Optimizing the Recognition of Eye States in Real Time Video

[1]Padiyappa Koogati & [2]K. Vinay Kumar

[1]Dept of Computer Science and Engineering, Maratha Mandal Engineering College, Belgaum, India.
[2]Dept. of CSE, National Institute Of Technology Karnataka, Surathkal, India
E-mail : padiyappabkoogati@gmail.com

***Abstract* - Most of the current eye tracking system suffers from slow response time and less accuracy which are not proper for real time applications. In order to overcome these problems a new algorithm has been proposed. In this proposed algorithm, the search area is limited by Region of Interest and then detects the face region followed by extracting the eyes and identifies the eye states.**

Keywords - face detection, Haar-like features, eye state, ROI.

I. INTRODUCTION

Many research results are published on real-time face and eye recognition application. Recent applications like for detecting the drowsiness of a driver also many efforts have been made in the past, but all they suffers from low response time and accuracy [1]. In real time applications this is not feasible. In this proposed method, the search area is limited by Region of Interest and then detects the face and eye region by Haar features followed by identifying the eye states by finding the intensity changes on the face. This algorithm is to implement a video monitoring system to identify the opened and shut states of eye in real-time video to determine the eye states, which is to be used as an important reference of getting drowsy of a driver.

The rest of this paper is organized as follows. Section 2 presents related work. Section 3 presents the proposed method. Section 4 describes the experimental results and finally in Section 5 conclusions are stated.

II. RELATED WORK

The role of smart and intelligent devices, capable of monitoring human eyes is increasing with the wide popularity of user centric applications. Eye tracking has been applied to new fields, such as drowsy-driver detection, display energy management in addition to traditional applications such as human-computer interface, security, health care and commercial applications [2].

The approach used in this proposed method is to track eyes with the aid of ordinary camera [3]. The techniques can be broadly classified into three groups: template-matching, appearance-based and feature-based. The template-based techniques are based on generic eye models and their recursive positioning, translation, rotation and deformation to best fit the eye representation in the image. While these techniques can detect eyes accurately, they require good image contrast to converge correctly and are computationally expensive. The appearance-based techniques detect eyes based on their photometric appearance. These techniques usually employ large amount of training data representing the eyes of different subjects under different face orientations and illumination conditions. These data are used to train a classifier such as Neural Networks (NN) or Support Vector Machine (SVM) and detection is achieved via classification. Since NN treat each frame independently, many redundant data have to be processed.

The feature-based techniques identify eyes by exploring their distinctive features, such as eye-corner points, color distribution of the eyes, edge and intensity of iris, intensity variation around the eyes, etc.

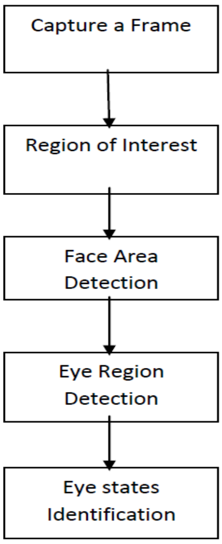

Figure 1. Proposed Method

III. PROPOSED METHOD

At first limit the search area by Region of Interest and then detect the face and eye region by Haar features followed by identifying the eye states finding the intensity changes on the face.

Figure 1 shows a general process for eye state identification from real time video. The video frames are captured sequentially, and for each frame, the region of interest is obtained. Then the face area is detected and extracted from the obtained frame, and finally the eye regions are separated. Some pre-processing steps are applied on eye regions to obtain the exact eyes without eyebrow or glasses borders. And then, the eye state is identified as opened or shut.

A. Region of Interest

The region of interest is a particular region in a scene in which we are interested shown in below figure 2. Autonomous agents have limited capability in memory, computation, decision making and communication for efficient co-operations. Therefore, it is essential to extract that region from the scene which has significant information. In order to extract significant region there need to determine its cognitive boundary.

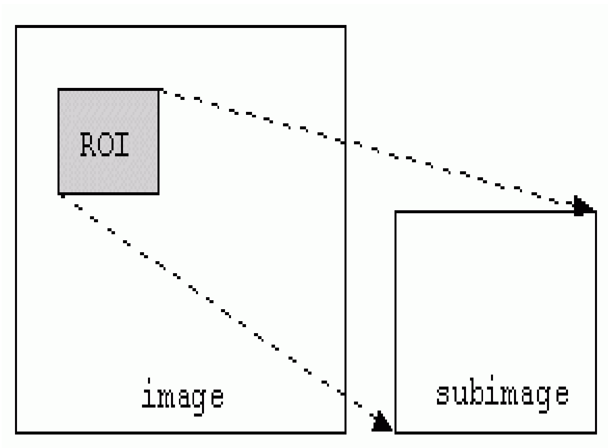

Figure 2. An image with Region of Interest defined

The search area reduction can be done by detecting the region of interest (ROI) [4]. Since we are interested in tracking eye locations in a sequence of real-time images, either we detect the movement in the images if there is any or limit the search area from the previously detected face area otherwise. If there is no significant motion in the current image, we see for the previously tracked face area if there is any, or set the entire image as the search area otherwise.

B. Face Detection

Yang et al. [5] presented an overview of face detection methods developed in recent years. Papageorgiou et al. [6] introduced a scheme of boosted cascade algorithm using Haar-like features as identification patterns. Lienhart and Maid [7] enhanced the algorithm, while Viola and Jones [8], [9] used the algorithm in face detection system.

A recognition process can be much more efficient if it is based on the detection of features that encode some information about the class to be detected. This is the case of Haar-like features that encode the existence of oriented contrasts between regions in the image. A set of these features can be used to encode the contrasts exhibited by a human face and their special relationships. Haar-like features are so called because they are computed similar to the coefficients in Haar wavelet transforms. The Haar-like feature value is computed as the difference between the sum of pixel values of white area and black area. Haar-like features encode differences in average intensities between different regions in an image. They may contain rich information about an object such as an eye region is darker than that of cheek and the bridge of nose is brighter than its sides.

The classifier in this paper we used is taken from the Emgu CV, a cross platform .Net wrapper to the Intel Open Source Computer Vision Library (OpenCV) for face detection and eye detection [10]. Classifier is an object trained with a few hundreds of sample views of a particular objet (a face or a car), called positive examples. This trained classifier can be applied to a region of interest in an image. The classifier outputs '1' if the region is likely to show the object (face/car) and '0' otherwise. The frame will be simply discarded if no face area is detected in a frame. Although some of frames which contain faces are discarded wrongly, the accuracy achieved is about 99% for face area detection.

C. Eye Detection

The Haar-like features have been used to detect both face and eyes successfully in many applications [1]. The Haar-like feature value is computed as the difference between the sum of pixel values of white area and black area. Two values computed by applying Haar-like filters are threshold to detect the eye region. The eye region is shown by dark in below figure 3.

The average detection accuracy achieved for the normal eye regions are about 98%, for closed eyes and those eyes with glasses are 95%.

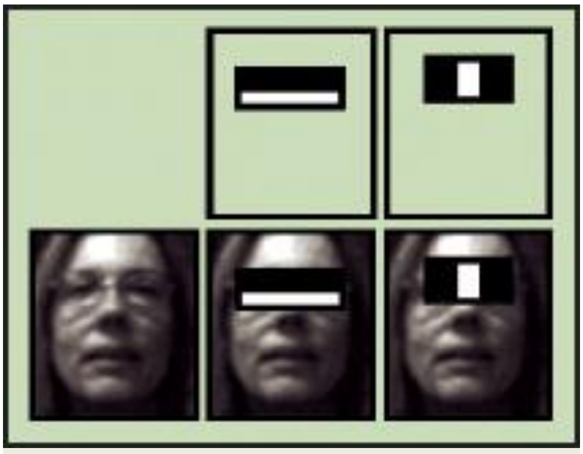

Figure 3. Haar features and result of eyes detection

D. *Eye states Identification*

The next step in locating the eyes is finding the intensity changes on the face. This is done using the gray scale image. Once the eyes are located, by using the intensity variations the distance between eyebrow and eyelids is measured [12]. This distance is maximum when eyes are completely closed and minimum when eyes are opened as shown in below figure. 4. The state of the eyes (whether they are open or closed) is determined by the distance between the first two intensity changes found.

(a) Open Eye (b) Closed Eye

Figure 4. Eye States Identification

IV. Experimental Results

An ordinary webcam is used to capture the videos, with its resolution: 800 x 600, frame frequency (FPS) is 30. The size of frames captured is 320 x 240 pixels. We used a database of about 2000 positive and 1000 negative example images for checking the accuracy. The frame will be simply discarded if no face area is detected in a frame. Although some of frames which contain faces are discarded wrongly, the accuracy achieved is about 99% for face area detection and the average eye state detection accuracy is more than 95%.

Table 1 shows accuracy rate of detection for different methods.

Method	Accuracy
Template Matching	85.67 %
KNN classifier	92 %
Color Segmentation And Hough Transform	93.5 %
Proposed Method	95 %

Table 1 Accuracy rate of detection for different methods

V. References

[1] Lei Yunqi, Yuan Meiling, Song Xiaobing, Liu Xiuxia, Ouyang Jiangfan, "Recognition of Eye States in Real Time Video", International Conference on Computer Engineering and Technology, 2009.

[2] Kara Pernice and Jakob Nielsen, "Eyetracking Methodology: How to Conduct and Evaluate Usability Studies Using Eyetracking", International Conference on Image Processing, August 2009.

[3] Shinji Yamamoto and Vasily G.Moshnyaga, "Algorithm Optimizations for Low-Complexity Eye Tracking", Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics San Antonio, TX, USA - October 2009.

[4] Young-Shin Lim and Hyonam Joo, Joon-Seek Kim, "A Realtime Pattern Detection using Fast Feature Matching Algorithm", International Conference on Control, Automation and Systems 2008, Hoseo University, Asan, Korea.

[5] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey", Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol 24(1), 2002, pp. 34- 58.

[6] C.P. Papageorgiou, M. Oren, and T. Poggio, "A general framework for object detection", A general framework for object detection, 1998, pp. 55-562.

[7] R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection", Image Processing.2002. Proceedings 2002 International Conference, vol 1, 2002, pp. I-900-I-903.

[8] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", Proceedings of the 2001 IEEE Computer Society Conference, vol 1, 2001, pp. I-511-I-518.

[9] The Ephraim Tristan Himmelman Kaleem Siddiqi, "Real-Time Viola-Jones Face Detection in a Web Browser", Canadian Conference on Computer and Robot Vision- 2009.

[10] Intel, Emgu CV, "Open source computer vision library".

[11] Weston Sewell, Oleg Komogortsev, "Real-Time Eye Gaze Tracking With an Unmodified Commodity Webcam Employing a Neural Network", ACM Conference on Human Factors in Computing Systems (CHI), Atlanta, GA, 2010.

[12] Mandalapu Sarada Devi, Dr Preeti R Bajaj, "Driver Fatigue Detection Based on Eye Tracking", First International Conference on Emerging Trends in Engineering and Technology- 2008.

◈◈◈