Proc. 1st International Conference on
Machine Learning and Data Engineering (iCMLDE2017)
20-22 Nov 2017, Sydney, Australia
ISBN: 978-0-6480147-3-7

Real-Time Iris Center Detection and Tracking From Video Images

Dipankar Das[1], Md. Golam Rashed[1], Md. Rezaul Bashar[2], and Yoshinori Kuno[3]

[1]University of Rajshahi, Rajshahi, Bangladesh
[2]Senior Research Fellow, Science, Technology and Management Crest, NSW 2566, Australia
[3]Saitama University, Saitama-shi, Saitama, Japan

Corresponding author's E-mail: dipankar@ru.ac.bd or dipankar_ru@yahoo.com

Abstract

In this paper, an efficient real-time iris center detection and tracking system has been proposed based on the vector field of image gradients (VFOID). The system uses video frames from cost-effective low-resolution USB cameras and natural lighting condition instead of high resolution sensors and infrared illumination. The proposed algorithm includes multi-stage approaches. Firstly, the AdaBoost-based cascaded classifiers are used to detect and track face region and its location in the image. Then, the facial feature points in the face region are extracted to locate the rough area of the eyes. Finally, the VFOID are used to accurately detect and track the iris center within the estimated eye areas. The experimental result shows a high accuracy on existing database and in a real-world environment.

Keywords: *Iris Center Detection, Iris Tracking, Image Gradients, Face Detection and Tracking*.

1. INTRODUCTION AND MOTIVATION

Analysis and understanding of human behavior, particularly of head movement and eye gaze behavior, has been a subject of interest for many years in cognitive psychology and neurophysiology (J. S. Agustin, J. P. Hansen & J. Mateo, 2008); however, a full understanding of the causes, dynamics, and control mechanisms of head and eye movements is still a subject of active research (E. G. Freedman, 2008). Recently, researchers in the computer vision community have increasingly sought to incorporate information gleaned from patterns of human behaviors into intelligent human-machine interfaces (A. Doshi & M. M. Trivedi, 2012) (R. Valenti, N. Sebe & T. Gevers, 2012). Iris center estimation is important in many application areas spanning from human-computer interaction (HCI) to human behavior analysis. The visual gaze, determined by detecting and tracking the center of the iris, provides important information about the interest of the subject. It is commonly used in various application areas such as: controlling devices for disabled people (J. S. Agustin, J. P. Hansen & J. Mateo, 2008), to analyze the user attention while driving (A. Doshi & M. M. Trivedi, 2009), HCI.

In this paper, we propose a real-time iris center detection and tracking approach that uses the gradient vectors of an image within an estimated region of eyes. The existing researches on iris center detection and tracking system can be classified into two broad categories: (i) non-intrusive, and (ii) head mounted devices. In non-intrusive approach, the researchers use external cameras to detect the eye center of the target human (E. Murphy-Chutorian & M. Trivedi 2009) (N. Robertson & I. Reid, 2006). However, in this approach the iris center detection task is very difficult if we use only eye or head information for iris because of the variability of the human head orientation. On the other hand, the head mounted device allows detecting and tracking iris center without having the difficulties of the orientation of the subject's head. In this research area, most of the existing system uses infra-red (IR) lighting properties to accurately detect the pupil position in the eyes (Z. Zhu, K. Fujimura & Q. Ji, 2002) (T. Ohno & N. Mukawa, 2004). However, when head mounted devices are difficult to use or IR lighting is impracticable, we can use image based methods to estimate the center of the iris position. In this research, we use a multi-stage image based method to detect and track the center of the iris position. In our approach we first detect and track the face area in the input video images using efficient face detection and tracking technique. Then based on the detected face area, we extract the facial feature points on the face images. Depending on the facial feature point, we estimate the possible eye location and rectangular area around the eye location. Finally, within the estimated eye region, we detect and track the iris center using the VFOID.

2. PROPOSED APPROACH

The proposed iris center detection and tracking system is illustrated in Figure 1. Here, we use a USB camera situated in an environment to record the eye images of the target person. The recorded video image is used to detect and track the iris center. For this purpose, the system first detects and tracks the face in the input video images. From the detected face region, we also calculate the head position in the image. Within the detected face region, the system tries to find the facial feature points. Based on the facial feature points, we estimate the eye rectangular regions. The rectangular eye regions are used to detect and track the iris center.

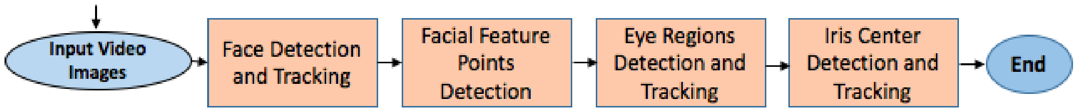

Figure 1. Overview of the proposed iris detection system.

3. FACE DETECTION AND TRACKING

We use 3D head tracking approach proposed by Kobayashi et al. in (Kobayashi, et al., 2006). In the approach, first the AdaBoost-based cascaded classifiers are used for hypothesis evaluation. Then, multiple cascaded classifiers, trained respectively to detect one direction of a human head, are adaptively used in hypothesis evaluation by considering the relationship between human head orientation and camera position. Consequently, tracking a human head robustly and accurately is realized even in the case of observation with low resolution or under varying light conditions.

The AdaBoost-based face detector proposed by Viola and Jones (Viola & Jones, Robust real-time face detection, 2004) has a cascade structure to reduce detection time (Figure 2). This cascade is effective in the evaluation phase even in the particle filter framework. In Figure 2, S_i represents a strong classifier. Each strong classifier classifies an input image into a positive or a negative. Only positive images are used as the input of the next strong classifier. At each stage, a strong classifier is trained to detect almost all face images while rejecting a certain fraction of non-face images. Thus the face detector detects almost all the face images and rejects almost all the non-face images. A strong classifier $S_i(x)$ at each stage of the cascade consists of many weak classifiers $s_t(x)$. This can be described as follows:

$$S_i(x) = \mathrm{sgn}\left(\sum_{t=1}^{T} \alpha_t s_t(x)\right) \quad (1)$$

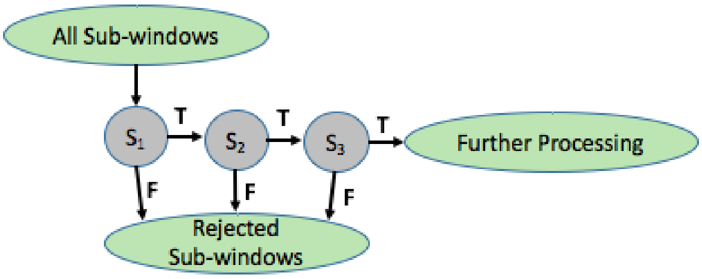

Figure 2. Cascade of classifier.

where T is the number of weak classifiers and $\alpha_t = \log\frac{1-\varepsilon_t}{\varepsilon_t}$. Here ε_t is an error rate specified in the training phase. Each weak classifier $s_t(x)$ evaluates a target image region by using Haar-like features. The detected head position, H_t using the head tracker in the input video frames is shown in Figure 3.

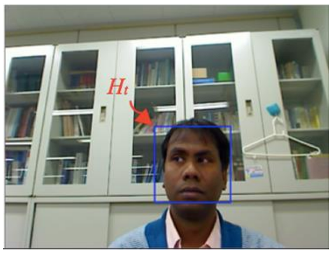

Figure 3. Detected head and its position within a video image.

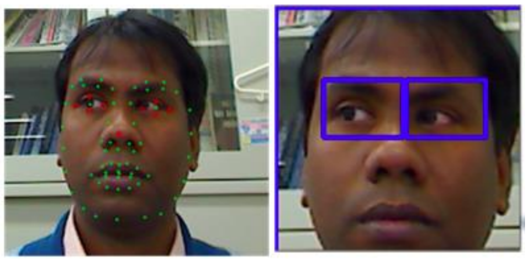

Figure 4. (a) Detected 68 facial feature points, and (b) eye regions detection based on facial feature points.

4. FACIAL FEATURE POINTS DETECTION

The proposed iris detection consists of identifying the pixels that belong to the iris boundary of an image. Because we are in presence of an image representing not only one eye but the full face and even extra scene components, we should consider the detection of face region using the face detection technique as described in the previous section, then the left and right eye regions, before detecting the iris position. To automatically detect the left and right eye regions, we determine the facial feature points. Based on the head location and its area, we detect and track the facial feature points using the active shape model (ASM) (Cootes, Taylor, Cooper, & Graham, 1995) (Figure 4 (a)). The facial feature points are used to roughly estimate the eye regions on the face (Figure 4 (b)).

5. IRIS CENTER DETECTION AND TRACKING

The VFOID provides important information for circular object. The center of a circular object can be detected by analyzing the VFOID as illustrated in Figure 5. We can use same technique for iris center detection if the iris is considered as a circular object. For instant, Kothari et al. (Kothari & Mitchell, 1996) proposed a method that utilizes the flow field characteristics that generates due to the strong contrast between iris and sclera. In their approach, the orientation of each gradient vector is used to draw a line through the whole image and then increases an accumulator bin each time one such line passes through it. The iris center is estimated as the accumulator bin where most of the lines intersect. However, their approach is only defined in the discrete image space and a mathematical formulation is missing. Moreover, they don't consider problems that arise due to eyebrows, eyelids, or glasses. In this research, we use the VFOID within the eye regions to detect the center of the iris within an eye as indicated by the red points in the Figure 6(b).

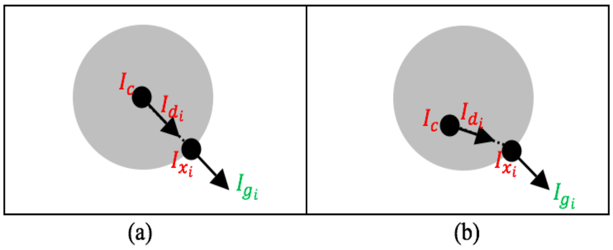

Figure 5. Abstract view with a dark circle (iris) on a light background (sclera): (a) displacement and the gradient vectors have the same orientations; (b) displacement and the gradient vectors do not have the same orientation.

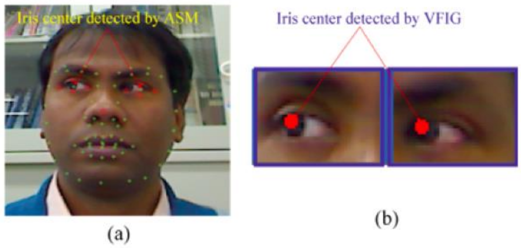

Figure 6. Iris center detected by (a) ASM landmark points, and (b) VFIG.

Although the facial feature points detected by the ASM model include the eye pupil points, their accuracy is not good for the eye pupil detection as illustrated in Figure 6(a). Thus, we propose the VFIG iris center detection method to accurately detect the iris center in the eye image as follows.

Let I_c be the possible iris center and I_{g_i} the gradient vector at position I_{x_i} . If I_{d_i} is the normalized displacement vector, then it should have the same absolute orientation as the gradient I_{g_i}. We can determine the optimal center I_c^* of the iris (darkest position of the eye) by computing the dot products of $I_{d_i}$and I_{g_i} and finding the global maximum of the dot product over the eye image. Mathematically,

$$I_c^* = \underset{I_c}{\operatorname{argmax}} \left\{ \frac{1}{N} \sum_{i=1}^{N} (I_{d_i}^T I_{g_i})^2 \right\} \quad (2)$$

where, $I_{d_i} = \frac{(I_{x_i} - I_c)}{(\|I_{x_i} - I_c\|_2)}$

$i = 1, \cdots\cdots, N$ and $\forall_x: \|I_{g_i}\|_2 = 1$. The displacement vector, I_{d_i} is scaled to unit length in order to obtain an equal weight for all pixel positions in the image.

We also scaled the gradient vector I_{g_i} to unit to tolerate the method against linearity relationship between illumination and contrast of the eye images. To reduce the computational complexity, the proposed method considers only gradient vectors with a significant magnitude, i.e. ignoring gradients vector in homogeneous regions. To obtain the image gradients in an eye image, we compute the partial derivatives using the following equation:

$$I_{g_i} = \left(\frac{\partial I(x_i, y_i)}{\partial x_i}, \frac{\partial I(x_i, y_i)}{\partial y_i} \right)^T \quad (3)$$

However, other methods for computing image gradients will not change the behavior of the objective function significantly.

6. EXPERIMENTAL RESULTS

6.1. Database

To evaluate the proposed iris center detection technique, we used the BioID face database (BioID face database). It is one of the most challenging databases and many eye center detection researchers use this database to evaluate their system. The database consisted of 1521 face images of 23 different peoples. The faces of the database were in gray-level and were collected from different locations at different times at different poses. In some face images eyes were completely hidden by strong reflection on the glasses. The face image sizes were 286×384.

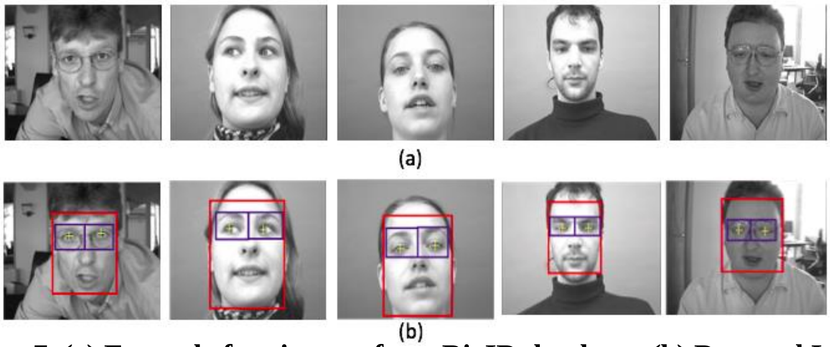

Figure 7. (a) Example face images from BioID database, (b) Detected Iris center.

The database also contained eye position files. The files indicated the *x* and the *y* coordinate of the left eye center (iris center) and the *x* and the *y* coordinate of the right eye center (iris center). Figure 7(a) show some face images from BioID database.

6.2. Evaluation Strategy

We used our face detection technique to detect the face rectangle on the images. Then based on the based on the facial feature points we estimated the eye regions. Finally, the eye region was used to detect the iris center. The accuracy of the iris centers were measured by the normalized error. The technique of this evaluation measure was proposed by Jesorsky et al. (Jesorsky, Kirchberg, & Frischholz, 2001) and is defined as:

$$e \leq \frac{1}{d}\max(e_l, e_r) \quad (4)$$

where e_l = Euclidean distances between estimated and correct left iris center position, e_r = Euclidean distances between estimated and correct right iris center position, and d is the distance between the correct iris's centers.

6.3. Experimental Results and Comparison

Our approach finds the performance for iris center localization which is obtained for $e \leq 0.05$. Figure 7(b) shows some example iris center detection results on the face images from BioID database. The proposed system accurately detected the iris center on the face in different looking conditions. Table 1 shows the performance comparison of our method with other methods for iris center localization. The proposed method performed better than the other three methods presented here.

Table 1: Performance comparison of our method with others.

Methods	Accuracy for $e \leq 0.05$
Asadifard & Shanbezadeh, 2010	47.0%
Kroon, Hanjalic, & Maas, 2008	65.0%
Valenti & Gevers, 2008	77.2%
Our method	***86.1%***

6.4 Real-Time Iris Tracking from Video Data

In real-time analysis of iris center detection, we experimented in a laboratory environment. Here, we placed a webcam in front of the target person's face and s/he was asked to look at different locations in the laboratory. The target person was asked to change her/his position of the head between 0.5 to 2.0 meters from the camera. The person also changed her/his head/eye in left and right directions. Our proposed system detected and tracked the center of the iris position continuously. Figure 8 shows some video frames to illustrate the performance of real-time iris center detection and tracking of our proposed method.

Figure 8. Iris center detection and tracking results on a video images

7. CONCLUSION

In this paper, fully automatic and real-time iris center detection and tracking system is presented. The proposed method can detects and tracks the iris center in various conditions such as background movement, scale changing, and environment illumination changing. The proposed face detection and tracking method effectively detect the face in different posed and scale variations. Then, the system automatically estimates the rough left and right eye regions from the facial feature points. Finally, the VFOID accurately detects the iris center. The proposed system is compared with some state-of-the-art to show the better performance of the system. We also evaluate the proposed method on a real-time tracking of iris center. The future work will focus on the processing of illumination variation, which can cause some highlight reflection on the cornea and increase the difficulty of robust detection of the eye center.

REFERENCES

[1] J. S. Agustin, J. P. Hansen and J. Mateo (2008), "Gaze beats mouse: Hands-free selection by combining gaze and emg," in CHI.

[2] E. G. Freedman (2008), "Coordination of the eyes and head during visual orienting," Experimental Brain Research, vol. 190, pp. 369-387.

[3] A. Doshi and M. M. Trivedi (2012), "Head and eye gaze dynamics during visual attention shifts in complex environments," Journal of Vision, vol. 12, no. 2, pp. 1-16.

[4] R. Valenti, N. Sebe and T. Gevers (2012), "Combining Head Pose and Eye Location Information for Gaze Estimation," IEEE Trans. on Image Proc., vol. 21, no. 2, pp. 802-815.

[5] J. S. Agustin, J. P. Hansen and J. Mateo (2008), "Gaze beats mouse: Hands-free selection by combining gaze and emg," in CHI.

[6] A. Doshi and M. M. Trivedi (2009), "On the roles of eye gaze and head pose in predicting driver's intent to change lanes," IEEE Trans. Intell. Transp. Syst., vol. 10, no. 3, pp. 453-462.

[7] E. Murphy-Chutorian and M. Trivedi (2009), "Head pose estimation in computer vision: A survey," IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 4, pp. 607-626.

[8] N. Robertson and I. Reid (2006), "Estimating gaze direction from low-resolution faces in video," in ECCV.

[9] Z. Zhu, K. Fujimura and Q. Ji (2002), "Real-time eye detection and tracking under various light conditions", The symposium on Eye tracking research & app. (ETRA '02)," NY, USA.

[10] T. Ohno and N. Mukawa (2004), "A free-head, simple calibration, gaze tracking system that enables gaze-based interaction," The symposium on Eye tracking research & app., NY, USA.

[11] D. W. Hansen and J. P. Hansen (2006), "Eye typing with common cameras," in Symp. Eye Track. Res. Appl..

[12] Y. Kobayashi, D. Sugimura, Y. Sato, H. Hirasawa, N. Suzuki, H. Kage and A. Sugimoto (2006), "3D Head Tracking using the Particle Filter with Cascaded Classifiers," in Proc. of the British Machine Vision Conference (BMVC2006).

[13] P. Viola and M. J. Jones (2004), "Robust real-time face detection," International Journal of Computer Vision, vol. 57, no. 2, pp. 137-154.

[14] T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham (1995), "Active shape models-their training and application," Computer Vis. and Image Understanding, vol. 61, no. 1, pp. 38-59.

[15] R. Kothari and J. Mitchell (1996), "Detection of eye locations in unconstrained visual images," in IEEE ICIP.

[16] "BioID face database," https://www.bioid.com/downloads/software/bioid-face-database.html.

[17] O. Jesorsky, K. J. Kirchberg and R. W. Frischholz (2001), "Robust Face Detection Using the Hausdorff Distance," in Audio and Video based Person Authentication.

[18] M. Asadifard and J. Shanbezadeh (2010), "Automatic adaptive center of pupil detection using face detection and cdf analysis," in IMECS.

[19] B. Kroon, A. Hanjalic and S. Maas (2008), "Eye localization for face matching: is it always useful and under what conditions?," in CIVR.