Utilizing image processing in eye tracking

1- Introduction

Eye Tracking project is a Matlab program that localizes the eye centers and plots the estimated circle on the screen using Viola-Jones algorithim for face detection and finding the approximated center of the eye using some geometry and filters. The program is thought to work in real-time with any laptop's camera, such those you can find on PCs, tablets and smartphones. This project can be useful in many applications like alert systems for sleeping drivers if he closes his eyes, rather than blinking, or tracking the eye for some games with human interface and so on.

2- Implementation

There are some other tries to track the eye using image processing tools, like template matching using cross correleation, but they were easier but not time efficient, as its computational time was, relatively, high to make it a real time eye tracking (using live acquisition from the camera).

So the method implemented was as followed **(for each received frame)**:

a- Face Detecting using Viola-Jones algorithm

Viola Jones algorithm was a revolution in image processing as they proposed the concept of integral image to enhance the time consumed in detection, so cascade object detector was utilized to acquire the face of the image, and in our application we needed to take only one face of the image, which is the largest face available in the video in case there are more than one face.

Figure 1: Cropped face

b- By using the size of the face cropped out of the image, by using constant numerical ratios has been acquired by try and error from the face to get the eye left and the right eye.

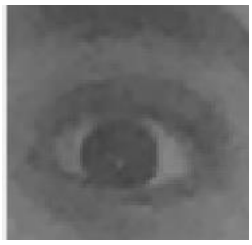

Figure 2: Cropped left eye

Figure 3: Cropped right eye

c- After that, Histogram equalization is performed to each eye and binarizing it - changing it to black and white - with a threshold of 0.2 to make the majority of the image is white and the blackest region is kept which is the pupil of the eye we are searching for.

Result of this operation is:

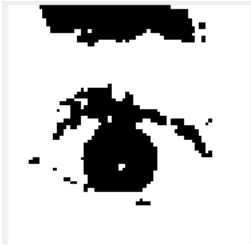

Figure 4: Operation done for left eye

Figure 5: Operation done for right eye

d- An area opening is applied for each eye to remove noise resulting from the center of the pupil which is, most of the time, very bright.

Figure 6: Median filter for black and white left eye

Figure 7: Median filter for black and white right eye

e- After that, Hough transform for circular detection is applied to get the biggest black circle in each image and get its radius and center, hence the pupil is detected.

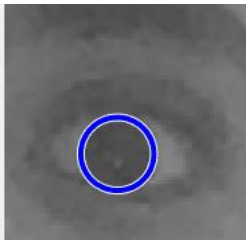

Figure 8: Circle drawn based on hough transform detection for left eye

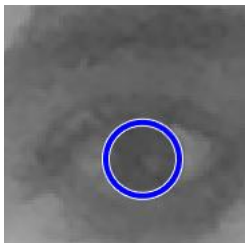

Figure 9: Circle drawn based on hough transform detection for right eye

f- The last step is after getting the radius and the center of the pupil with respect to each eye the offset is calculated from each eye box and added to the off set of each eye box with respect to the face box.

Figure 10: eye detecting in the face box

and after these steps all these offsets are added to the coordinates of the cropped face in order to draw the circle in the right place for each eye in the whole frame.

Figure 11: eye detecting with respect to the whole frame.

3- Analysis

This method has some limitations that the exposure of light must be good, the face must not be rotated or tilted, and it may not produce accurate results of the pupil if you are not opening your eye completely.

On the other hand, it can detect the pupil even if you are not looking directly into the camera, it is very time efficient as for each frame the computing time on my laptop is approximately 1.1 second which is fairly good for live detection which was implemented as well for a live video acquisition and it worked well.

4- Recommendation for enhancing this code

a- More optimization for the code for hough transform, as an example to make the range of radii smaller and limit the code for a certain application that we know that in this depth between the camera and the user the eye pupil will vary in a low radius range which will enhance the computation time a lot.
b- More morphological operations and filtration may be good to be implemented on the eyes to produce more accurate results.
c- More training for the cascade object detector to detect the face if it is tilted by angle of even rotated for any reason, if required in a certain application, to detect the face and also get the eye from it, as this code used the conventional cascade detector which was not trained enough for acquiring tilted faces.