

Eye Detection and Tracking in Image with Complex Background

Mohammadali Azimi Kashani, Mahdi Mollaei Arani, Mohammad Reza Ramezanpour

Department of Computer Science & Research Islamic Azad University Shoushtar Branch
Shoushtar, Iran

Azimi.kashani@gmail.com

ABSTRACT

To detect and track eye images with complex background, distinctive features of user eye are used. Generally, an eye-tracking and detection system can be divided into four steps: Face detection, eye region detection, pupil detection and eye tracking. To find the position of pupil, first, face region must be separated from the rest of the image, this will cause the images background to be non effective in our next steps. We used the horizontal projection obtained from face region, to separate a region of face containing eyes and eyebrow. This will result in decreasing the computational complexity and ignoring some factors such as bread. Finally, by proposed algorithm we will obtain the pupil position. In the next step, we perform eye tracking. In the proposed method, eye detection and tracking are applied on testing sets, gathered from different images of face data with complex backgrounds. Experiments indicate correct detection rate of 94.9%, which is indicative of the method's superiority and high robustness.

Keywords: *eye detection; eye tracking; kallman filter*

I. INTRODUCTION

Automatic tracking of eyes and gaze direction is an interesting topic in computer vision with its application in biometric, security, intelligent human-computer interfaces, and driver's drowsiness detection system. Localization and extraction of eyes are operations requisite for solving problem. In [1] eye localization methods have been classified into five main categories:(1) Shape-based Approaches which described by its shape, which includes the iris and pupil contours and the exterior shape of the eye (eyelids).(2) Feature-Based Shape Methods which explore the characteristics of the human eye to identify a set of distinctive features around the eyes. The limbus, pupil (dark/bright pupil images) and cornea reflections are common features used for eye localization.(3) Appearance-Based Methods which detect and track eyes directly, based on the photometric appearance as characterized by the color distribution or filter responses of the eye and its surroundings. These methods are independent of the actual object of interest and are in principle capable of modeling other objects besides eyes. (4) Hybrid Models which aim at combining the advantages of different eye-models within a single system to overcome their respective shortcomings. (5) Other Methods which employing temporal information and active light. In [2], skin-color filter is used to extract face. The eye position is gained by gradient characteristic projection and corresponding conditions setting. In [3] use the Haar-like features to detect the eye. This method trains classifiers with a few thousands of sample view images of object and construct a cascade of classifiers to detect eye rapidly.

II. SKIN COLOR SEGMENTATION

Since color processing is faster than other face features and color data remain the same as the face turns, skin

color will be used to detect face region. One of the choices for the color space without illumination is the chromatic color space YCbCr because the light intensity can be separated from the color space. Hence, the image is first converted

from RGB color space into YCbCr. At the time of operation, only Cb and distribution. RGB color space changes into YCbCr by equation (1)[6]:

$$\begin{cases} Y = 0.299 R - 0.587 G - 0.114 B \\ Cb = R - Y \\ Cr = B - Y \end{cases} \quad (1)$$

A. Skin color Model Construction

1) Simple Gaussian Model

A likelihood function could be defined based on the skin color model. This function can determine which skin region a pixel belongs to. Skin color distribution may be reconstructed in YCBCR color space through a normal (Gaussian) distribution with \sum and μ parameters. The multi variable normal distribution of incidental D-dimensional vector \mathcal{X} is defined as equation (2):

$$N(x, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right] \quad (2)$$

Where μ denotes the mean (average) vector and Σ stands for the covariance matrix of vector \mathcal{X} . The model parameters are obtained from the instructional data according to equation (3):

$$\mu = \frac{1}{n} \sum_{i=1}^n c_i \quad (3)$$

$$\Sigma = \frac{1}{n-1} \sum_{i=1}^n (c_i - \mu)(c_i - \mu)^T$$

The Mahalanobis distance from the color vector C to vector μ and the measured covariance Σ denote the belonging of pixel C to the skin color.

2) Gaussian Mixture Model

The simple Gaussian model works well in detecting features and distinguishing between different clusters. However, severe calculation errors may occur in this model when the respective components do not vary evenly compared to the mean. Generally stated, Gaussian mixture model is a good approximation of phenomena with natural distribution. In this model, some Gaussian functions are used to model skin color [4]. Gaussian mixture which equals total weight μ is part of Gaussian distribution. The possibility of the feature vector x to λ_i model is calculated through equation (4).

$$f(x | \lambda_i) = \sum_{i=1}^M p_i g_i(x) \quad (4)$$

$g_i(x)$ Denotes the possibility of observing vector x . This observation results from the i th mixture, and p_i is the weight of the i th mixture. Generally, each of the mean vector (μ_i), covariance matrix (Σ_i) and the mixture weight (p_i) equation (5). Each of the above parameters represents a skin color.

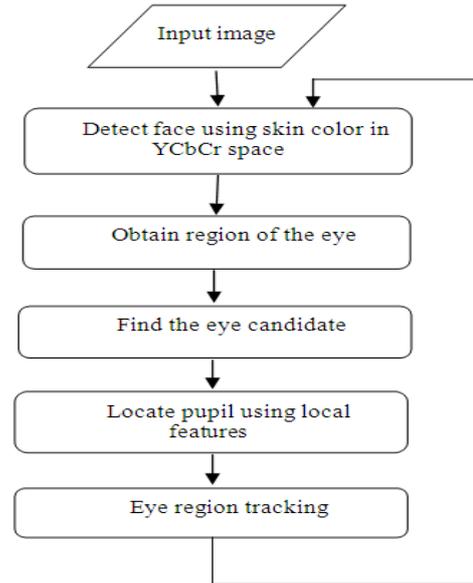
$$\lambda_i = (\mu_i, \Sigma_i, p_i) \quad (5)$$

III. PROPOSED METHOD

In the proposed method, color information in YCBCR environment and Gaussian mixture model has been used. Since the region above the eyebrows and below the eyes is brighter, the eye region is cropped. This region, which starts above the eyebrows and extends to the region below the eyelids, includes two eyebrows, two eyes, and their characteristic parts like eyelids and eye corners. Applying Harris method, we roughly locate the pupil. Having identified the pupil in the initial frame, we detect the eye region by kallman filtering. Fig 1 illustrates the proposed method.

A. Extracting Eye Region

One of the main face regions with a lot of feature is the eye region. Starting above the eyebrows and extending below the eyelids, this region incorporates two eyebrows, two eyes, and their characteristic parts like eyelids and eye corners. To obtain the eye region, the face region is cropped out of the image and then the horizontal projection of the face is worked out.



Proposed algorithm for eye detection and tracking

The horizontal projection for an image of $N_{Rows} \times N_{Columns}$ is a sum of pixel intensities in each line. Equation (6)

$$vp = \{vp_x | 1 \leq x \leq N_{Rows}\}$$

$$vp_x = \sum_{y=1}^{N_{Columns}} f(x, y) \quad (6)$$

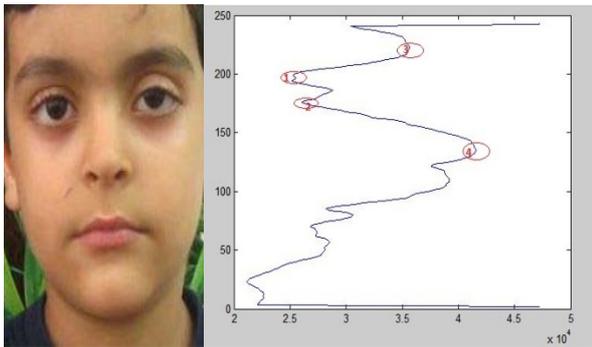
Fig2 is an image of the face along with its horizontal projection. As observed in this figure, the uppermost valley since the region above the eyebrows and below the eyes is brighter, the eye region is cropped. This region, which starts above the eyebrows and extends to the region below the eyelids, includes two eyebrows, two eyes, and their characteristic parts like eyelids and eye corners. Applying Harris method, we roughly locate the pupil. Having identified the pupil in the initial frame, we detect the eye region by kallman filtering. Fig 1 illustrates the proposed method.

B. Eye detection algorithm

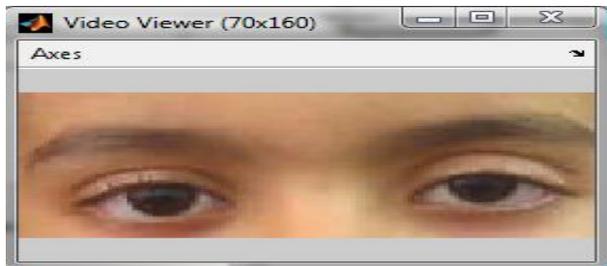
In this stage, eye candidates are extracted through the following procedure: First, the values of A, B, C are identified through equation (7).

$$A = (I_x)^2, B = (I_y)^2, C = (I_x I_y) \quad (7)$$

Where I_x, I_y denote the image gradients in x, y directions, respectively.



Horizontal projection of face: the horizontal axis illustrates the total pixels of each line of the face and the vertical axis the number of each line.



Extraction of the eye region obtained from the horizontal projection of the face.

Then R is obtained from equation (8):

$$R = Det(M) - KTr^2(M) \quad (8)$$

$$Tr = \alpha + \beta = A + B, Det(M) = \alpha\beta = AB - C^2$$

The points with the highest values of R in the specified region are selected as the eye candidate's. In the next stage, the eye region is well detected among these points. Color entropy in the eye region is used to eliminate the irrelevant candidates

C. Eliminating Candidates Irrelevant to the Eye

The image of eye has a lot of features. Variations in lightness levels and the form of eye component

complicate the detection of these features. One factor useful for locating the eye is the irregular variations in lightness levels of sclera (eye's white area). The blood vessels in this area contribute to this entropy on the surface of the sclera. To ensure correct detection of pupil, some points with the highest values of r are specified. These points are considered as the pupil candidates. To eliminate the candidates irrelevant to the pupil, color entropy is considered for each candidate, and color entropy is calculated. The probability of entropy is obtained from equation (9).

$$H(S) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i) \quad (9)$$

Where $p(x_i)$ denotes the probability of pixel x_i , and n stands for the number of pixels in each section. The probability function of each pixel, $p(x_i)$, is the frequency of the pixel color intensity in the selected area divided by the total frequency of pixels. Two regions with the highest entropy in the eye's candidate are decided as the area including the pupil. The center of these areas is regarded as the center of the pupil. fig.4.



Identifying the pupil by the proposed method

D. Eye tracking

In most application, eye tracking or face tracking follows eye detection. To cover head movement, the eye itself is not tracked. Rather, part of the face including two eyes is tracked. Among the method available, kallman filtering is faster and computationally less intricate, and therefore will be used for detecting the eye region. Kallman filter predicts the next time state vector based on the movement information in the previous stages. For detection, the state vector is considered $X = (x, y, v_x, v_y)$. x, y Are the coordinates of the center of the region including the eye, v_x, v_y the coordinate difference of the center of the eye region lies in two successive frames. Kallman filter estimates the state vector at $t + 1$ from the time state vector t . Equation (10).

$$X_{t+1} = \Phi X_t \quad (10)$$

The observation vector is supposed as $z = (x, y)$. The equation (11) relating the state vector and observation vector is:

$$Z_{t+1} = H X_{t+1} \quad (11)$$

Where Φ Matrix is as follows:

$$\Phi = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

H Is the matrix that connects X, Z vectors and is defined as:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

The proposed method will be put into practice in a region selected based on the estimated observation matrix (Z_{t+1}). This narrows the operational domain and requires fewer calculations. After the exact coordinates of the region including the eye have been obtained, the new state vector is derived from equation (12) using the derived state vector. Kallman filtering corrects its parameters for estimation in the following stages.

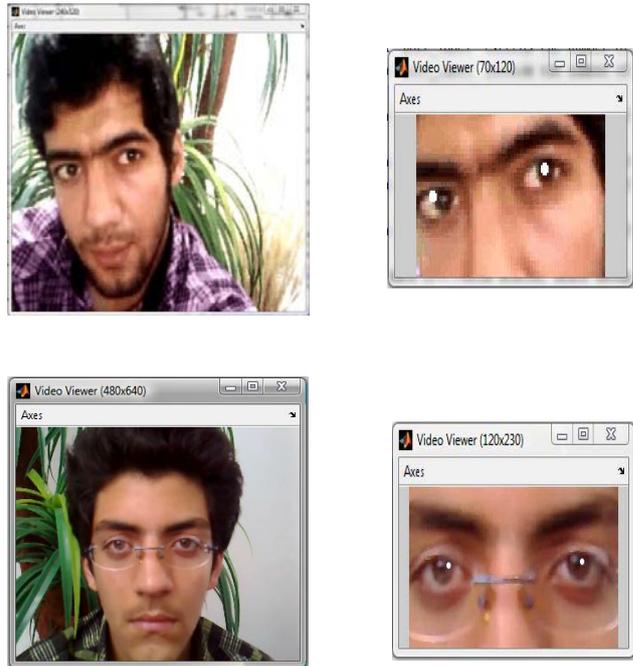
$$x_{t+1} = x_{t+1}^{-1} + k_{t+1} (z_{t+1} - Hx_{t+1}^{-1}) \quad (12)$$

In this equation x_{t+1}^{-1} is the estimated state vector, and k_{t+1} represents the gain matrix, which is derived from equation (13)

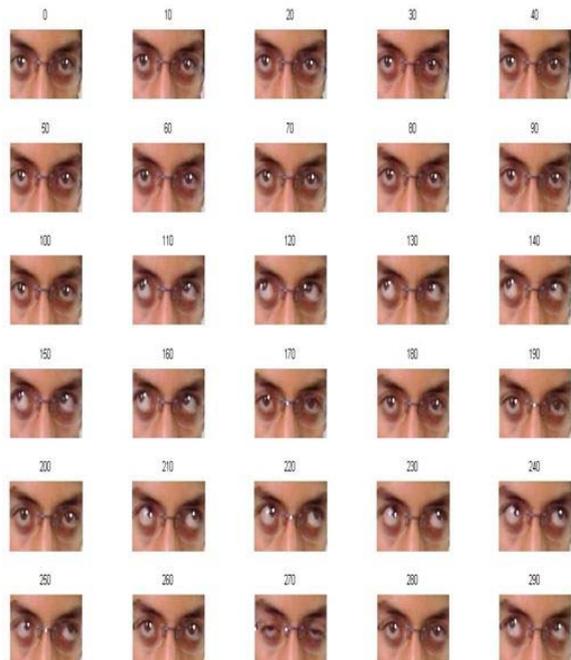
$$k_{t+1} = \frac{\sum_{t+1}^{-1} H^T}{H \sum_{t+1}^{-1} H^T} \quad (13)$$

IV. OBSERVATION

In this section the experimental results obtained on a large number of images are shown. The images taken from Banka database. No constraints have been imposed on the background. Fig.6 shows the result of eye tracking by using kallman filter and pupil detection in different frames. In fig.7 and fig. 8 some results of correct detection are shown, in particular fig.8 shown the eye detection on image of people wearing glasses.



Pupil detection



result of the detection of a person wearing glasses in different frames.



some result with people wearing glasses



some result

The proposed method is a step-by-step process and the results of a subsequent step necessarily depend on the previous step. Accordingly, if the first stage is not precise, the final detection will not be accurate enough, and the offered procedure will failed. The result statics of eye detection are tabulated in table 1. Based on the derived results, it could be claimed that the method enjoys a high detection rate.

Table 1 Rate of Correct and Incorrect Eye Detection

	Two eyes are correctly detected	one eyes are wrongly detected	Two eyes are wrongly detected
Person 1	92.4%	3.4%	4.2%
Person 2	89.6%	5.4%	5.0%
Person 3	90.8%	3.1%	6.1%
Person 4	94.7%	3.7%	1.6%
Person 5	91.1%	2.3%	6.6%
Person 6	88.6%	4.3%	7.1%
mean	91.2%	3.7%	5.1%

V. CONCLUSION

The purpose of this paper was to offer a novel method to detect and track human eye movements. To locate the pupil, first the face was extracted from the whole image. this neutralized the impact of image background in the later stages of the process. Then the pupil position was found using the proposed method. The position of the eye region in the next frame was estimated by using kallman filter. An area surrounding the estimated location was selected and the proposed method was carried out within it. This narrowed the operational domain and required fewer calculations. Based on the obtained practical results, the proposed method could be assessed as an appropriate method to eye detect and track. Results indicated that the method could locate the position of the pupil well.

REFERENCES

- [1] D. Hansen Q.Ji," In the Eye of the Beholder: A Survey of Models for Eyes and Gaze", Pattern Analysis and Machine Intelligence, IEEE Transactions on March 2010
- [2] E. Rosten and T. Drummond. "Fusing Points and Lines for High Performance Tracking"; Proceedings of the IEEE International Conference on Computer Vision Vol. 2 (October 2005): pp. 1508–1511.
- [3] Ba Linh NGUYEN, "Eye Gaze Tracking"; IEEE vol. 2,2009.
- [4] P. Gejgus, J. Placek, M. Sperka," Skin color segmentation method based on mixture of Gaussians and its application in Learning System for Finger Alphabet",International Conference on Computer Systems and Technologies - CompSysTech'2004
- [5] Z Zhu, Q. Ji, K. Fujimura," Combining Kalman Filtering and Mean Shift for Real Time Eye Tracking Under Active IR Illumination", IEEE 2002
- [6] J. Zhang, Y. Liu and S. Ha," A Novel Approach of Face Detection Based on Skin Color Segmentation and PCA",The 9th International Conference for Young Computer Scientists, IEEE 2008