Iris Region and Bayes Classifier for Robust Open or Closed Eye Detection

Prasertsak Tiawongsombat*

Abstract

This paper presents a robust method to detect sequence of state open or closed of eye in low-resolution image which can finally lead to efficient eye blink detection for practical use. Eye states and eye blink detection play an important role in human-computer interaction (HCI) systems. Eye blinks can be used as communication method for people with severe disability providing an alternate input modality to control a computer or as detection method for a driver's drowsiness. The proposed approach is based on an analysis of eye and skin in eye region image. Evidently, the iris and sclera regions increase as a person opens an eye and decrease while an eye is closing. In particular, the distributions of these eye components, during each eye state, form a bell-like shape. By using color tone differences, the iris and sclera regions can be extracted from the skin. Next, a naive Bayes classifier effectively classifies the eye states. Further, a study also shows that iris region as a feature gives better detection rate over sclera region as a feature. The approach works online with low-resolution image and in typical lighting conditions. It was successfully tested in 9 image sequences (2,210 frames) and achieved high accuracy of over 92% for open eye and over 86% for closed eye compared to the ground truth. In particular, it improves almost 15% in terms of open eye state detection compared to a recent commonly used approach, template matching algorithm.

Keywords : human-computer interaction, open or closed eye detection, eye blink detection, iris region, naive Bayes classifier

Department of Electronics Engineering Technology, College of Industrial Technology, King Mongkut's University of Technology North Bangkok.

* Corresponding author, E-mail: prasertsak.t@cit.kmutnb.ac.th Received 7 November 2014, Accepted 15 January 2015

1. Introduction

In recent years, there has been an effort to augment traditional human-computer interfaces like keyboard and mouse with natural interfaces allowing users to interact with the computer more naturally and effectively. This intelligent system also opens an opportunity for people who cannot use the keyboard or mouse due to severe disabilities to gain access and control over a computer again.

As is known, the people with severe disabilities often have great difficulties communicating their desires, thoughts, and needs. In particular, the ability to control eye muscles is sometimes the only remaining voluntary movement they have. Hence, to determine eye statuses (i.e., open or closed eye) and use them as natural communicative cues, computer vision techniques seem to be an effective solution since they are nonintrusive, comfortable, reliable, and inexpensive.

Preliminary approaches to video-based computer interfaces for people with disabilities are described in [1-5]. In general, an appearance is a popular feature used for shape description in an image. However, the appearance-based approach requires proper image quality to accurately obtain the position features. Therefore, in the cases of low-resolution images or inaccurate eye location images, the approach obviously has difficulty in efficiently building a model representing eye shape. Furthermore, it also has difficulty in locating/tracking the eye contours of different individuals and under different illumination.

J. Betke et al. [6] creates a system called “Camera Mouse” which detects body motion as feature in order to control mouse pointer on a computer. Template matching approach using normalized correlation coefficient algorithm is employed to track the chosen feature. Grauman et al. [7] also employs a similar technique to determine eye blink patterns.

Benoit and Caplier [8] propose method to estimate state open or closed of eye to determine blinking for application of hypovigilence state of driver detection. A particular electronic circuit, called retina filter [9-10] is used to enhance contours in eye region image. As a result, calculated level of energy from the filtered image, which is proportional to the quantity of contours, are used to determine eye states. The spectrum energy in case of an open eye is always higher than those obtained in case of a closed eye.

Bacivarov et al. [11] develop a statistical active appearance model (AAM) to track and detect eye blinking. The approach has been designed to be robust to variations of head pose. However, a proper image resolution is required to annotate landmark points for a well descriptive eye model. Furthermore, AAM is usually computationally expensive.

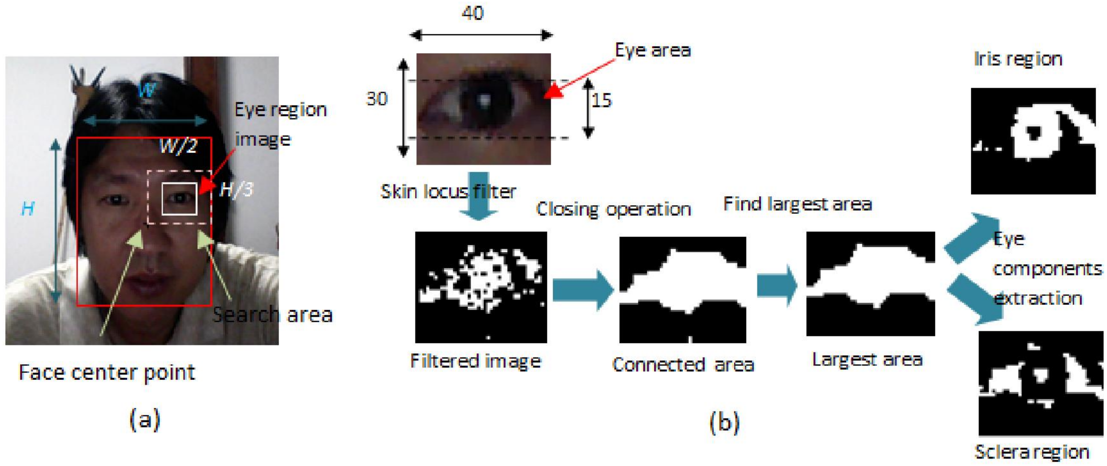

Fig. 1. Eye components extraction procedure.

To overcome the need of high-resolution eye region image, which is not always available in practice, and to avoid excessive computation, this paper proposes the use of eye component region (i.e., iris or sclera) as a robust feature and the use of Bayes classifier for simple and fast eye state classification approach.

This paper is organized as follows. Section 2 describes a procedure for extracting eye components. In Section 3, the probability distributions of eye component regions are introduced. The proposed method for open or closed eye classification is detailed in Section 4. The experimental results and the conclusion can be found in Section 5 and 6, respectively.

2. Eye component extracting procedure

In this work, in order to acquire the proposed eye components (i.e., iris and sclera regions) from an eye region image, several processing steps are required (see Fig. 1).

At the very first frame of a detected frontal face, an eye region image and face center point are found. The face and eye detector used in this work is based on Haar Cascade Classifier for face and eye [12-13]. In particular, by utilizing knowledge of human face geometry, the candidate search area for an eye is determined to have a height of *H/3* and a width of *W/2* (see Fig. 1a), where *H* and *W* indicate height and width of a rectangle surrounding the detected face. After the eye region is detected, the rectangle enclosing it is anchored by the face center point, which is tracked using an optical flow algorithm, described in [14]. Fig.

2 depicts samples of detected eye region image in various illumination and skin colors.

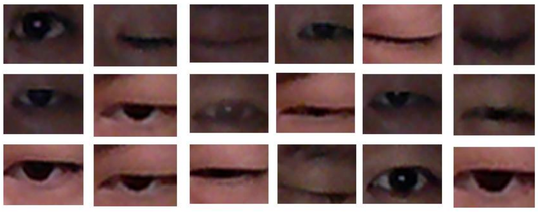

Fig. 2. Sample images of detected eye region images.

Consider the detected eye region image which has a normalized size of 40×30 pixels. An eye location is assumed to be located in the middle of the image (see Fig. 1b). To separate skin pixels from eye pixels, skin locus is created and used as a filter. The skin locus, described in [15], is the range of skin colors under light sources of varying correlated color temperature, in normalized color coordinates, that follows the curvature of the Planckian locus [16].

Next, closing operation is applied to the image to form the connected area. As a result, the largest area found within the eye area is considered as an eye and used as a mask on the original eye region image to narrow the search for eye components. The iris pixels can be extracted by searching for the pixels in the eye area with their luma Y' lesser than the mean illumination $\overline{Y}'$ of detected eye area (i.e., iris is generally darker than sclera). Hence, the remaining pixels with Y' greater or equal to $\overline{Y}'$ are considered as sclera pixels.

3. Probability distributions of eye components during different eye states

Generally, an appearance is a popular feature used for shape description in an image. However, the appearance-based approach requires proper image quality, which practically not always available, to accurately obtain the position features. Hence, low performance might often be a result in most cases in practice.

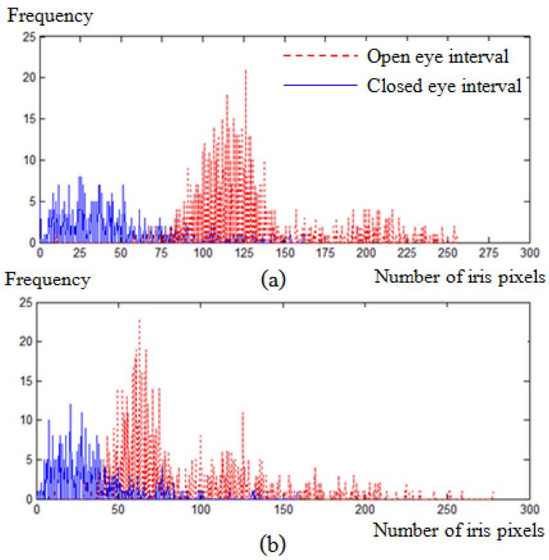

Fig. 3. Distributions of eye components during open and closed eye intervals. (a) Number of iris pixels and (b) number of sclera pixels.

In this work, two eye components, iris and sclera, are observed and studied in practical low-resolution eye region image. Intuitively, there are changes in the number of eye component pixels during eye blink, in which the number of eye component pixels decreases while a person is closing an eye and increases while an eye opens.

According to the study [17] and the frequency plots shown in Fig. 3, it can be seen that the two distributions for open and closed eye intervals of each eye component are clearly separable. The sample data were collected from four individual without disabilities consisting of a total of 986 frames. The room's illumination conditions were natural and not controlled and the blinking actions were not directed in all collected image sequences. Explicitly, the overlapped area in the case of sclera pixels (see Fig. 3b) seems to be larger than the ones in the case of iris pixels (see Fig. 3a). As expected, there seems to be more false detected pixels (i.e., skin pixels recognized as iris or sclera pixels) in the extraction of sclera than in the extraction of iris pixels.

Note also that the distributions during each eye interval evidently form "bell-shaped" curve. Hence, they can be sufficiently estimated by Gaussian function and finally can be employed for eye state classification purpose. In this paper, only the iris pixels are further considered and used as feature since the study [17] reported that using iris as feature achieved higher accuracy for classifying open or closed eye state than using sclera as feature.

4. Eye states classification using naive Bayes classifier

As is known, a difficulty in classification problem generally depends on how complex observations are. In particular, if the chosen observations are not meaningful and not well-descriptive for classified categories, the system might result in low performance of classifying rate.

In this paper, a naive Bayes classifier is chosen to be applied as a classifier since it is a simple, fast, and efficient probabilistic classifier. Particularly, a single feature is only considered in this work. Hence, the corresponding Bayes classifier can be simplified and by combining it with a decision rule, which is to pick a class with the most probable hypothesis, Bayes classifier for eye states classification is defined as follows:

$$eye_t^* = \underset{eye_t}{\operatorname{argmax}} \; [p(eye_t)\, p(f_t|eye_t)] \quad (1)$$

where $p(eye_t)$ is the prior probability distribution of eye state at time t and $eye_t \in \{eye_o, eye_c\}$ is the eye state at t, open and closed eye, respectively. $p(f_t|eye_t)$ denotes the probability of observation f_t given eye state eye_t where f_t is an observed number of iris pixels at t. eye_t^* indicates the classified eye state of a person at t.

5. Experimental results

To evaluate the performance of the proposed method, the method was tested in 9 image sequences from four individuals without disabilities consisting of a total of 2,210 frames, displaying a single individual sitting and blinking in front of a camera in a typical room with uncontrolled lighting conditions (see Fig. 4).

Fig. 4. Sample images.

Particularly, the blinking actions were not directed in all collected image sequences and all 9 image sequences were not used during the training phase of Gaussian distribution estimation of each eye state. Further, the ground truth has been visually determined by manually marking frames of the image sequences, as open or closed eye states. Finally, to evaluate the performance of the proposed approach, the classification results were compared against to a recent commonly used approach, template matching algorithm [6] and the ground truth.

In the experiments, the skin locus was created from the skin pixels collecting from UBIRIS [18], noisy iris image database, and from our database. The probability distributions of iris region for each eye state were estimated by Gaussian function during the training phase and $p(eye_t)$ was set to 0.5 for each eye state because they were assumed to have equal chances to occur in our case. For the template matching algorithm [6], the predefined classification threshold was set to 0.5. If the calculated correlation value is greater than or equal to the threshold, it is classified as an open eye state, otherwise it is defined as a closed eye state.

Fig. 5 depicts the eye state classification outcomes of one image sequence of an individual. In particular, Fig. 5a illustrates the sequence of the number of iris pixels over time. Fig. 5b shows the open or closed eye classification results. Furthermore, the ground truth is also indicated and the recognized blinking events are also shown.

As expected, the number of iris pixels changed respectively during each eye state interval. Note that there were several times that inaccurate observation of the number of iris pixels occurred, such as during frame 75 - 83 and frame 222 - 240. Understandably, this caused by bad lighting condition - either too bright or too dark. However, if the incorrect observation of iris pixels did not repeatedly and consecutively occur, the proposed approach could still robustly perform good eye state classification outcomes which could successfully lead to reliable eye blink recognition (see Fig. 5b).

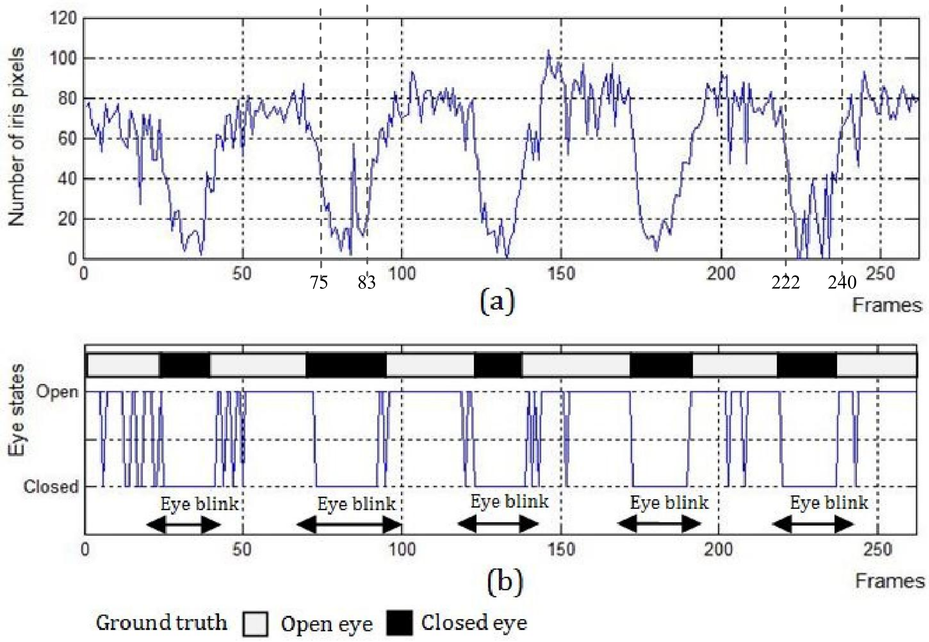

Fig. 5. Open or closed eye classification result from one of the image sequences.

Intuitively, there exist false detections of eye states that may lead to incorrect eye blink detection. As a result, post-processing may be required to filter out undesired frequent changes of eye states, for example a predefined 10 consecutive frames might be set as window scope to ensure the smooth state transition of one eye state to another. This means if the currently recognized eye state does not occur over consecutive 10 frames, the outcome of recognized eye state remains as the previous detected one. As is known, eye blink simply consists of open-closed-open eye sequence. Therefore, to effectively determine the eye blink event, one can simply determine the open-closed-open sequence from post-processed outcome.

Table 1 summarizes results comparing the performance of the proposed approach and the template matching algorithm against the ground truth. It shows that the proposed method performs well with very high classification rates for open eye ($>92\%$) and closed eye ($>86\%$). In particular, most incorrectly classified eye states often occurred when the lighting conditions were extremely poor (too bright or too dark), since in these conditions, it becomes difficult to separate the eye area from the skin. As a result, incorrect detected iris region may often be a result.

Table 1 Summary of results

Eye state interval	**Accuracy (%)**	
	The proposed approach	Template matching approach [6]
Open eye	92.5%	76.2%
Closed eye	86.3%	85.1%

6. Conclusion

A robust approach for open or closed eye classification, which can lead to an effective application for eye blink detection, has been presented. The proposed approach is based on an analysis of eye and skin in eye region image. Evidently, eye component regions (i.e., iris and sclera) change their sizes in eye region image as a person is closing or opening an eye. In particular, probability distributions of these eye components during open and closed eye states are clearly separable and showed robustness for classification purpose. A study also showed that using iris as feature is preferable and produced a better classification results than using sclera as a feature. Finally, by utilizing a Bayes classifier combined with decision rule, the eye states can be efficiently recognized. Consequently, the eye blink can be determined from the detected frequencies of eye states. The approach works online with low-resolution image and in typical lighting conditions. It was successfully tested in 9 image sequences (2,210 frames) and achieved high accuracy of over 92% for open eye and over 86% for closed eye. Particularly, it improves almost 15% in terms of open eye state detection compared to a recent commonly used approach, template matching algorithm.

In future work, an author plans to develop also a robust approach for classifying eye movements (i.e., moving left or right). As a result, with the proposed approach and the ongoing research method, the practical eye-controlled application, which gives various options and ways for users with disabilities to command computer with eyes, can be achieved.

7. Acknowledgements

This research was funded by King Mongkut's University of Technology North Bangkok. Contact no. KMUTNB-NEW-57-10. Furthermore, the author thanks Asst. Prof. Dr. Supot Sookpotharom and Dr. Pakorn Ubolkosold for their supports.

8. Reference

[1] H. R. Arabnia, "A computer input device for medically impaired users of computer," in Proceedings of Johns Hopkins National Search for Computing Applications to Assist Persons With Disabilities, 1992, pp. 131-134.

[2] H. Sako, M. Whitehouse, A. Smith, and A. Sutherland, "Real-time facial-feature tracking based on matching techniques and its applications," in Proceedings of 12th IAPR International Conference, Pattern Recognition, 2, Conference B; Computer Vision & Image Processing, 1994, pp. 320-324.

[3] O. Takami, K. Morimoto, T. Ochiai, and T. Ishimatsu, "Computer interface to use head and eyeball movement for handicapped people," IEEE International Conference Systems, Man and Cybernetics: Intelligent Systems for the 21st Century, 2, 1995, pp. 1119-1123.

[4] T. Ohtani, H. Ichihashi, T. Miyoshi, and N. Tani, "A pointing device using coordinate transformation for neurofuzzy GMDH," in Proceedings of 2nd International Conference Knowledge-Based Intelligent Electronics Systems, 2, 1998, pp. 108-115.

[5] J. Gips, M. Betke, and P. A. DiMattia, "Early experiences using visual tracking for computer access by people with profound physical disabilities," in Universal Access in HCI: Toward an Information Society for All, C. Stephanidis, Ed. Mahwah, NJ: Lawrence Erlbaum, 3, 2001, pp. 914-918.

[6] M. Betke, J. Gips, and P. Fleming, "The Camera Mouse: Visual tracking of body features to provide computer access for People with Severe disabilities," IEEE Transactions on Neural Systems and Rehabilitation Engineering, 10(1), 2002.

[7] K. Grauman, M. Betke, J. Gips, and G. R. Bradski, "Communication via Eye Blinks - Detection and duration analysis in real time," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2001.

[8] A. Benoit and A. Caplier, "Hypovigilence Analysis: Open or closed eye or mouth? blinking or yawning frequency?," in Proceedings of IEEE Conference on Advanced Video and Signal based Surveillance, 2005, pp. 207-212.

[9] J. Hérault and W. Beaudot, "Motion processing in the retina: about velocity matched filter," in European Symposium on Artificial Neural Networks, 1993.

[10] J. Hérault and B. Durette, "Modeling visual perception for image processing," Lecture Notes in Computer Science, 4507, 2007, pp. 662-675.

[11] I. Bacivarov, M. Ionita, and P. Corcoran, "Statistical models of appearance for eye tracking and eye-blink detection and measurement," IEEE Transactions on Consumer Electronics, 54(3), 2008.

[12] P. Viola and M. Jones, "Robust real-time face detection," International Journal of Computer Vision, 57, 2004, pp. 137-154.

[13] OpenCV, Available: http://opencv.org/

[14] J. Y. Bouguet, "Pyramidal Implementation of the Lucas Kanade Feature Tracker," Intel Corporation, Microprocessor Research Labs, 2000.

[15] M. Soriano, B. Martinkauppi, S. Huovinen, and M. Laaksonen, "Adaptive skin color modeling using the skin locus for selecting training pixels," Pattern Recognition, 36, 2003, pp. 681-690.

[16] E. Marszalec, B. Martinkauppi, M. Soriano, and Pietikäinen, "Physics-based face database for color research," Journal of Electronic Imaging, 9(1), 2000, pp. 32-38.

[17] P. Tiawongsombat and C. Rattanapoka, "A study of two robust features for effective open or closed eye classification", the 2015 International Electrical Engineering Congress (iEECON2015), 2015.

[18] H. Proença and L. A. Alexandre, "UBIRIS: a noisy iris image database," in Proceedings of 13th International Conference on Image Analysis and Processing, 1, 2005, pp. 970-977.