Real Time Eye Tracking in Unconstrained Environments

Majid Tajeri[1], Mahmoud Mahlouji[2] and Hossein Ebrahimpour-Komleh[3]

**[1] Department of Computer Engineering, Arak Branch, Islamic Azad University
Arak, Iran**

**[2] Department of Electrical and Computer Engineering, Kashan Branch, Islamic Azad University
Kashan, Iran**

**[3] Department of Computer Engineering, The University of Kashan
Kashan, Iran**

Abstract

In this paper, an effective method for human eye tracking and also decreasing the current challenges and problems in its algorithms, possibly as real time and for unconstrained environments has been proposed. In this method, firstly face has been detected and segmented from the remaining parts to make the searching area in tracking stage, narrower and processing speed higher. Then eye area is determined and eye pupils are detected in the specified area. In the proposed method, to support tracking in eye occlusion state, corner detection has been additionally used. Experimental results show the potential of this method for real time eye tracking in unconstrained environments with existence of complex background, head and face rotation, beard, makeup, eye glasses and veil, even while the eyes are closed. The correct recognition rate of the proposed method is about 91.9%.

Keywords: *Eye Detection, Face Detection, Eye Tracking, Kalman Filter, Corner Detection, Harris Algorithm.*

1. Introduction

Nowadays by ever-increasing development of scientific discussions in the content of computer vision and pattern recognition, automatic eye tracking in video images has been significantly considered because of eye organism and structure. Since eye is an important and critical organ of the body and manifests some internal moods and their reflexes to physical and non-physical factors, by using it some moods like fatigue, sleepiness, drunkenness, happiness, and sadness, stare, etc can be distinguished. So, design and implementation of eye tracking systems are applicable to the following cases: safety systems to control peoples entrance and exit, tracking, detecting and determining social misdemeanant identity, traffic systems to recognize drunken and narcotic-taken drivers, lie-detector systems and in cars to detect driver fatigue and sleepiness. An eye tracking system includes the following stages: image acquisition; face detection; eye detection; and eye tracking.

Eye detection methods are divided into four classes [1]: shape-based methods; feature-based methods; appearance-based methods and hybrid methods. When eye is open, it has some special form. Eye's corners, pupil and eyelids cause it to be distinguished from other parts of face. Shape-based methods use these characteristics and express some models for eye. This model can be a simple ellipsoid or a complex structure [2, 3]. Feature-based methods use some distinguishing features around the eye. Iris, pupil, eyelids, eyelashes, limbus (border between iris and sclera), pupil expression and reflection of light in cornea are some features used for detecting the eye [4, 5]. Appearance-based methods which are known as pattern matching methods or general methods detect and track the eyes directly based on their appearance states. These methods are independent of the considered objects and can model every other object beside the eye [6, 7]. The main purpose of hybrid methods is to compose various advantages of eye models at a system to overcome relative constraints. To improve these algorithms, the proposed algorithms try to reduce search area for pattern matching by combining methods, or propose an approach reducing the time of this pattern matching [8,9].

The available approaches for eye tracking can be divided into 3 classes [10]: knowledge-based methods, learning-based methods and motion estimation methods. In knowledge-based methods, tracking methods are defined and developed based on the rules obtained by studies and researches on face components. Simply we can utilize some rules to describe face features and their relations [11, 12]. Learning-based methods can be classified into the three following categories [13, 14]: neural networks, Adaboost classifiers and support vector machines. Motion estimation methods estimate the object position in current frame, and then determine its exact position by a local searching. In the next stages, by a correct estimation in the first stage, the accuracy and effectiveness of the method are increased [15-17].

In this paper, an effective method is presented for real time eye tracking in unconstrained environments. To perform this, in eye and face detection stages, color feature has been used. Then in tracking stage, pupil position estimation method and Kalman filter have been used. To support eye occlusion in tracking stage, shape-based method is combined with pupil position estimation method and parallel to Kalman filter, corner detection has been used. The remainder of this paper is as follows: in the second section, an eye tracking system will be introduced. In the third section, the proposed method is presented. In the fourth part, the results obtained by experiments are offered and the fifth section is for conclusion.

2. Eye Tracking System

In general, an eye tracking system can be outlined in four stages:

- Image acquisition
- Face detection
- Eye detection
- Eye tracking

Image acquisition stage is data or image entry to system which can be a video taking by a camera. The output of this stage is an image in RGB space. After image acquisition, face detection is performed. One current method for face detection is by using color feature [18]. As the RGB space, besides having pixel color data, has light intensity data, and also, face brightness is different in various people and environments, use this space to detect face color is not effective and makes some problems. So we should find a space not dependent on light intensity. One option is YCbCr chromatic color space. Firstly the image is transformed to YCbCr. RGB color space is transformed to YCbCr as:

$$\begin{cases} Y = 0.299R - 0.587G - 0.11B \\ Cb = R - Y \\ Cr = B - Y \end{cases} \qquad (1)$$

After transforming to YCbCr, the mean and covariance matrix of Cb and Cr components are computed. Eq. (2) shows how to compute the mean and covariance matrix:

$$mean: \mu = E(x)$$
$$\text{cov}: c = E\left((x-\mu)(x-\mu)^T\right) \qquad (2)$$

in which E is mathematical expectation, μ is mean, c is covariance matrix, x is the main matrix and $(x-\mu)^T$ is $(x-\mu)$ transposed. By using mean and covariance measures, a Gaussian model for skin color can be determined. After obtaining skin color model, you can use Gaussian function and Euclidean distance for face detection and recognize a pixel belong or non-belong to the considered model. Suppose two vectors $A = \{a_1, a_2, \ldots, a_n\}$ and $B = \{b_1, b_2, \ldots, b_n\}$. Then their Euclidean distance is as:

$$D(A,B) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2} \qquad (3)$$

After face detection, eye detection is performed. Since eye color differs from other parts of face, this feature can be used for eye detection. For eye detection based on color feature, we can use horizontal and vertical projection [19]. Horizontal projection for an image with M rows and N columns is obtained from adding pixels intensities in each row according to Eq. (4), and vertical projection is obtained from adding pixels intensities in each column according to Eq. (5) as:

$$hp = \{hp_x \mid 1 \le x \le M\}$$
$$hp_x = \sum_{y=1}^{N} f(x,y) \qquad (4)$$

$$vp = \{vp_y \mid 1 \le y \le N\}$$
$$vp_y = \sum_{x=1}^{M} f(x,y) \qquad (5)$$

Another method of eye detection is by using eye corners. In [20], Harris algorithm is used to specify the pupil's center. In this algorithm, a circle mask in each point of image is considered. By displacing mask in different directions the mean variation rate of image intensity in each window, compared with the main window, is computed and the minimum variation is considered as corner response. Depending on the position of each point, three states are created:

- If the windowed image patch is flat (approximately constant in intensity), then all shifts will result in only a small change.
- If the window straddles an edge, then a shift along the edge will result in a small change, but a shift perpendicular to the edge will result in a large change.
- If the windowed patch is a corner or isolated point, then all shifts will result in a large change. A corner can thus be detected by finding where the minimum change produced by any of the shifts is large.

Harris algorithm is mathematically expressed as:

$$E_{x,y} = \sum_{u,v} w_{u,v} \mid I_{x+u,y+v} - I_{u,v} \mid^2 \qquad (6)$$

in which w is a mask imposed on image, which is considered as circle mask with coefficients 1. I is the image gradient and E is the variation made by (x,y) displacement. (x,y) displacement in four directions includes the following set:

$\{(1,0),(1,1),(0,1),(0,0)\}$

Harris corner detector introduces local maximum of $\min_{x,y}\{E_{x,y}\}$ that is more than a special threshold, as a corner. Eq. (7), includes all possible little displacement as:

$$E(x,y) = Ax^2 + 2Cxy + By^2$$
$$A = (I_x)^2 \otimes w$$
$$B = (I_y)^2 \otimes w$$
$$C = (I_x I_y) \otimes w \qquad (7)$$

in which, w is Gaussian mask, I_x, I_y is image gradient in x and y directions respectively and $\otimes$ is convolution operator. E can be expressed as:

$$E(x,y) = (x,y)M(x,y)^T$$
$$M = \begin{bmatrix} A & C \\ C & B \end{bmatrix} \qquad (8)$$

in which M is a 2*2 matrix in (x, y) coordinates and E is local auto-correlation function.

In [20], it has been shown that corner response, R, can be obtained in each image point as:

$$R = Det(M) - KTr^2(M)$$
$$Tr = \alpha - \beta = A + B$$
$$Det(M) = \alpha\beta = AB - C^2 \qquad (9)$$

in which K is constant. The corner response in corner area is positive, in edge area is negative and in flat area, has a minimal absolute magnitude. To reduce the computation time, the existing mask in Harris algorithm is neglected. So, firstly A, B, C are obtained as:

$$A = (I_x)^2$$
$$B = (I_y)^2$$
$$C = (I_x I_y) \qquad (10)$$

After computing these measures, R is obtained by Eq. (9). Experiments show two points with the highest value of R in an eye area are eye iris corners [20].

The last stage in eye tracking system is eye tracking stage. One most favored and applicable method useful for real time applications is Kalman filter [14, 21]. A Kalman filter is a set of recursive algorithms estimating the position and uncertainty of moving objects in the next time frame. It means that where the pupil should be searched and how wide the image in the next frame, around the estimated positions, should be searched to find pupils confidentially. This method recursively limits the present estimation to all previous computations, and this trend is repeated with previous estimations which are used for prediction. This recursive nature is one attractive characteristic of Kalman filter, which makes its implementation simple and possible. In Kalman filter, a state vector is considered. This vector expresses the system state. Kalman filter, in each stage, by using movement data in previous time, predicates the next time state vector. State vector is $X = (x, y, v_x, v_y)$, in which x, y are components of moving coordinates and v_x and v_y are speed components in x and y directions, respectively. State equation is as:

$$X_{t+1} = \Phi_t X_t + w_t \qquad (11)$$

in which w_t is Gaussian white noise with zero mean, X_{t+1} and X_t are state vectors in $t+1$ and t times, respectively, showing movement speed and Φ_t matrix is considered as:

$$\Phi_t = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (12)$$

in which T is time distance between frames. Measurement vector is considered as z=(x, y), and relation between state vector and measurement vector is as:

$$Z_t = HX_t + e_t \qquad (13)$$

in which e_t is measurement error, X_t is state vector in t time, Z_t is measurement vector in t time, H is a matrix which relates state vector to observation vector, and is defined as:

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \qquad (14)$$

By using Kalman filter and from movement value in previous times, the new position of head in the next frame is predicated. In [22], Kalman filter is used for eye tracking in video streams. The pupil state in each time is expressed by speed and location coordinates. If the coordinates of pupil center shown with x and y, and speed in x and y directions shown by u and v, respectively, the state vector in t time is shown with $X_t = (x_t, y_t, u_t, v_t)$.

In Kalman filter, X_{t+1}^{-1} is the next time state vector predicated as:

$$X_{t+1}^{-1} = \Phi_t X_t \qquad (15)$$

By replacing Eq. (13) with Eq. (15), observation matrix is obtained. Error covariance matrix showing the covariance between the predicated and actual states is obtained as:

$$\sum_{t+1}^{-1} = \Phi_t \sum_t (\Phi_t)^T + e_t \qquad (16)$$

in which $(\Phi_t)^T$ is Φ_t transposed and $\sum_t$ is error covariance matrix in *t* time. Based on this matrix, firstly, an area is defined and eye pupil is searched in this area. Then, the actual state vector is obtained as:

$$X_{t+1} = X_{t+1}^{-1} + K_{t+1}\left(Z_{t+1} - HX_{t+1}^{-1}\right) \quad (17)$$

in which K_{t+1} is a matrix obtained as:

$$K_{t+1} = \frac{\Sigma_{t+1}^{-1} H^T}{H\,\Sigma_{t+1}^{-1} H^T} \quad (18)$$

and finally, parameters predicted for the next stages are modified.

3. The Proposed Method

At the first, to make the intensity around ineffective in the proposed algorithm, the color space of input image is transformed from RGB to YCbCr by Eq. (1). Then a color Gaussian model, according to Eq. (2) is defined for face, and finally, the face area is detected. The input image used here, has a complex background, no constraint is considered for it, head can rotate in different orientations, eye can be occluded and person can have eyeglasses, beard, veil or make up. In the next stage, by using horizontal projection according to Eq. (4), the region containing eyes is determined and divided into right and left areas, to precisely find the pupils' center and in each segmented area; the eye pupil's center is searched. This operation causes the searching area to be narrower and the algorithm becomes closer to real time state. After that, by using vertical and horizontal projection via Eq. (4) and Eq. (5), the pupils' center is obtained. In the last stage, two Kalman filters track the eyes in parallel. Also, corners detections are used for eye tracking.
The proposed method includes 4 stages: face detection, detection of area including eyes, obtaining the eye's pupils and corner detection, and eye tracking.

3.1 Face Detection

Since color is fixed when the head rotates and its processing is faster than other features of face, this feature is used for face detection. The most current color space is RGB which has many applications and is the standard space to present color images. But because of what is explained in section 2, the image is transformed to YCbCr space by Eq. (1). Fig. 1 shows the differences between a person's images in color space RGB and YCbCr, Cb component and Cr component.

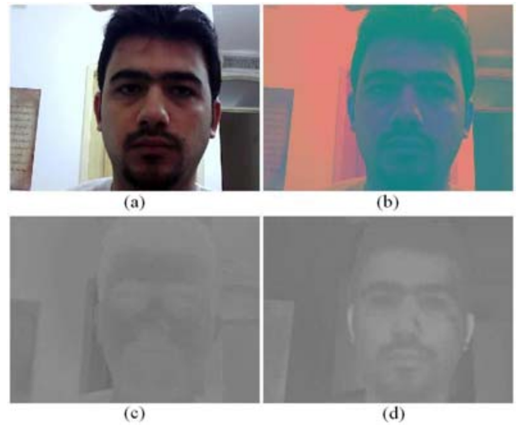

Fig. 1 Differences between a person's images in various color spaces: (a) RGB space (b) YCbCr space (c) Cb component (d) Cr component

3.1.1 Obtaining Skin Color Model

To create skin color model, firstly a set of various skin images with different color and texture is prepared, then is transformed to YCbCr color space. After that by using Eq. (2), skin color model is obtained. In Fig. 2 there are some samples of skin color.

3.1.2 Euclidean Distance

After obtaining skin color model, by using a Gaussian model and Euclidean distance computed by Eq. (3), every pixel's belong or non-belong to the model is determined. Fig. 3 shows the results obtained by face detection of different people.

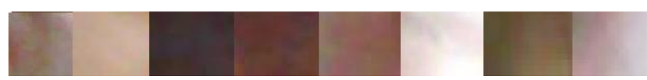

Fig. 2 Some samples of skin color

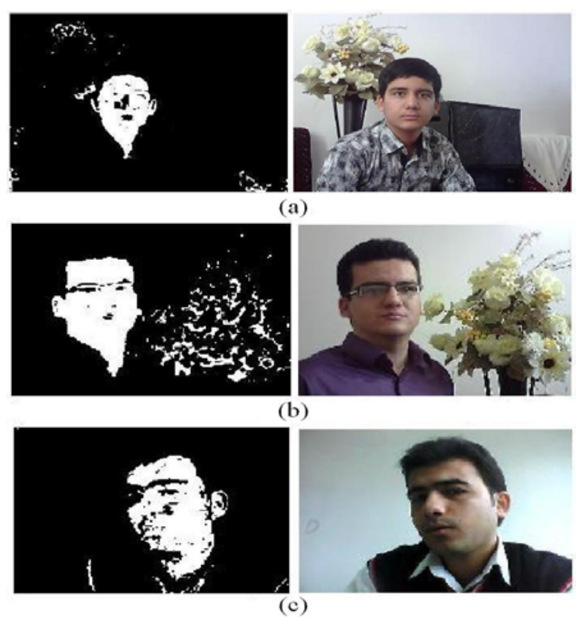

Fig. 3 Results obtained by face detection of different people with various brightness

3.2 Eye Area Extraction

Eye area starts from the above of eyebrows to under the lower eyelid including two eyes, two eyebrows and their distinguishing features as eyelids and eye corners. To extract this area horizontal projection of face that computed by (4) is used. In Fig. 4 results of the horizontal projection is shown. As could be seen in Fig. 4, from above, the first valley (No. 2) belongs to eyebrows; the second valley (No. 4) belongs to eyes pupils. As mentioned in section (3), since the image has no limitation and face may rotate, it cannot be claimed that the second valley always belongs to pupils. As seen in the face image forehead and under the eyes area (the distance between nose ridges under the eyes) have more brightness comparing to the eyes. This causes the horizontal projection to have two peaks. The first peak (No. 1), belongs to forehead, the second peak (No. 5) belongs to the area under eyes. So, maximum points (peaks) are used to extract eye area. As the second peak is above the nose ridges area, if face has beard, the algorithm will be properly operated. Fig. 5 shows the result of eye area detection by using horizontal projection.

3.3 Finding Eye Candidates

To find eye candidates, for image in Fig. 5, horizontal and vertical projection are obtained according to Eq. (4) and Eq. (5) and the results are shown in Fig. 6. In Fig. 6(b), there are three peaks (Nos. 6, 8, 10) and two valleys (Nos. 7, 9). In this figure, the peak (No. 8) is related to the area between two eyebrows. This point and points 6 and 10 are used for dividing the area including eyes into right and left areas, so that the area between points 6 and 8 is considered as left eye area and the area between points 8 and 10 is considered as right eye area. In Fig. 6(a), there are two valleys. From above, the first valley (No. 2) belongs to eyebrow and second valley (No. 4) belongs to pupil. This is the case for two left and right areas. Also in Fig. 6(b), the valley (No. 7) belongs to left eye's eyebrow and pupil and the valley (No. 9) belongs to right eye's eyebrow and pupil. So the mentioned points are used and eyes candidates are obtained.

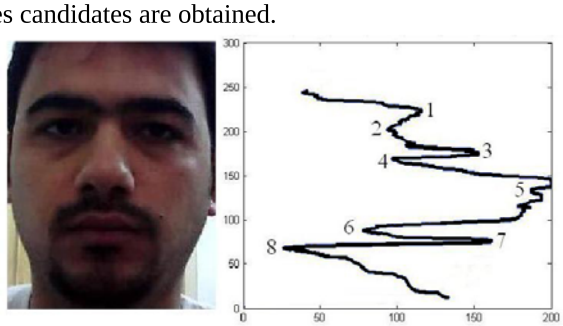

Fig. 4 Horizontal projection of face: 1) Forehead 2) Eyebrows 3) Distance between eyebrow and eye 4) Pupils 5) Distance between the ridges of nose under the eyes 6) Moustache 7) Lips 8) Area under lips

Fig. 5 Eye area detection by using horizontal projection

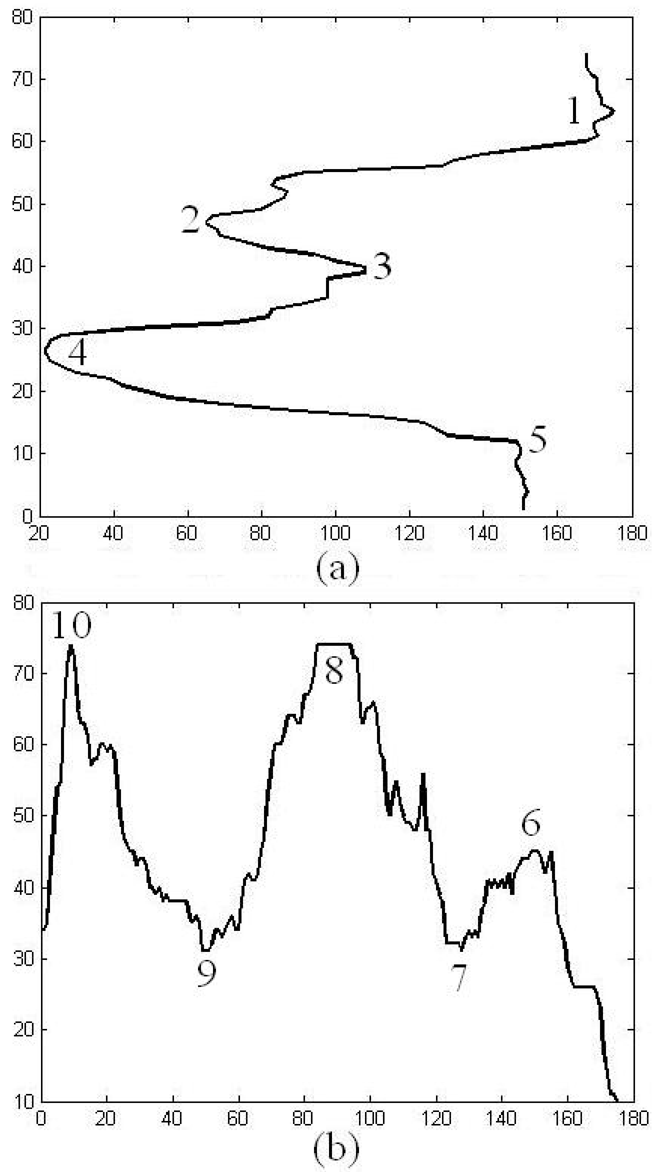

Fig. 6 Horizontal and vertical projection of Fig. 5:
(a) Horizontal projection: 1-Forehead 2-Eyebrows 3- Distance between eyebrow and eye 4-Pupils 5-Area under eyes (b) Vertical projection: 6-Beside the left eye 7-Left pupil and eyebrow 8-Distance between two eyebrows 9-Right pupil and eyebrow 10-Beside the right eye

Regarding to the mentioned subject and vertical and horizontal projections of eyes area shown in Fig. 6, there are two candidates for eyes. The first candidate is eyebrows and the second one is eyes. The reason why eyebrows are considered as eye candidate is that eyebrows, like pupil, has some darkness and makes valley. Also eyebrows are located within the area containing the eyes. To omit eyebrows as candidates, horizontal projection of face area shown in Fig. 6.a is used. In this figure, there are two valleys. Regarding face geometry and eyebrows locating above the eye, we can conclude that from above, the first valley (No. 2) belongs to eyebrows. So, to recognize the eye pupil, the first valley is neglected and the second one (No. 4) surely belonging to the eyes, is examined.

3.4 Finding the Pupil Center

By having a pixel from iris or pupil and having pupil circle, pupil center can be obtained. To find pupil center, lines intersection can be used. In this method, only one row and column of eye is processed instead of total eye area. So, a line is drown, along with *x* and *y* axis to the border edge of iris and sclera. As shown in Fig. 7, if the intersection center of drown lines is considered as (x,y) and the end points of lines are considered as $(x_1,y'),(x_2,y'),(x',y_1),(x',y_2)$, the pupil center can be obtained as:

$$x_c = \frac{x_1+x_2}{2}$$
$$y_c = \frac{y_1+y_2}{2} \quad (19)$$

In Fig. 8, pupil recognition is examined in different states. In this figure, some states in which head has rotating movement or person may have eyeglasses, make up, beard or veil, have been shown.

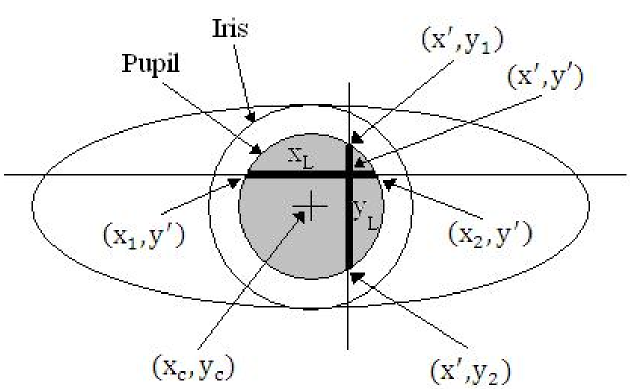

Fig. 7 Using lines intersection to find pupil center

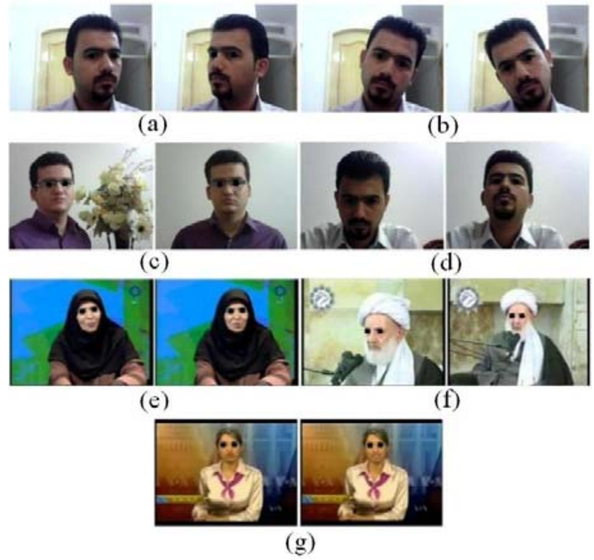

Fig. 8 Pupil detection: (a) Horizontal head rotation (b) Vertical head rotation (c) With eyeglasses (d) Up and down head rotation (e) With veil (f) With beard (g) With make up

3.5 Eye Tracking

In most applications, after eye detection, tracking of eye or face movements is considered. The same problems in eye detection state are in eye tracking, by a difference that in the tracking cases, the problems appears stronger and there is more destructive effect on tracking process. As an example, changing the environment light conditions cause loss of eye position on the image and consequently eye detection process is repeated. Light reflection on eyeglasses can cause the same problem.

For eye tracking, in addition to the estimation and location of eye positions in sequential frames, head and face movements should be considered and a method not sensitive to these changes and problems should be proposed. In addition to head, face components like eyelids have independent movement which should be considered.

To support head movements, the eye itself was not tracked and an area containing eye was tracked. Among the presented methods, Kalman filter has less computational complexities and higher speed, and is suitable for real time applications. So, Kalman filter is used for tracking. In the proposed method, two filters used in parallel, track the eyes separately because, as explained in section 3-3, the eyes are recognized separately. Kalman filter in each stage, by using the movement data in previous times, predicts the state vector of next time according to Eqs. (11) to (18). Kalman filter estimates form state vectors of t and t+1 time. Fig. 9 shows tracking result of eye area by using Kalman filter and pupil recognition in different frames. In these frames, the person's background is varied and it has head movements. As seen, the proposed method also supports head and face movements.

Fig. 9 Result of eye area tracking by using Kalman filter and pupil recognition in different frames

A problem here is that if both eyes are occluded or closed, Kalman filter does not work. To overcome this problem, besides Kalman filter, the eye corners features are also used. To perform this, in each left and right area, two corners are recognized. To find the corners, Harris algorithm, explained in section 2 has been used. So, the proposed method, is a combination of Kalman filter and corner detection. In the proposed method, after estimating eye position by Kalman filter, the estimated position is compared with the previous position, and if their coordinates are higher than an experimental threshold level, tracking will be wrong. In this state, the position determined by corners is used.
Although in this method, the required hardware has been increased, and for each eye an additional filter has been used, but tracking in the case of face and head rotating movements and occlusion and loss of one of the eyes, is performed more precisely. Also, as the filters operate in parallel, no significant change in working speed happens and tracking is about real time.

4. Experimental Results

Fig. 10 shows the result of eye tracking and pupil recognition by using different frames. In this figure, the eyes are closed in the first three rows and as observed in figures, eye closeness problem is resolved and the eye in closing state is also tracked.
The proposed method supports various states of head and face movements, eye occlusion and closeness and tracking persons having eyeglasses, veil and make up, shown in Fig. 11 to Fig. 13, respectively.

Fig. 10 Result of eye tracking in different frames with proposed method

Fig. 11 Result of tracking with eyeglasses

Fig. 12 Result of tracking with make up

Fig. 13 Result of tracking with veil

Statistical results of eye recognition rate have been expressed in Table 1. In this table, for each set, three films with different conditions are considered and the results are the mean correct or incorrect recognition for these sets. Regarding the obtained results, it can be claimed that this method mostly supports unconstrained environments with existence of makeup, veil, beard, eyeglasses and rotating movement's states.

Following results are concluded from Table 1:

- As a horizontal projection of face and maximum measure are used to find eye areas, and this measure is for under the eye area, between lower eyelid and nose ridges, if person has beard the results are good and acceptable.

Table 1: Correct and incorrect eye recognition rate

	Correct recognition of two eyes	Incorrect recognition of one eye	Incorrect recognition of two eyes	Total frames
First set (without beard)	94.7%	2.02%	3.28%	1980
Second set (with beard)	92.86%	3.4%	3.74%	1470
Third set (with eyeglasses)	87.1%	5.35%	7.55%	1590
Fourth set (with makeup)	92.0%	3.33%	4.67%	1500
Fifth set (with veil)	92.12%	3.38%	4.5%	2220
Average	91.9%	3.42%	4.68%	Total = 8760

- As the results made by eyeglasses, comparing to other cases, are less when person is wearing eyeglasses and his eyes is occluded or closed, since the eyeglasses cover the eye corners and corners are loosed, tracking encounters problem.
- Since to recognize pupil, color feature is used and make up does not change the pupil color, with makeup, acceptable results has been obtained. But as make up changes the eye corners, in this state the results obtained is less than non-beard state.
- The reason why the results obtained by veil is less than those of without beard is that if person has veil and its head rotates and its eye occluded, since veil covers eye corners, tracking encounters problem and recognition rate decreases.

Comparison between the proposed method and other ones is difficult because of their performing in different platforms and databases. So, the main parts of these methods are compared theoretically. In Table 2, the proposed method is used with 3 similar methods and supporting various conditions including light conditions, head rotation, eyeglasses, beard, make up, veil, eye occlusion and images resolution, had been compared.

Table 2: Comparison between the proposed method with 3 different methods

	Used Method for detection and tracking	Detection in different light conditions, head rotation, eyeglasses, beard, make up, veil and eye occlusion	Image with high resolution
Paper [5]	Entropy of eye and darkness of iris	Supports the different light conditions, head rotation and eyeglasses	Need
Paper [8]	Combine eye form and image intensity	Supports eyeglasses and beard	No Need
Paper [12]	Eye form structure and corners	Supports head rotation	Need
Proposed Method	Combine color intensity around the eye, Kalman filter and corner detection	Supports head rotation to 40 degree, different light conditions, eyeglasses, beard, make up, veil and eye occlusion	No Need

Following results are concluded from Table 2:

- In [5] entropy of eye and darkness of iris is used. This method supports various light conditions, head rotation and existence of eyeglasses and in case of existing make up, beard and eyes occlusion, tracking is not possible. Also the images used by this method should be in high resolution.
- In [8] a combination of eye form and image intensity is used for tracking and detecting. This method supports eyeglasses and beard existence and in case of head rotation, makeup, veil, eyes occlusion, it cannot work.
- In [12] the eye form, structure and corners are used for detection. This method only supports head rotation and in case of different light conditions, eyeglasses, beard, make up, veil and eye occlusion, it cannot track. Also, the images used by this method should have high resolution.
- In the proposed method the combined color intensity around the eye, Kalman filter and corner detection are used for detection and tracking. The proposed method supports various light conditions, head rotation, eyeglasses, beard, and make up, veil and eye occlusion. Also in this method there is no need to the high resolution images.

5. Conclusions

In this paper, a new method for human eye tracking in unconstrained environments has been presented. Considering the experiments in different conditions including images with complex background, changing light conditions, head rotation, beard existence, veil, eyeglasses, make up and eye closeness, good results were obtained. Also, experimental results show the proposed method has low computational complexity and maximum stability. Also, it is convenience for unconstrained environments and real time applications.

References

[1] S. Yamamoto, V. G. Moshnyaga, "Algorithm Optimizations for Low-Complexity Eye Tracking", in IEEE International Conference on Systems, Man, and Cybernetics, 2009.
[2] K. N. Kim, R. S. Ramakrishna, "Vision-Based Eye-Gaze Tracking for Human Computer Interface", in IEEE International Conference on Systems, Man, and Cybernetics, 1999.
[3] R. Kothari, J. L. Mitchell, "Detection of Eye Locations in Unconstrained Visual Images", in IEEE International Conference on Image Processing, 1996.
[4] W. Liting, D. Xiaoqing, F. Chi, "Face Live Detection Method Based on Physiological Motion Analysis", Tsinghua Science and Technology, Vol. 14, No. 6, 2009.
[5] M. Hassaballah, S. Ido, "Eye Detection Using Intensity and Appearance Information", in IAPR Conference on Machine Vision Applications, 2009.
[6] Z. Zhu, K. Fujimura, Q. Ji, "Real-Time Eye Detection and Tracking under Various Light Conditions", in Eye Tracking Research & Applications Symposium, 2002.
[7] A. Hyrskykari, "Utilizing Eye Movements: Overcoming Inaccuracy while Tracking the Focus of Attention during Reading", Computers in Human Behavior in Attention aware systems, Vol. 22, No. 4, 2006.
[8] K. Peng, L. Chen, S. Ruan, G. Kukharev, "A Robust Algorithm for Eye Detection on Gray Intensity Face without Spectacles", JCS&T, Vol. 5 No. 3, 2005.
[9] K. C. K. Lee, J. Schiffman, B. Zheng, W. C. Lee, H. V. Leong, "A System for Tracking Nearest Surrenders in Moving Object Environments", in IEEE International Conference on Pervasive Services, 2006.
[10] Z. Qian, D. Xu, "Automatic Eye Detection Using Intensity Filtering and K-means Clustering", Pattern Recognition Letters, Vol. 31, No. 12, 2010.
[11] B. Kroon, S. Maas, S. Boughorbel, A. Hanjalic, "Eye Localization in Low and Standard Definition Content with Application to Face Matching", Computer Vision and Image Understanding, Vol. 113, No. 8, 2009.
[12] Y. Tian, T. Kanade, J. F. Cohn, "Dual State Parametric Eye Tracking", in IEEE International Conference on Automatic Face and Gesture Recognition, 2000.
[13] F. H. C. Tivive, A. Bouzerdoum, "A Fast Neural-Based Eye Detection System", in IEEE International Symposium on Intelligent Signal Processing and Communication Systems, 2005.
[14] Z. Niu, S. Shan, S. Yan, X. Chen, W. Gao, "2D Cascaded Adaboost for Eye Localization", in IEEE International Conference on Pattern Recognition, 2006.
[15] D. Comaniciu, V. Ramesh, P. Meer, "Kernel-Based Object Tracking", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 5, 2003.
[16] Z. Zhu, Q. Ji, K. Fujimura, "Combining Kalman Filtering and Mean Shift for Real Time Eye Tracking under Active IR Illumination", in IEEE International Conference on Pattern Recognition, 2002.
[17] G. Veneri, P. Federighi, F. Rosini, A. Federico, A. Rufa, "Influences of Data Filtering on Human–Computer Interaction by Gaze-Contingent Display and Eye-Tracking Applications", Computers in Human Behavior, Vol. 26, No. 6, 2010.
[18] X. Long, O. K. Tonguz, A. Kiderman, "A High Speed Eye Tracking System with Robust Pupil Center Estimation Algorithm", in IEEE International Conference Engineering in Medicine and Biology Society, 2007.
[19] Z. Zhu, Q. Ji, "Eye and Gaze Tracking for Interactive Graphic Display", Machine Vision and Applications, Vol. 15, No. 3, 2004.
[20] T. Liu, S. Zhu, "Eyes Detection and Tracking Based on Entropy in Particle Filter", in International Conference on Control and Automation, Vol. 2, 2005.
[21] W. Zheng, Z. Lu, X. Xu, "A Novel Skin Clustering Method for Face Detection", in First International Conference on Innovative Computing, Information and Control, Vol. 3, 2006.
[22] Q. Ji, X. Yang, "Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance", Elsevier Science, Vol. 8, No. 5, 2002.

Majid Tajeri received the B.S. degree in computer engineering from Kashan Branch, Islamic Azad University, Kashan, Iran, in 2009. He is currently a graduate student in the Department of Computer Engineering at Arak Branch, Islamic Azad University, Arak, Iran. His research interests include image processing and pattern recognition.

Mahmoud Mahlouji received the B.S. degree in telecommunications engineering from Sharif University of Technology, Tehran, Iran, in 1990, the M.Sc. degree in electronics engineering from Sharif University of Technology, Tehran, Iran, in 1993, and the Ph.D. degree in telecommunications engineering from Science and Research Branch, Islamic Azad University, Tehran, Iran, in 2008. At present he is an assistant professor of the Electrical and Computer Engineering Department, Kashan Branch, Islamic Azad University, Kashan, Iran. His current interests are in image processing, pattern recognition, neural networks, computer vision, and iris recognition.

Hossein Ebrahimpour-Komleh is currently an assistant professor at the Department of Electrical and Computer Engineering at the University of Kashan, Kashan, Iran. His main area of research includes Computer Vision, Image Processing, Pattern Recognition, Biometrics, Robotics, Fractals, chaos theory and applications of Artificial Intelligence in Engineering. He received his Ph.D. degree in Computer engineering from Queensland University of technology, Brisbane, Australia in 2006. His Ph.D. research work is on the Fractal Techniques for face recognition. From 2005 to 2007 and prior to joining the University of Kashan, he was working as a Post-doc researcher in the University of Newcastle, NSW, Australia and as a visiting scientist in CSRIO Sydney. Hossein Ebrahimpour-Komleh has B.Sc. and M.Sc. degrees both in computer engineering from Isfahan University of Technology (Isfahan, Iran) and Amirkabir University of Technology (Tehran, Iran,) respectively. He has served as the editorial board member and reviewer of several journals and international and national conferences.