CSCI 8810: Eye Trackers

Presentation Report

Prof. Arabnia

Sina Solaimanpour

March, 2015

Contents

I. Introduction

To start our discussion about Eye Trackers, we first need to address the motivation behind the act of tracking human eyes. In other words, Why eye tracking is important? Simply put, we move our eyes to bring a particular portion of the visible field of view into high resolution so that we may see in fine detail whatever is at the central direction of gaze. Most often we also divert our attention to that point so that we can focus our concentration on the object or region of interest. Thus, we may presume that if we can track someones eye movements, we can follow along the path of attention deployed by the observer. This may give us some insight into what the observer found interesting.
In this paper, a model for human vision will be discussed. Then, we will talk a little bit about human eye's anatomy and from there, we will start talking about some classic and also some state of the art examples of eye trackers. Finally, we will end our discussion with introducing an eye tracking system called Starburst.

II. Bottom-up Vision Model

Vision might behave in a cyclical process composed of the following steps:

1. Given a stimulus, such as an image, the entire scene is first seen mostly in parallel through peripheral vision and thus mostly at low resolution. At this stage, interesting features may *pop out* in the field of view for further detailed inspection.
2. Attention is thus turned off or disengaged from the foveal location and the eyes are quickly repositioned to the first region that attracted attention.
3. Once the eyes complete their movement, the fovea is now directed at the region of interest, and attention is now engaged to perceive the feature under inspection at high resolution.

This is a bottom-up model or concept of visual attention. The bottom-up model is at least correct in the sense that it can be said to be a component of natural human vision. The bottom-up view of visual attention is, however, incomplete and has some shortcomings. There are several key questions that are not addressed in this model.

1. Exactly what types of features attract attention? Human Visual System responds strongly to some types of stimuli, and weakly to others.
2. Is the visual stimulus solely responsible for attracting attention? Would we ever need the capability of making voluntary eye movements?
3. What is the link between attention and eye movements? Humans can voluntarily dissociate attention from the foveal direction of gaze. In fact, astronomers do this regularly to detect faint constellations with the naked eye by looking *Off The Fovea*.

Besides these shortcomings, the bottom-up model forms a powerful basis for computational models of visual search and has shown to be very useful specially in the Image Processing field of study.

III. The Eye

"The human eye is an organ that reacts to light and has several purposes."[1] In fact, human eye has a very complicated structure which will not fit into our discussion to talk about it in detail but, there are some parts of the eye which are essential for our study and these parts will be discussed here. You can see an overview of the human eye structure in **Figure 1**.
In the following sub-sections, we will be discussing about Cornea, Pupil, Iris and Sclera.

[1]Wikipedia: `http://en.wikipedia.org/wiki/Human_eye`

III. The Eye continued on next page...

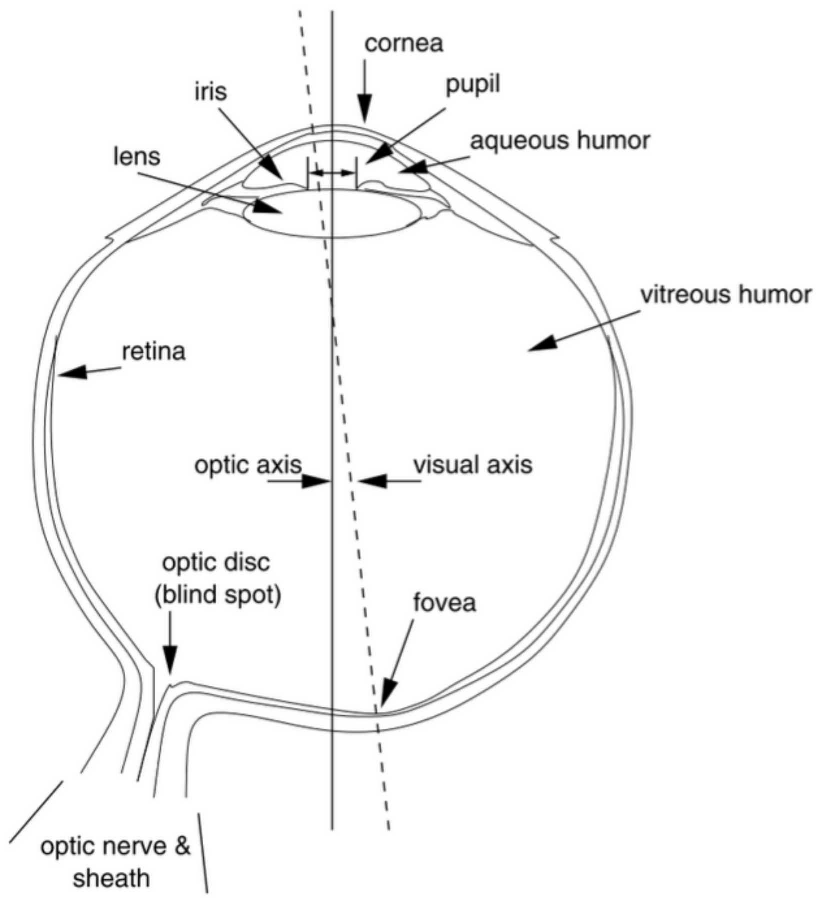

Figure 1: Human Eye structure overview.

(a) Cornea

"The cornea is the transparent front part of the eye that covers the iris, pupil, and anterior chamber."[2] Cornea acts like the outer crystal surface of a watch, protecting the inner components of the eye. Cornea itself is also made out of 5 layers listed below.

1. Corneal epithelium
2. Bowman's layer
3. Corneal stroma
4. Descemet's membrane
5. Corneal endothelium

(b) Pupil

The apparently black circular opening in the center of the iris of the eye, through which light passes to the retina. It appears black because light rays entering the pupil are either absorbed by the tissues inside the eye directly, or absorbed after diffuse reflections within the eye that mostly miss exiting the narrow pupil.

[2]Wikipedia: `http://en.wikipedia.org/wiki/Cornea`

(c) Iris

"The iris is a thin, circular structure in the eye, responsible for controlling the diameter and size of the pupil and thus the amount of light reaching the retina." [3] The color of the iris gives the eye its color. In optical terms, the pupil is the eye's aperture and the iris is the diaphragm that serves as the aperture stop.

(d) Sclera

"The sclera , also known as the white of the eye, is the opaque, fibrous, protective, outer layer of the eye containing collagen and elastic fiber." [4] In humans the whole sclera is white, contrasting with the coloured iris. The human eye is relatively rare for having an iris that is small enough for its position to be plainly visible against the sclera. This makes it easier for one individual to infer where another individual is looking, and the cooperative eye hypothesis suggests this has evolved as a method of nonverbal communication. Also, this fact is usually used in eye tracking systems.

IV. Eye Trackers

Now, we come up to this question, what is an eye tracker? Any measurement device used for measuring eye position and movements is commonly known as an Eye Tracker. Eye trackers are used in research on the visual system, in psychology, in cognitive linguistics, marketing, as an input device for human computer interaction, and in product design. There are generally two types of Eye Trackers:

- Those that measure the position of the eye relative to the head
- Those that measure the orientation of the eye in space

The second type of the trackers are typically used when we want to identify elements in a visual scene. The first usage of eye trackers recorded in human history was done in 1879 in Paris by Louis mile Javal. He observed that reading does not involve a smooth sweeping of the eyes along the text, as previously assumed, but a series of short stops (called fixations) and quick saccades. This observation raised important questions about reading, which were explored during the 1900s: On which words do the eyes stop? For how long? When does it regress back to already seen words? **Figure 2** shows the process of reading a text which was resulted from the above mentioned experiment.

Another example of classic eye tracking is done in the 1950s by Alfred L. Yarbus which is considered a very important eye tracking research. He showed the task given to a subject has a very large influence on the subject's eye movement. He asked the subjects to determine some aspects of the persons entering a room like, Are the people in the image related? What are they wearing?

He also wrote about the relation between fixations and interest: "All the records ... show conclusively that the character of the eye movement is either completely independent of or only very slightly dependent on the material of the picture and how it was made, provided that it is flat or nearly flat." The cyclical pattern in the examination of pictures "is dependent not only on what is shown on the picture, but also on the problem facing the observer and the information that he hopes to gain from the picture." [5] **Figure 3** shows the results from Yarbus's experiments.

To mention other recent usages of eye trackers, we can think of the Microsoft's HoloLens and using eye trackers in many other daily aspects of human like controlling an interface, placing advertisements in the right place and many other examples.

Now let's see a taxonomy of eye trackers.

[3] Wikipedia: `http://en.wikipedia.org/wiki/Iris_(anatomy)`
[4] Wikipedia: `http://en.wikipedia.org/wiki/Sclera`
[5] Yarbus 1967, p. 194

IV. Eye Trackers continued on next page...

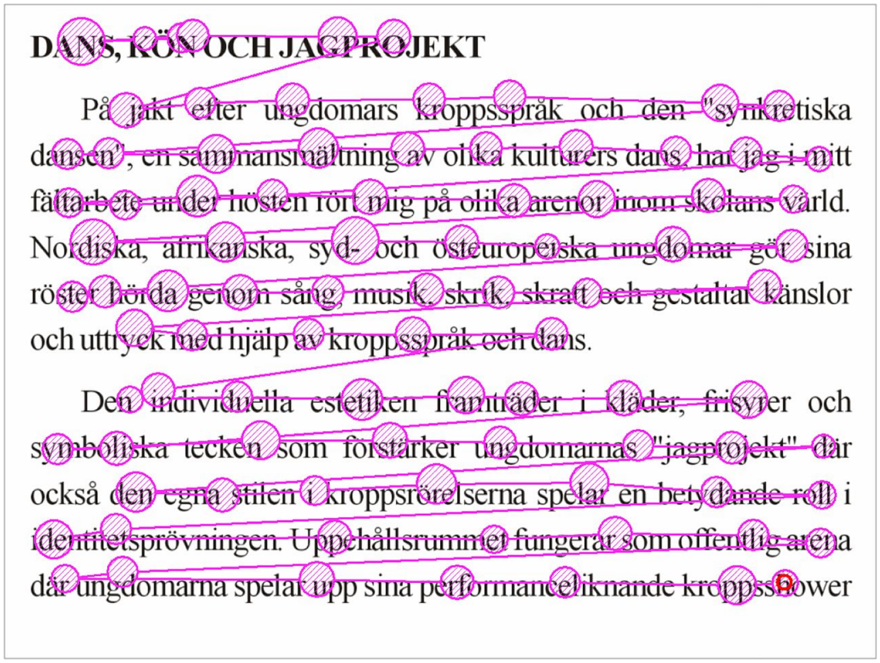

Figure 2: The result from Louis mile Javal's research on reading a text.

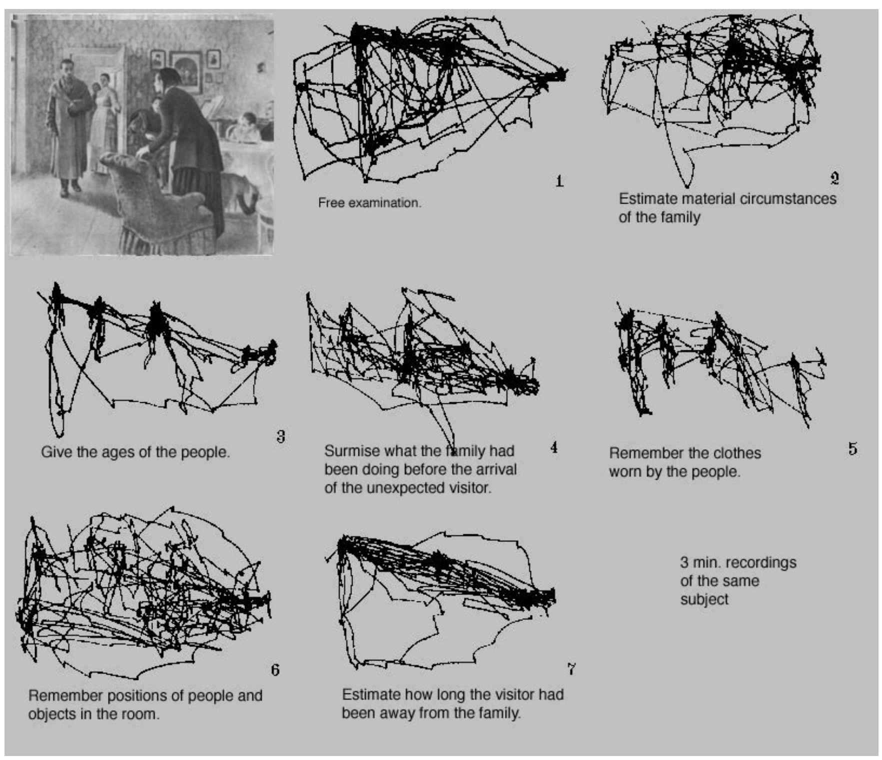

Figure 3: The result from Yarbus's experiments.

(a) Electro-oculography (EOG)

It relies on measurement of the skins electric potential differences, of electrodes placed around the ocular cavity. During the mid-1970s, this technique was the most widely applied eye movement method.It measures eye movements relative to head position. An example of this class is shown in **Figure 4**.

Figure 4: Electro-oculography (EOG)

(b) Scleral Contact Lens/Search Coil

One of the most precise eye movement measurement methods which involves attaching a mechanical or optical reference object mounted on a contact lens which is then worn directly on the eye. As the problems for this method, it is considered the most intrusive method. An example of this class is shown in **Figure 5**.

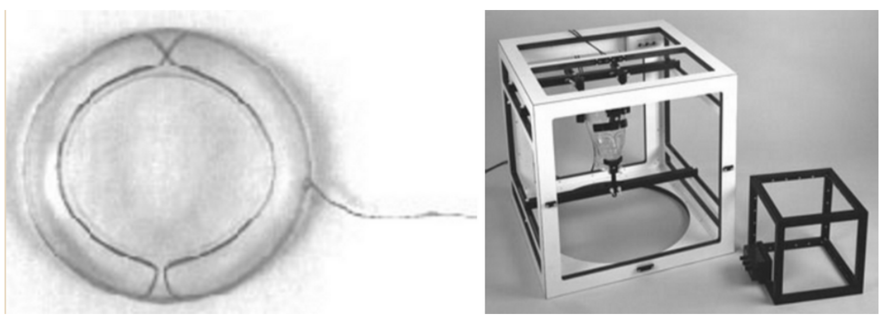

Figure 5: Scleral Contact Lens/Search Coil

(c) Photo-OculoGraphy (POG) or Video-OculoGraphy (VOG)

These are two groups of eye trackers involving the measurement of distinguishable features of the eyes under rotation / translation. Examples of these features are,

- The apparent shape of the pupil
- The position of the limbus (the iris-sclera boundary)

These techniques are grouped together because they often do not provide point of regard measurement.

(d) Video-Based Combined Pupil/Corneal Reflection

To measure the Point of Regard either,

- Subjects head should be fixed, or

IV. Eye Trackers [(d) Video-Based Combined Pupil/Corneal Reflection] continued on next page.

- Multiple ocular features must be measured in order to disambiguate head movement from eye rotation

Two features that can help find the point of regard are,

- The Corneal Reflection
- The Pupil Center

Most of the Video-based trackers utilize relatively inexpensive cameras and image processing hardware to compute the point of regard in real-time. This technique is most suitable to be used in interactive applications.
We mentioned The Corneal Reflections or also called The Purkinje Images. Purkinje Images are reflections of objects from the structure of the eye. At least four Purkinje images are usually visible:

1. The first Purkinje image (P1) is the reflection from the outer surface of the cornea
2. The second Purkinje image (P2) is the reflection from the inner surface of the cornea
3. The third Purkinje image (P3) is the reflection from the outer (anterior) surface of the lens
4. The fourth Purkinje image (P4) is the reflection from the inner (posterior) surface of the lens

We need two points of reference on the eye to separate eye movements from head movements. As the positional difference between the pupil center and corneal reflection **changes with pure eye rotation** but **remains relatively constant with minor head movements** most Video-Based eye trackers measure the First and sometimes the Fourth (in generation-V systems) Corneal Reflections and based on these measurements, they separate translational and rotational eye movements. **Figure 6** shows the first, third and the fourth Purkinje Images.

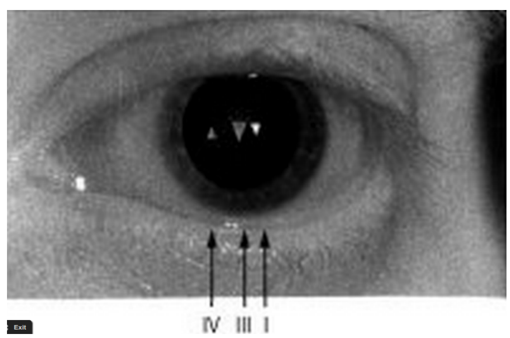

Figure 6: the first, third and the fourth Purkinje Images

V. Starburst

The primary obstacle of integrating eye movements into todays interfaces is the availability of a reliable, low-cost open-source eye-tracking system. Towards making such a system available to interface designers, a hybrid eye tracking algorithm that integrates feature-based and model-based approaches has been introduced in this study is made available in an open-source package. This algorithm is referred to as starburst because of the novel way in which pupil features are detected. This starburst algorithm is more accurate than pure feature-based approaches yet is significantly less time consuming than pure model-based approaches.
The bird eye view of the algorithm is as depicted in **Figure 7**. In the following subsections, the different parts of this algorithm are explained in more detail.

V. Starburst continued on next page...

Input: Eye image, Scene image
Output: Point of gaze
Procedure:
Detect the corneal reflection
Localize the corneal reflection
Remove the corneal reflection
Iterative detection of candidate feature points
Apply RANSAC to find feature point consensus set
Determine best-fitting ellipse using consensus set
Model-based optimization of ellipse parameters
Apply calibration to estimate point of gaze

Figure 7: Starburst Algorithm

(a) Noise Reduction

Due to the use of cheap consumer grade CCD cameras, we need to start by reducing the noise in the images. There Are two types of noise in the images:

- Shot noise: Reduced by using a 5x5 Gaussian filter with a standard deviation of 2 pixels.
- Line noise: This noise is reduced by applying a normalization factor (C) line by line to shift the mean intensity of the line to the running average derived from the previous frame.

An example of the effect which the noise reduction process has on the image can be seen in **Figure 8**.

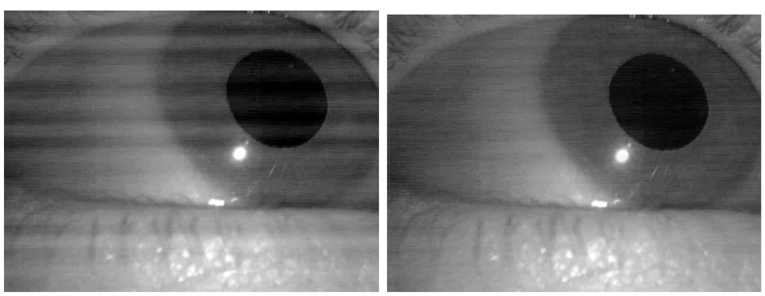

Figure 8: Noise Reduction effect

(b) Corneal reflection detection and localization

(c) Corneal reflection removal

(d) Pupil contour detection

(e) Ellipse fitting

(f) Model-based optimization

(g) Calibration