used subtly in the background.
Advantages	▪ Selection speed ▪ Free hands interaction ▪ Very useful for disable users ▪ Relatively low learning curve	▪ Natural human-computer conversation ▪ Non-intrusive interaction ▪ Reduce user's cognitive load ▪ System adaptability
Challenges	▪ Accuracy and responsiveness ▪ 'Midas touch' ▪ Feedback	▪ Feedback ▪ Understanding of gaze patterns ▪ Integration with devices and sensors
Common integrations	▪ Mouse and keyboards ▪ Speech recognition ▪ Face and body gestures	▪ Artificial Intelligence ▪ Positional sensors (GPS, Accelerometer)
Examples	▪ ERICA, Hutchinson et al. (1989) ▪ GazeTalk. Hansen (2006) ▪ EASE. Zhai (2003)	▪ Attentive Windows. Bolt (1987) ▪ Interactive narratives. Vesterby (2005) ▪ MAGIC. Zhai et al. (1999) ▪ iDict. Hyrskykari (2006)

FIG 17 COMPARITION BETWEEN EXPLICIT INPUT APPLICATIONS AND ATTENTIVER USER INTERFACES

CHAPTER 5

Multimodal gaze-based interaction systems

How the eye tracking experiences can be enhanced with additional input methods.

COMBINATION OF EYE TRACKING technologies and traditional input methods have been implemented into multi-modal applications in order to fulfil interaction flaws and potentiate even more the benefits for the user. Below a description of the frequent integrated systems with specific examples.

5.1 Eye tracking and traditional input methods

The common approach for these applications is to translate the gaze position information retrieved by an eye tracker, into a pointer within an interface. Then, by pressing a key or clicking, the user sends a signal to the system so it activates a command over an element or an area.

EyePoint (Kumar, Paepcke & Winograd 2007) is a prototype that combines zooming techniques with multimodal input methods. To initiate, the user locates the target by looking around the interface, then by pressing a key the observer brings up a magnified version of the region that was looked at. After that, the user locates the specific target within the zoomed window and finally he or she releases the keyboard key to issue the click. Other applications with similar behaviour have been created to simplify the switch between windows when using a computer. EyeExposé (Kumar 2007) window selection method is based on pressing a key to show all the windows available. Then the user aligns the gaze with the desired window, and finally releases the key to complete the selection.

In regards to systems combining mouse and gaze interactions, Räihä & Spakov (2009) proposed a technique to select targets when using multiple monitors. To overcome the long mouse movements, it was suggested to include a pointer in each monitor, and activate the pointer that was closer to the user's point of regard. Users indicated preference for the gaze-based technique instead of the regular one, due to improvements in speed of selection.

5.2 Eye tracking and speech

Combining gaze input information with speech recognition is an attractive option to improve the robustness of the interaction within a system. Overcoming the implementation flaws of gaze-based systems with the advantages of the speech input could be an option, especially when considering issues such as the selection of small targets and the activation of commands using only the gaze. The common approach to combine this two interaction channels has been to rely selection tasks to the user's gaze and to perform the initiation of actions through voice commands.

Castellina, Corno & Pellegrino (2008) proposed a very interesting approach by testing a system that integrates gaze and speech within the context of realistic desktop environments. The system architecture of that test is based on five elements: 1) the eye tracker which is responsible for the identification of the user's fixation point; 2) the screen reader which receives the input from the eye tracker and returns a series of objects base on the observer's gaze position and defines a corresponding name, role, state, default action and position. 3) the grammar generator produces an appropriate VXML grammar for the speech recognition module, based on the characteristics of the objects. 4) a vocal unit receives the possible commands defined by the grammar generator and supplies the commands pronounced by the user. This unit analyse every spoken word and interpret them in order to notify the user if the command is recognized, ambiguous or wrong. If the command is recognized properly by the vocal unit the command is passed to the 5) action executer unit, which perform an action responding to the command retrieved by the data structure and matching it with the object selected.

This approach revealed positive conclusions about the complementary characteristics of both gaze pointing and speech recognition working together in one interaction system. From one side, accuracy problems of eye tracking devices (specially with small targets) are overcome with the correct use of vocal commands; and from the other side, the assistance of pointing and selecting using the gaze enhance the efficiency of ambiguous commands using the voice. In a similar experiment Zhang et al. (2004) proposed the simultaneously use of gaze and speech command to select an object on a system's interface, based on the follow concept: *"The gaze information gives a spatial limit to the multimodal integration, whereas the speech information gives a semantic limit"*. Their goal was to gain insights into multimodal performance errors and possible ways to mitigate the perceived flaws. One of the key concepts in this approach was to consider an "eye-operative" region. This is defined as an area surrounding the eye cursor point that allows the system to pre-activate targets inside so they would be ready to be fully activated through voice commands. In order to minimize errors, they suggested that size of the eye-operative region should be large enough to include the desired target but also limited to exclude possible irrelevant targets. They emphasised the importance of continuous measuring and error analysis of multimodal integration strategies in order to refine their functionality and effectiveness.

5.3 Eye tracking and facial muscles gestures

Another popular multimodal system technique that has been developed recently is the combination of facial gestures commands (not only made by eye muscles) read by facial electromyography (EMG) devices, with the gaze input retrieved by eye trackers. This method provides users the possibility to adopt additional mechanisms to improve their communication with the interface of an application. In some cases the interaction has been proved to be more efficient that using dwell-based selection. The muscles of the forehead, the ones activated while smiling and the muscles of the jaw, perform the most frequently studied gestures within this kind of systems.

San Agustin et al. (2009) performed an interesting evaluation of a system combining gaze pointing and EMG clicking within the context of game interaction. They compared the performance of the combination of mouse and gaze pointing with button and EMG clicking in a target acquisition task, divided in two main activities: pointing and selection. So four input combinations were proposed by using two pointing methods (mouse of gaze) and two selection methods (mouse button or EMG switch). Each of the five participants of the test completed a block of trials for the four input combinations. Target's diameter was between 100 and 150 pixels, and the EMG activity was measured with a CyberlinkTM system, which sent a click command to the computer each time participants slightly frowned or tightened their jaw.

The experiment yielded relevant results related with the throughput of the input variables. They found that gaze-EMG combination could perform as well as the mouse while allowing the user's hands to be used to control other functions. This in spite or the face that users received limited practice with a novel interaction device and technique. In the same way, they mentioned the importance of the eye tracker's accuracy, being this variable vital to the level of engagement or frustration from part of the observer. In other research made by Tuisku et al. (2012), gaze-based input was combined with facial gestures. More precisely, during this experiment, the muscles activated when smile *(zygomaticus major)* were monitored in order to execute commands over a digital keyboard presented within an interface. Using a prototype called Face Interface, ten users were asked to type the finish word "aurinko" (sun in English) using their gaze to point the targets in the screen and the smiling gesture as trigger for selection. Every time the users finished the type task, the characters were randomly rearranged so it was impossible for them to memorize the position of the letters.

The prototype was built on the frames of protective glasses and it was composed by two cameras, an infrared light, a series of capacitive sensors to detect facial movements and a Bluetooth module. Before conducting the experiment each participant had five trials to get on-board with the system and the interaction techniques. After using the Face Interface technique to perform the task, participants were asked to repeat the typing task using a mouse as pointing device, so differences between both input methods could be measured. The text entry experiment showed that the overall mean text input speeds with Face Interface and mouse were 20 cpm (characters per minute) and 27 cpm, respectively. One interesting finding for Face Interfaces was that one participant had a maximum text entry speed value of 47.4 cpm with was even higher than that participant's maximum speed with mouse, that was, 43.3 cpm. Considering the low familiarity of the participants with the prototype, results revealed a promising future for text entry with this kind of devices. Also they confirmed that smiling could be used as the selection technique, which can be exploited in future systems.

Combination between gaze-based systems and facial gestures monitoring can be very promising principally in scenarios when users are partially of fully unable to perform physical movements or their hands are used to execute other actions. However is important to consider the user's lack of experience interacting with this type of techniques and thus the possible challenging learning curve.

CHAPTER 6

Promising future applications

What are the most promising fields for gaze-based applications for the following years.

IN THE PREVIOUS CHAPTERS the most common domains for gaze-based interactive applications have been mentioned. Assistive systems for disable users and attentive pointers to facilitate selection tasks on monitors, make part of the traditional approaches to applications of eye tracking. However, during the last years, a new wave of application fields have grown as consequence of technological developments on equipment and software. Low cost sensors, smaller video cameras and accessible software are just some examples of the tools that designers, developers and companies can now approach to integrate eye tracking with other technologies like virtual/augmented/mixed reality or wearable devices. Industries such as health, entertainment and security have started to incorporate the use of gaze as an important component at their research and development departments, demonstrating the promising potential of eye tracking systems into a wider range of interactive application fields.

6.1. Videogame applications

The videogame industry is constantly looking for generate more immersive and natural experiences for their users. For that reason, Tobii has recently developed an eye tracker specifically to enhance users' gameplay, by using their gaze position as an additional input control. So far, more that 150 games have integrated this technology as a complement for traditional controllers like keyboards, mice and gamepads. There are some key features that developers had consider to exploit the potential of eye tracking in video games, especially in action-adventure genre:

Infinite screen: The camera angles will automatically rotate depending on the eye position. For instance, if the gaze reaches a region close to the edge of the screen the camera will move accordingly in order to expand the surroundings of the virtual world more a fluently and natural way.

Direct interaction: It is possible to aim, fire and tag characters and objects using the gaze. The camera will automatically centre on the target the user is looking at to make possible to activate some command through the manual controller.

Immersive graphics and gaze awareness: The sound and graphics of the virtual environment will adapt dynamically according to the user's point of regard. For instance,

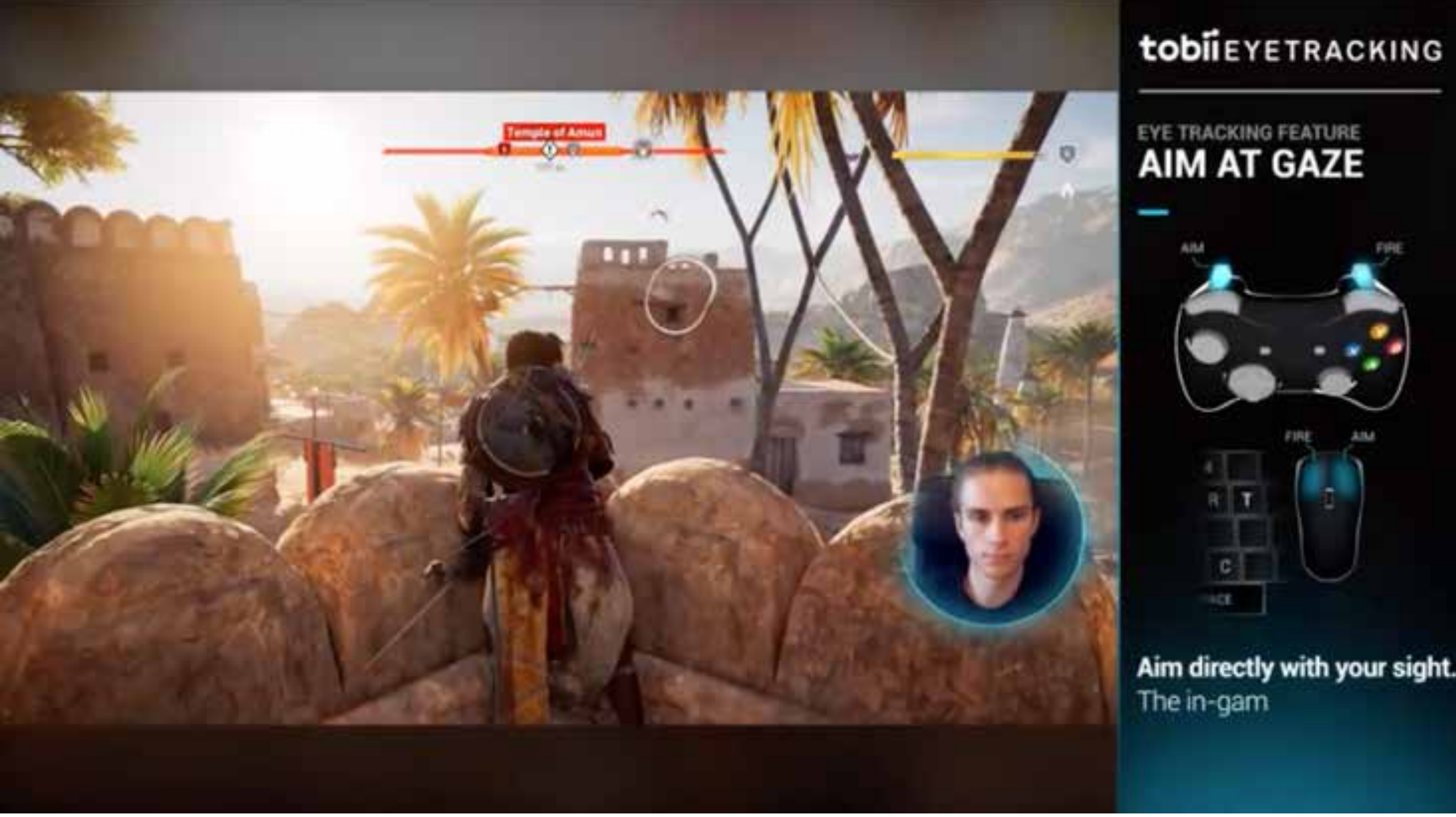

FIG 18 *DEMONSTRATION OF ASASSIN'S CREED © ORIGINS GAMEPLAY, SHOWING THE INTERACTION FEATURES BY USING TOBII'S EYE TRACKING TECHNOLOGY.* FULL VIDEO EXPLANATION

some elements like maps and stats could be transparent until the user align the gaze to them. Resulting in cleaner user interfaces and more immersive experience.

Further developments are likely to keep exploiting eye tracking potential into different video game genres. It is important to consider that the way the gaze information is used should depend on the specific characteristic of the gameplay, so the gaze interaction blends properly with the game experience and the new features could be perceived as natural as possible.

6.2. Virtual reality applications

The integration with virtual reality (VR) is one the most promising application fields for eye tracking technologies. The possibility to monitoring and control the human behaviour in semi-realistic conditions opens novel methods for conduction research concerning human cognitive processes. By using geometric calculations, it is possible to calculate the position of gaze of the subject in 3D spaces, define regions of interest and trace the points in time to the establish which objects and areas were looked at. This provides researchers with the tools to modify and test controlled scenarios more efficiently, which means shorter iteration times and higher quality data collected.

Clay Clay, König & Koenig (2019) presented an example of integration between eye tracking with VR applications. The output of the experiment was to examine the behaviour of 27 participants while they interacted with a virtual city. First, the users were asked to explore the different areas of the virtual environment for 30 minutes using a head mounted display. Then, the research team presented aleatory photos of the virtual houses to the participants asking them to provide the level of familiarity with each house. Finally, the qualitative data was compared with the information retrieved by the eye tracker about the time that users spent looking at every particular region, in order to understand participant's navigational behaviour. They suggest a huge potential for human cognitive processes research by merging these two technologies: "*The possibility to model various environments and control every aspect of it is highly valuable in research and should be further exploited in the future*". However, it is also mentioned during the research certain issues to overcome in future projects such as motion sickness or fatigue by wearing a headset for long periods of time.

FIG 19 *VISUALIZATION OF GAZE POINTS FROM ONE SUBJECT WHILE EXPLORING THE VIRTUAL CITY. SPHERES REPRESENT HIT POINTS OF THE GAZE VECTOR.* TAKEN FROM CLAY ET AL. (2019)

It is possible for researchers to create customized three-dimensional environments that accurately simulate real life scenarios by using virtual reality technology. Additionally, the use of eye tracking functionalities into these experiences offers the opportunity to obtain valid findings in user behaviour. Meißner et al. (2017) presented a system test based on the virtual experience of shopping in a retail store. By simulating the physical characteristics of product positioning and measuring the eye movements of the participants, the decision making process was monitored and analysed. Additionally, it was possible to test new shopping features, such as the display of product information through simulated augmented reality functionalities. Although the construction of virtual reality environments represent a significant technical efforts, applications like the one mentioned above, suggest that by integrating eye tracking features the amount of validations and reproducibility of the experiments could bring meaningful insights and important benefits for researchers.

The case studies mentioned before combine eye tracking and virtual reality technologies to simulate real life experiences. Both of the projects are oriented to observe and analyse human behaviour by taking advantage of the 3D perception and movement sensation that VR provides. From an interaction point of view, the approach of these projects seems to be user-research oriented. Further projects should explore

FIG 20 *USER INTERACTS WITH THE PRODUCTS INTO THE VIRTUAL ENVIRONMENT. SIMULATED AR INFORMATION IS POSITIONED OVER WHEN THE GAZE-RAYS ARE ORIENTED TO A PREDEFINED AREA..* TAKEN FROM MEISSNER ET AL. (2017)

the possibility to use explicit interaction gaze-based techniques or even attentive systems features, in order to take full advantage of eye tracking benefits. This certainly means substantial efforts within the creation of interaction techniques that fit with virtual environments' characteristics and support natural and meaningful user experiences.

6.3. Gaze-based mobility control

Gaze-based applications have a huge impact in assistive systems for disable users, as mentioned in previous chapters of this work. For that reason, the right to self-directed mobility using eye tracking devices is an aspect that has been matter of research within in the last years. For users with physical and cognitive impairments, gaze-based applications integrated with smart wheelchairs could become a potential way for mobility, without the need to relay in the assistance from caretakers (Bates et al. 2012).

Current wheelchairs systems are commonly manipulated through joysticks or sip-puff switches. However, there is a group of users that cannot fully operate this type of controls due to their particular disabilities, such as paraplegia, amputations and restricted body movement. Thus, combining gaze-based features with mobility platforms could be potentially beneficial for the independence and the general well being of these users. Tuisku et al. (2008) introduced two types of user interfaces that can be adopted in gaze-based mobility systems: In the eyes down technique, the user manipulates the wheelchair indirectly by using a computer screen with on-screen buttons; while in the eyes down the users utilize gaze pointing over the real world to indicate the direction where they wish to go. Both techniques have potential risks related with safety, precision and reaction speed, which represent a significant challenge for the interaction design of these devices.

Purwanto, Mardiyanto and Arai (2009) proposed en electric wheelchair controlled by gaze direction and eye blinking. In this system, the user's point of regard provide information about the direction of movement of the wheelchair, while timing and velocity instructions are giving by eye blink commands. Tsui et al. (2007) presented a hands-free control system of an electric-powered wheelchair based on electromyography signals recorded from eye movements. The system allowed the user to choose either control state or non-control state so non-intended movements could be ignored. Wanluk et al. (2016) developed a prototype of a smart wheelchair that could be controlled using eye movements and additionally allowed the user to communicate with a caretaker via text message.

The safety of the user has been considered as a priority in the design of assistive-mobility products. Serious accidents could result from computer failures or undesired activation of commands from the user. For this reason, safety recommendations most be precisely considered to avoid any kind of hazard or potential danger that could affect users. Bates (2009) suggested some parameters to take into account to try to cover the situations that could arise while using a gaze-based mobility system including the changes between indoor and outdoor light, the system response times and the incorporation of collision detection to alert the user about possible obstacles.

Tailored implementations should be done in future projects so it can be possible to provide accurate solutions to disable users by integrating eye tracking technologies with mobility platforms. There is a huge potential in improving the life quality of users with special needs, especially it terms of independence, learning and rehabilitation.

CHAPTER 7

Conclusions

Research outcomes and reflexions.

BREAKTHROUGH INNOVATIONS are created when technological developments and accurate approaches to use that technology merge in a product or a service. Eye tracking systems and gaze-based applications are not the exception. The most recent techniques to track, measure and analyse users' gaze such as video ocular system, allow to propose and test new interaction techniques that could re define experiences in human-computer communication by retrieving and provide multi channel input and output information. Easier and faster calibration, higher acuity in real time, and modelling of patterns of use are some of the technology's characteristics that as interaction designers should be exploit into the generation of smart systems that consider eye position and movements as the main or partial input method.

In order to take full advantage of interaction techniques using eye tracking technology it is fundamental for designers to understand the eye's physiology, the movements it perform and the strong relation it possesses with attention and cognition processes. For instance, comprehending the limitations in size of the high-resolution area into the visual field and the distortions of peripheral vision caused by the distribution of photosensitive cells inside the retina must be permanently considered within the creation of the visual elements that configure the interfaces into gaze-based applications. In the same way, it becomes imperative to recognize and interpret the role of fixations, saccades, smooth pursuits and other eye movements in order to design interactions that blend naturally within applications and respond efficiently to the characteristics of human physiology. Furthermore, noticing the differences between the distinct types of voluntary and unconscious shifts in attention and their causes, could help to construct meticulous algorithms and patterns of use that enhance the user experience by extending the system's capacity of response according to people's context, tasks and goals.

The various categories of applications presented in this work, could serve as a guide to orientate future gaze-based system projects. However, it worth to highlight that the complexion of every project must attend the whole context of the users involved, which means that hybrid-category products that merge diverse technologies should be explored in order to accurately fulfil people needs and make a real impact in their daily lives. As mentioned in the chapter related with explicit input applications, an extensive amount of gaze-based systems are created to support users with some restriction in their motor functions. In those cases, techniques such as dwell-time and zooming respond to the specific need of restriction of movement, where gaze becomes practically the only input method that users possess to interact with an interface. In other cases, it can be possible to use additional input methods such as speech recognition or facial gestures in order to overcome the eye tracking system's flaws, as long as the user capacities allow it. Using multichannel information in an optimal way, can improve significantly the interaction in terms of speed, efficiency and accuracy.

Moreover, it is crucial for designers to understand the changes in the paradigms that attentive interfaces and systems bring to the products and services that are currently built. The concept of systems that are aware of users' attention and intentions and are able to translate that information into accurate and unobtrusive responses should be highly explored in the search of novel solutions that propose a double-way communication between humans and computers. This of course, means a major effort in the integration of eye tracking technologies with artificial intelligence. The more precise can be the algorithms that read and interpret the user's behaviour; more accurate could be feedback that the system deploys. That is an example of how important is to scrutinize the possibilities of assemble different kind of technologies that support new meanings in people's experiences.

The passive types of gaze-based applications are also a great source of insights related with user's perception and cognitive processes of information. This is a major opportunity to understand human visual behaviour, thus to trace and analyse patterns presented in the interaction with computers. Every piece of information obtained about people's performance is important to propose models of prediction that support automatic behaviour in systems. For that reason, usability tests performed with eye trackers, likewise, could be source of knowledge about what, where and how users look at, providing additional information to increase the acuity within eye movement patterns. Additionally, the integration

with virtual reality environments introduces infinite possibilities to perform user research by measuring human visual behaviour under simulated real life experiences.

Finally, it is important to contemplate intrinsic advantages of eye tracking systems such as the contactless interaction. Considering situations as the current worldwide health alert related by the rapid spread of COVID-19, it is probably a context where this kind of technology could be approached and used for the general wellbeing of the society. Replacing buttons, touch screens and other manual input methods, for gaze-based applications could become in the near future a common standard given by strict hygiene and sanitation regulations. Consequently, the expansion of applications based on eye tracking technologies seems to be very promising and it is likely to expand even more its range of domain fields.

references

Ashmore, M, Duchowski, AT & Shoemaker, G 2005, *'Efficient eye pointing with a fisheye lens.'*, Proceedings of Graphics Interface 2005, pp. 203-210.

Ball, L & Poole, A 2006, *'Eye tracking in human-computer interaction and usability research: Current status and future prospects'*, in Encyclopedia of Human Computer Interaction, Idea Group Reference.

Bates, R & Istance, H 2002, *'Zooming Interfaces! Enhancing the Performance of Eye Controlled Pointing Devices'*, Assets, pp. 119-126.

Bates, R, Castellina, E, Corno, F, Novák, P & Stepánková, O 2012, *'Beyond communication and control: environmental control and mobility by gaze'*, in Gaze interactio and applications of eye tracking, IGI Global, Hershey.

Bates, R, Donegan, M, Istance, HO, Hansen, JP & Räihä, J 2006, *'Introducing COGAIN - Communcation by Gaze Interaction'*, in J Clarkson, P Langdon, P Robinson (eds.), Designing accesible Technology, Springer, London.

Bee, N, Prendinger, H, Nakasone, A, André, E & Ishizuka, M 2006, *'Automatic preference detection by analizing the gaze 'cascade effect''*, Proceedings of the 2nd conference on Communication by Gaze Interaction COGAIN, pp. 61-64.

Benel, DCR, Ottens, DJ & Horst, R 1991, *'Use of an Eyetracking System in the Usability Laboratory'*, Proceedings of the Human Factors Society Annual Meeting, vol 35, no. 6, pp. 461-465.

Bergstrom, J & Schall, AJ 2014, *Eye Tracking in User Experience Design*, Elsevier Inc., Waltham, MA, USA.

Blascheck, T, Kurzhals, K, Raschke, M, Burch, M, Weiskopf, T & Ertl, D 2014, *'State-of-the-Art of Visualization for Eye Tracking Data'*, Eurographics Conference on Visualization (EuroVis).

Blascheck, T, Kurzhals, K, Rashcke, M, Burch, M, Weiskopf, D & Ertl, T 2014, *'State-of-the-Art of Visualization for Eye Tracking Data'*, Eurographics Conference on Visualization (EuroVis).

Bolt, R 1982, *'Eyes at the interface'*, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 360-362.

Bolt, RA 1981, *'Gaze-orchestrated dynamic windows'*, Computer Graphics, vol 15, no. 3, pp. 109-119.

Card, SK 1984, *'Visual search of computer command menus:'*, in Attention and Performance X, Control of Language Processes., Lawrance Eirlbaum Associates, Hillsdale, NJ.

Castellina, E, Corno, P & Pellegrino, P 2008, *'Integrated Speech and Gaze Control for Realistic Desktop Environments'*, Proceedings of the Eye Tracking Research & Application Symposium, ETRA.

Chennamma, HR & Yuan, X 2013, *'A survey on eye-gaze tracking techniques'*, Indian Journal of Computer Science and Engineering (IJCSE), vol 4, no. 5, pp. 388-393.

Clay, V, König, P & Koenig, S 2019, 'Eye tracking in Virtual Reality', Journal of Eye Movement Research, vol 12, no. 1.

Corno, F, Gale, A, Majaranta, P & Räihä, K-J 2010, 'Eye-based Direct Interaction for Environmental Control in Heterogeneous Smart Environments', in H Nakashima (ed.), Handbook of ambient intelligence and smart environments, Springer, New York.

Cornsweet, TN & Crane, HD 1973, *'Accurate two-dimensional eye tracker using first and fourth Purkinje images'*, Journal of the Optical Society of America, vol 63, no. 8, pp. 921-928.

Cowen, L, Ball, LJ & Delin, J 2002, *'An eye-movement analysis of web-page usability'*.

Donegan, M 2010, *'Gaze controlled technology - How far have we come? Where are we going?'*, in A Wilson, R Gow (eds.), The eye have it! The use of eye gaze to support communication, CALL Scotland, Edinburgh, UK.

Donegan, M, Morris, DJCF, Signorile, I, Chió, A, Pasian, V, Vignola, A, Buchholz, M & Holmqvist, E 2009, *'Understanding users and their needs'*, Univ Access Inf Soc, vol 8, no. 4, pp. 259–275.

Donegan, M, Oozthuizen, L, Daunys, G, R., I, Bates, R & Signorile, I 2006, *'Report on features of the different systems and development needs'*, Comunication by Gaze Interaction COGAIN.

Drewes, H & Schmidt, A 2007, *'Interacting with the Computer Using Gaze Gestures'*, Human-Computer Interaction - INTERACT, pp. 475-488.

Duchowski, AT 2017, *Eye Tracking Methodology*. Theory and Practice, 3rd edn, Springer International Publishing, London, UK.

Fairclough, S 2011, *'Physiological computing: interacting with the human nervous system'*, in M Ouwerkerk, J Westerlink, M Krans (eds.), Sensing emotions in context: the impact of context on behavioural and physiological experience measurements, Springer, Amsterdam.

Fairclough, SH, Gilleade, K (eds.) 2014, *Advances in Physiological Computing*, Springer-Verlag, London.

Fitts, P 1954, *'The information capacity of the human motor system in controlling the amplitude of movement.'*, Journal of Experimental Psychology, vol 47, pp. 381-391.

Goldberg, JH & Kotval, XP 1999, *'Computer interface evaluation using eye movements: methods and constructs'*, International Journal of Industrial Ergonomics, vol 24, pp. 631-645.

Goldberg, JH & Wichansky, AM 2002, *Eye Tracking in Usability Evaluation: A Practitioner's Guide*, Oracle Corporation, Redwood Shores, CA.

Hansen, J, Lund, H, Aoki, H & Itoh, K 2006, *'Gaze communication systems for people with ALS'*, IT University of Copenhagen.

Hansen, JP, Tørning, K, Johansen, AS, Kenji, I & Hirotaka, A 2004, *'Gaze Typing Compared with Input by Head and Hand'*, Proceedings of the 2004 symposium on eye tracking research and applications, ETRA'04., pp. 131–138.

Holmqvist, K, Nyström, M, Andersson, R, Dewhurst, R, Jarodska, H & Van de Weijer, J 2011, *Eye Tracking: A Comprehensie Guide to Methods and Measures*, 1st edn, Oxford University Press.

Hubel, DH 1988, *Eye, brain, and vision,* Scientific American Library, New York.

Hutchinson, T, White, KP, Martin, W, Reichert, K & Frey, L 1989, *'Human-Computer Interaction Using Eye-Gaze Input'*, IEEE Transactions on systems, man, and cybernetics, vol 19, no. 6, pp. 1527-1534.

Hyrskykari, A 2006, *'Eyes in attentive interfaces: experiences from creating iDict, a gaze-aware reading aid'*.

Hyrskykari, A, Majaranta, P, Aaltonen, A & Räihä, K-J 2000, *'Design Issues of iDict: a gaze-assisted translation aid'*, Proceedings of Eye Tracking Research and Applia-tions, pp. 9-14.

Istance, H & Hyrskykari, A 2012, *'Gaze-Aware Systems and Attentive Applications'*, in Gaze Interaction and Applications of Eye Tracking, IGI Global, Hershey.

Istance, H, Spinner, C & A., HP 1996, '*Providing motor impaired users with access to standard Graphical User Interface (GUI) software via eye-based interaction'*, Proceedings of the 1st European Conference on Disability, Virtual Reality and Associated Technologies, pp. 109-116.

Jacob, R & Karn, K 2003, *'Eye Tracking in Human-Computer Interaction and Usability Reseach: Ready to Deliver the Promises'*, in The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research, Elseiver Sciense BV.

Jacob, R 2003, *'Eye Movement-Based Human-Computer Interaction Techniques: Toward Non-Command Interfaces'*, Human-Computer Interaction lab, Naval Research Laboratory, Washington, D.C.

Jacob, R 2003, 'Eye Tracking in Advance Interface Design', Virtual Envirnments and Advance Interface Design.

Jacob, RJK 1990, '*What you look at is what you get: Eye movement-based interaction techniques'*, CHI '90 Proceedings, pp. 11-18.

Jacob, RJK 1995, *'Eye tracking in advance user interface design.'*, in W Barfield, T Furness (eds.), Virtual envirnment and advanced interface design, Oxford University Press, New York.

Just, MA & Carpenter, P 1976, *'Eye fixations and cognitive processes'*, Cognitive Psychology, vol 8, no. 4, pp. 441-480.

Kumar, M 2007, *'Gaze-enhanced user interface design'*, Doctoral dissertation, Standford University.

Kumar, M, Paepcke, A & Winograd, A 2007, *'EyePoint: Practical pointing and selection using gaze and keyboard'*, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 421-430.

Maglio, PP, Matlock, T, Campbell, CS, Zhai, S & Smith, BA 2000, *'Gaze and speech in attentive user interfaces'*, Proceedings of the 3rd International Conference on Advance in Multimodal Interfaces, vol 1, no. 7.

Majaranta, P & Bulling, A 2014, *'Eye Tracking and Eye-Based Human–Computer Interaction'*, in SH Fairclough, K Gilleade (eds.), Advances in Physiological Computing, Springer, London.

Majaranta, P & Donegan, M 2011, *'Introduction to Gaze Interaction'*, in Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies, Information Science Reference - Imprint of: IGI Publishing, Hershey, PA.

Meißner, M, Pfeiffer, J, Pfeiffer, T & Oppewal, H 2017, *'Combining virtual reality and mobile eye tracking to provide a naturalistic experimental environment for shopper research'*, Journal of business research, no. 100, pp. 445-458.

Morimoto, C & Mimica, M 2005, *'Eye gaze tracking techniques for interactive applications'*, Computer Vision and Image Understaning, vol 98, pp. 4-24.

Mulvey, F & Heubner, M 2012, *'Eye movements and attention'*, in Gaze interaction and applications of eye tracking, IGI Global, Hershey, PA, USA.

Mulvey, F 2012, *'Eye Anatomy, Eye Movements and Vision'*, in Gaze Interaction and Applications of Eye Tracking, IGI Global, Hershey.

Nielsen, J 1993, *'Noncommand user interfaces'*, Communications of the ACM, vol 36, no. 4, pp. 83-99.

Purwanto, D, Mardiyanto, R & Arai, K 2009, *'Electric wheelchair control with gaze direction and eye blinking'*, Artificial Life and Robotics, vol 14, pp. 397-400.

Räihä, K & O., S 2009, *'Disambiguating Ninja cursors with eye gaze'*, Proceedings of the 27th International Conference on Human Factors in Computing Systems, pp. 1411-1414.

Rayner, K & Pollatsek, A 1989, *The psycology of reading*, Prentice Hall, Englewood Cliffs, NJ.

Saito, S 1992, *'Does fatigue exist in a quantitative measurement of eye movements?'*, Ergonomis, vol 35, no. 5/6, pp. 607-615.

San Agustin, J, Mateo, J, Hansen, JP & Villanueva, A 2009, *'Evaluation of the Potential of Gaze Input for Game Interaction'*, PsychNology Journal, vol 7, no. 2, pp. 213 – 236.

Shen, J, Reingold, EM, Pomplun, M & Williams, DE 2003, *'Saccadic Selectivity During Visual Search: The Influence of Central Processing Difficulty'*, in The Mind's Eye. Cognitive and Applied Aspects of Eye Movement Research, North Holland.

Silbert, LE & Jacob, RJK 2000, *'Evaluation of eye gaze interaction'*, Proceedings of CHI, no. 2, pp. 281-288.

Skovsgaard, H, Räihä, K-J & Tall, M 2012, *'Computer Control by Gaze'*, in Gaze Interaction and Applications of Eye Tracking, IGI Global, Hershey.

Skovsgaard, HHT, Hansen, JP & Mateo, JC 2008, *'How can tiny buttons be hit using gaze only?'*, Proceedings of the 4th Conference on Communication by Gaze Interaction COGAIN, pp. 38-42.

Špakov, O & Miniotas, D 2005, *'Gaze-based selection of standard-size menu items'*, Proceedings of the 7th International Conference on Multimodal Interfaces, vol 4, no. 6, pp. 124-128.

Starker, I & Bold, R 1990, *'A gaze-responsive self-disclosing display'*, CHI '90 Proceedings.

Tong, H & Fisher, R 1984, *'Progress Report on an Eye-Slaved Area-of-Interest Visual Display.'*, Air Force Human Resources Laboratory.

Tsui, CSL, Jia, P, Gan, JQ, Hu, H & Yuan, K 2007, *'EMG-based Hands-free wheelchair control with EOG attention shift detection'*, Proceedings of the 2007 IEEE International Conference on Robotics and Biomimetics, Sanya, China.

Tuisku, O, Bates, R, Stepankova, O, Fejtova, M, Novák, P, Istance, H & Majaranta, P 2008, *'A survey of existing 'de-facto' standards and systems of gaze based mobility control'*, Communication by Gaze Interaction (COGAIN), Project IST-2003-511598.

Vertegaal, R & Shell, J 2008, *'Attentive user interfaces: the surveillance and sousveillance of gaze-aware objects'*, Social Science Information, vol 47, no. 3, pp. 275-298.

Vertegaal, R 2003, *'Attentive user interfaces'*, Communications of the ACM, vol 46, no. 3, pp. 30-33.

Vesterby, T, Voss, J, Hansen, JP, Glanstrup, A, Hansen, DW & Rudolph, M 2005, *'Gaze-guided viewing of interactive movies'*, Digital Creativity, vol 16, no. 4, pp. 193-204.

Wade, N & Tatler, B 2005, *The Moving Tablet of the Eye: The Origins of Modern Eye Movement Research*, USA.

Wanluk, N, Juhong, A, Visitsattapongse, S & Pintavirooj, C 2016, *'Smart wheelchair based on eye tracking'*, Biomedical engineering international conference.

Ware, C & Mikaelian, H 1987, *'An Evaluation of an Eye Tracker as a Device for Computer Input'*, Proceedings of CHI, pp. 183-188.

Yarbus, AL 1967, *Eye Movements and Vision*, Plenum Press, New York.

Yonezawa, T, Yamazoe, H, Utsumi, A & Abe, S 2008, *'Evaluations of interactive guideboard with gaze-communicative stuffed-toy robot.'*, Proceedings of the 4th Conference on Comunication by Gaze Interaction COGAIN, pp. 53-58.

Zhai, S 2003, *'What's in the eyes for attentive input'*, Communications of the ACM, vol 46, no. 3, pp. 34-39.

Zhai, S, Morimoto, C & Ihde, S 1999, '*Manual and gaze input cascaded (MAGIC) pointing*', Proceeding of the CHI '99 Conference on Human Factors in Computing Systems: The CHI is the Limit, Pittsburgh.

Zhang, Q, Imamiya, A, Go, K & Mao, X 2004, *'Overriding errors in a speech and gaze multimodal architecture'*, Proceedings of the 2004 International Conference on Intelligent User Interfaces.

Zülch, G & Stowasser, S 2003, *'Eye Tracking for Evaluation Industrial Human-Computer Interfaces', in The Mind's Eye*. Cognitive and Applied Aspects of Eye Movement Reseach, North Holland.