Correlation peak analysis applied to a sequence of images using two different filters for eye tracking model

Verónica A. Patrón A, Josué Álvarez Borrego b, Ángel Coronel Beltrán b


ABSTRACT

Eye tracking has many useful applications that range from biometrics to face recognition and human-computer interaction. The analysis of the characteristics of the eyes has become one of the methods to accomplish the location of the eyes and the tracking of the point of gaze. Characteristics such as the contrast between the iris and the sclera, the shape, and distribution of colors and dark/light zones in the area are the starting point for these analyses. In this work, the focus will be on the contrast between the iris and the sclera, performing a correlation in the frequency domain. The images are acquired with an ordinary camera, which with were taken images of thirty-one volunteers. The reference image is an image of the subjects looking to a point in front of them at 0° angle. Then sequences of images are taken with the subject looking at different angles. These images are processed in MATLAB, obtaining the maximum correlation peak for each image, using two different filters. Each filter were analyzed and then one was selected, which is the filter that gives the best performance in terms of the utility of the data, which is displayed in graphs that shows the decay of the correlation peak as the eye moves progressively at different angle. This data will be used to obtain a mathematical model or function that establishes a relationship between the angle of vision (AOV) and the maximum correlation peak (MCP). This model will be tested using different input images from other subject not contained in the initial database, being able to predict angle of vision using the maximum correlation peak data.

Keywords: Filtering, Fourier analysis, correlation, eye location, eye tracking.

1. INTRODUCTION

Eye tracking has become a challenging area of research, especially in determining the point of gaze in a video sequence. Regardless of major advances over the last 30 years, eye detection and location still face many difficulties mostly due to factors such as eye individuality, occlusions, scale variability, location and illumination conditions[1]. However many systems have been developed for biometrics applications, such as iris recognition, early detection of retinopathies and other eye conditions. Eye tracking and gaze estimation are used in other applications, from human-computer interaction, to the study of neurological diseases, but also for commercial use and web navigation.

Research is also focused on improving image processing tools for eye location, much of it going towards understanding the mechanism and dynamics of eye rotation. There are three major aspects to consider in eye detection: 1) detecting the presence of eyes in an image, 2) interpretation of the eyes position within the image and, 3) for video sequence, frame by frame eye tracking and gaze estimation.

In this work, the tracking of the eyes movement is analyzed using the digital filtering response of a sequence of images of the eyes moving from left to right. The sample space consisted on images from thirty-one subjects, taken under the same configuration and lighting conditions. The digital filters tested in this work were a non-linear filter and a phase filter, both computed in MATLAB. The response was measured using the MCP between the filter image (reference) and the image sequence, and then aim to obtain the mathematical model that describes the relationship between this MCP and AOV. Using the information from each filter, the one with no positive contribution to the model was discarded, and a mean from all samples of the filter with the best response was obtained.
2. SYSTEM SETUP

For the acquisition of the images to create the database, a Nikon Coolpix S6000 camera was used, mounted on a laboratory table. At 70cm from the camera, a chin support for the subject was placed facing the camera. On the walls marks from 0° to 45° were drawn, to the left and to the right, spaced in 5° each (19 marks total). Three frames per mark per subject were taken.

![Diagram of system setup](image)

Figure 1. a) Configuration of the system used to obtain the images. The camera is mounted on a floated table, 15cm above the surface of the table. The subject is asked to look at each mark on the wall, from 0° to 45° to both directions, b) image sample.

Later the images were preprocessed. First, using Adobe Photoshop Elements 13, from the image with the complete face, the eye area was cut, of both eyes separate. Figure 2 shows a sequence of these images of the right eye, going from 0° to 45° to the right.

![Sequence of images](image)

Figure 2. Sequence of these images of the right eye, going from 0° to 45° to the right.

Figure 3 shows the complete database, consisting of 31 subjects, 15 males and 16 females, ages 20 to 60 years old. From the database, 30 of the subjects were used to create a model, and the other one was used for testing.

The preprocessing also consisted on improving the contrast of the image by implementing functions of the MATLAB image processing toolbox [2]. Then the images, converted to grayscale, were subjected to a threshold, calculated empirically. Then the threshold image was processed with a Sobel Edge detector, then dilated and eroded.
Figure 3. Preprocessing of the images: a) Original image; b) threshold image; c) Sobel edge detector image; d) Image after the process of dilation and erosion.

Figure 4. Image database. The model was created using thirty of the subjects, leaving one as a testing subject.

3. IMAGE PROCESSING

Every sequence of preprocessed images was then processed in MATLAB. Two different digital filters were used, based on Fourier analysis using the FFT and the correlation theorem that states as follows:

\[ TF\{\int f(x) g^*(x - x_0) dx\} = F(u)G^*(u), \]

where \( F(u) \) and \( G^*(u) \) are the Fourier transform of \( f(x) \) and the conjugate of the Fourier transform of \( g(x) \) respectively.

The filter or transfer function for all image sequences was the reference image (0° image). For each subject, all images in the sequence were correlated to the reference image. The digital filters used were a non-linear filter and a phase filter. For the phase filter the correlation was computed using only the complex part of the FFT.
In the case of the non-linear filter the expression is

\[ (u, v) = e^{j\theta(u,v)}. \]  \hspace{1cm} (2)

The non-linearity factor was varied from 0.1 to 0.9, obtaining the best result at \( k = 0.4 \).

From each filter we obtained the MCP. Figure 4 shows the peak when the correlation spectrum is graphed.

![Correlation spectrum](image)

**Figure 5.** Correlation spectrum. We observe the MCP with a) non-linear filter and b) phase only filter.

![Filter graphs](image)

**Figure 6.** The filter graphs show the relationship between the MCP and AOV for both filters for one subject.

## 4. RESULTS

Given the comparison of the relationship between the MCP and AOV for both filters, the best results came from the non-linear filter, which curve is smoother and with greater slope.

With the non-linear filter we developed a model, taking the results from 30 subjects, with normalized data. Figure 7 shows the data with a Distance Weighted Least Squares fit, which gave an even smoother curve and greater slope.
Figure 7. Statistical model that gives us the relation MCP with AOV. This data only shows the results for the right eye of every subject.

Once the model was obtained, the data of one subject was used for testing said model. The figure 8 shows the results of the test. The arrows point at the values given by the filtering of the test subject, this data was input in the model to see if the prediction were accurate.

Figure 8. Test run of the model. The level of accuracy is not as high as expected, but the data points are in range for every AOV.
In the figure 8 we can see the data dispersion at 5°. This might be due to the FFT being invariant to position, so that what it is seen is changes in shape. The iris changes its shape while moving, going from a circular shape to a more elliptical the farther from the center it gets.

5. CONCLUSIONS AND FUTURE WORK

This work presented a method to estimate gaze by creating a model that establishes a relationship between the angle of vision (AOV) and a maximum correlation peak (MCP). This MCP was obtained by digital filtering, using FFT. Two filters were tested, a phase only filter and a non-linear filter, being the latter the one with better results and the one used to create a model of AOV given by a MCP. The model was tested using data from one subject not included in the model. The results showed that the dispersion exhibited by the data does not allow a high accurate prediction of angle of vision. This is due inherent characteristics of a filter based on the FFT, which is position invariant. Therefore the data established the relationship based on changes in shape of the iris that happens when it moves, changing from circular to elliptical. Our goal is to create a model that predicts the angle of vision, but further improvements to the model must be made in order to accomplish a more accurate and reliable result. Also more data should be integrated to the creation of the model, with more test subjects as well.

In addition to the prediction of point of gaze, another application of the model could be as a calibration device, having an alternative to detect the presence of eyes before starting the tracking process.

6. REFERENCES