An Eye Blink detection technique in video surveillance based on Eye Aspect Ratio

Tran Thanh Phuong*, Lam Thanh Hien*, Do Nang Toan**, Ngo Duc Vinh***

*Lac Hong University, Vietnam

**Intitute of Information Technology, Vietnam Academy of Science and Technology, Vietnam

***Hanoi University of Industry, Vietnam

thanhphuong@lhu.edu.vn, lthien@lhu.edu.vn, donangtoan@gmail.com, ngoducvinh@haui.edu.vn

***Abstract*— Detecting the open/closed state of the eyes is an important step in determining the state of the face in applications related to human facial expressions. This report deals with an eye open/closed detection technique based on the variation of the ratio rectangle surrounding the eye. The technique has been tested and proven effective in determining sleepiness based on images.**

***Keywords*— Eye detection, eye state recognition, eye blink detection**

I. Introduction

As one of the most prominent features of the human face, the eyes play an important role in analyzing and estimating the changing state and emotions of a person. The open/closed state of the eyes is considered a powerful feature of the human face and has been studied over the years for a large number of applications, such as monitoring driver fatigue and sleepiness for automobile accident prevention [1][2], psychiatric diagnosis [3], emotion recognition [4], Human Computer Interfaces (HCI) [5]. Psychological research reports that blinking is an interpretation of a person's activity [6], for example: we blink faster when excited, when speaking in public, tense while attending an interview interrogation and when lying. In addition, detecting the open/closed state of the eyes is also applied in the system to estimate the student's concentration during class time [7].

Although there have been many applications related to estimating the state of the eye, this is an issue that is still of interest to researchers because of the variety of data and applications. Various methods have been proposed and are divided into two main groups: direct and indirect.

As for the direct method, which is performed by attaching a device around the eye to record the changes in eye movements [8], this method has high accuracy, but is difficult to implement in practice because of the inconvenient when having to attach additional sensors around the eyes.

For the indirect method, the system does not require any additional hardware and physical contact with people, but only analyzes based on images or video captured during monitoring. Approaches for the method include: pattern-based [9], appearance-based [10] and feature-based [11].

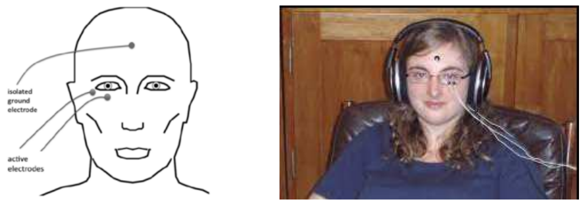


Figure 1. Position of electrodes to record blink intensity

i. Pattern-based techniques, the eye image matches the pattern of the common eye model. Patterns can distort to match the best representation of the eye in the image, mainly used to improve accuracy.
ii. In appearance-based techniques, eyes are detected based on a set of trained original images without eye patterns and features. A trained classifier such as a neural network or a support vector machine is used to detect eyes in different lighting conditions and face orientations.
iii. Feature-based techniques, using distinct features around the eye such as eyelid contour and intensity or color of the iris.

The performance of these techniques is highly dependent on sample data and is very sensitive to the resolution of the image. For technique (iii) has the advantage that there is no need to train the sample data first. However, there are some limitations on noise when detecting eye area such as: partially covered eyelashes, lighting conditions or eye diseases.

In this paper, we propose a technique to detect the open/closed state of the eyes based on the rectangular aspect ratio containing the eyes, instead of having to compare with a fixed threshold, the technique will flexibly calculate for all objects for decision making. The rest of the paper is presented as follows: Section 2 presents some related studies. Part 3 is the proposed technique. This is followed by testing and final is conclusions on the proposed technique.

II. Related Work

For blink detection, Grauman et al. [12] proposed a method based on pattern by using a trained open eye sample and then matching it to a real image for blink detection. They measured the similarity between actual blinks and eye samples. Kawato and Ohya [13] also detect blink and its duration based on the correlation score of the training sample and the actual image of the eye.

Mohanakrishnan et al [14] propose a method to detect blinks by analysing the motion vectors in the face and eye regions. First, the author applies a Viola-Jones detector to extract the eye area. Then the eye area was divided into 3×3 cells and generated 9 motion vectors. Of these 9 vectors, the above 6 specifies eyelid movements. From these vectors the variance related to eyelid movement is estimated and the obtained value compared with the selected threshold value. The author reported an accuracy of 91% on the ZJU dataset and 79% on the Eyeblink8 dataset. In a recent study, Fogelton and Benesova [15] used motion vectors in the eye region and blink sequence learning model using Recurrent Neural Network to detect blink levels and achieve similar performance compared to the above method.

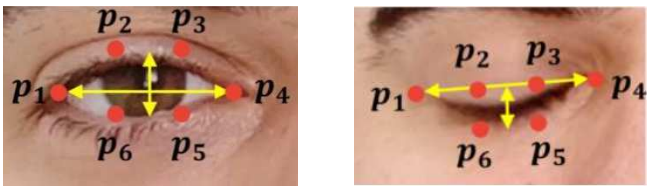
For the feature-based technique, Soukupova and colleagues [16] performed estimation of the open/closed state of the eye based on the ratio of the height and width of the eye. In the study [17], the curvature of the upper eyelid was used for this purpose. This approach is speed-optimized, however, this method is still limited because the decision of the state of the eye must be based on a fixed threshold, which will not be generalizable to all subjects due to the large geometry. The shape and characteristics of the eyes vary from person to person, resulting in a different ratio between the height and width of the eye in each person (see Figure 2).

Figure 2. Differences in eye shape and characteristics in each person

III. EYES OPEN/CLOSE STATUS DETECTION TECHNIQUE

A. Detect open/closed state based on varying aspect ratio

To calculate the aspect ratio of the eye, according to Figure 3, the human eye is represented by a set of 6 points, starting with the left corner of the eye and marked clockwise. Each point represents 1 coordinate in 2D space with $P_1(x_1, y_1)$ and corresponds to the remaining points from P_2 to P_6.



(a) Aspect ratio of eyes when open (b) Aspect ratio of eyes when closed

Figure 3. Coordinates and aspect ratio of the eye

The width of the eye is the distance between the points P_1 and P_4 as determined by the formula **Error! Reference source not found.** according to Euclid.

$$D(P_1, P_4) = \sqrt{\left(P_1.x - P_4.x\right)^2 + \left(P_1.y - P_4.y\right)^2} \qquad (1)$$

The eye height is the distance between the pairs of points (P_2, P_6) và (P_3, P_5) refer to the formula **Error! Reference source not found.**.

The distance from P_1 to P_4 is always constant, while the height will change depending on the opening or closing of the eyes. To detect a blink by calculating the ratio between the average height and the length of the eye. This ratio will remain almost constant when the eye is open, and it will rapidly decrease to zero when the eye is closed.

The formula for the aspect ratio of the eye (EAR – Eye Aspect Ratio) is determined as follows:

$$EAR = \frac{D(P_2,P_6)+D(P_3,P_5)}{2*D(P_1,P_4)} \qquad (2)$$

Where: P_i represents the corresponding coordinates on the eye

With the left eye, including the coordinates from P_{36} to P_{41}; the right eye covering the coordinates from P_{42} to P_{47} clockwise (see Figure 6).

Since the size of the eyes is not the same from one person to another, so a person with large eyes and a person with small eyes will not have the same height and coordinate values. This leads to inaccuracies in determining eye opening/closing behavior when comparing the EAR value with a fixed threshold (EAR is within a certain threshold, the eye will open or close). To improve this, we offer a solution to determine the dynamic eye opening/closing behavior based on the variation of the frame containing the eyes of each individual as follows:

The aspect ratio eyes in normal state:

$$EAR_{MAX} = \frac{D(P_2,P_6)_{max}+D(P_3,P_5)_{max}}{2*D(P_1,P_4)} \qquad (1)$$

The aspect ratio eyes on current state:

$$EAR_{CURRENT} = \frac{D(P_2,P_6)_{current}+D(P_3,P_5)_{current}}{2*D(P_1,P_4)} \qquad (2)$$

Assuming the eye-opening/closing behavior in humans is occurring simultaneously in both eyes, then the aspect ratio containing the eyes in the normal state and the current state is determined respectively as follows:

$$\overline{EAR}_{MAX} = \frac{EAR_{MAX\_L}+EAR_{MAX\_R}}{2} \qquad (5)$$

$$\overline{EAR}_{CURRENT} = \frac{EAR_{CURRENT\_L}+EAR_{CURRENT\_R}}{2} \qquad (6)$$

Where:

EAR_{MAX_L} is the normal aspect ratio of the left eye;
EAR_{MAX_R} is the normal aspect ratio of the right eye;
$EAR_{CURRENT\_L}$ is the current aspect ratio of the left eye;
$EAR_{CURRENT\_R}$ is the current aspect ratio of the right eye.

Finally, to decide whether the eyes are closed or open, we compare the $\overline{EAR}_{MAX}$ value with the $\overline{EAR}_{CURRENT}$ value as follows:

$$x = \overline{EAR}_{CURRENT} - ½ * \overline{EAR}_{MAX} = \begin{cases} \text{EYE CLOSE}, & x < 0 \\ \text{EYE OPEN}, & x \geq 0 \end{cases} \qquad (7)$$

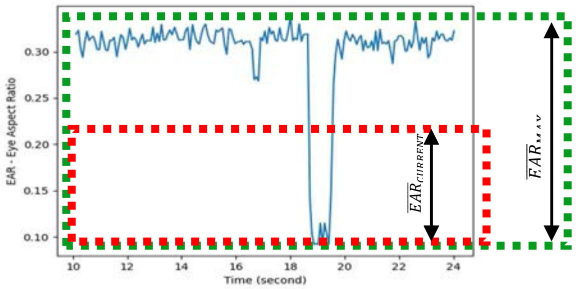


Figure 4. Variation of aspect ratio in eye opening/closing behavior

B. Algorithm to detect the state of opening and closing eyes Blink detection+

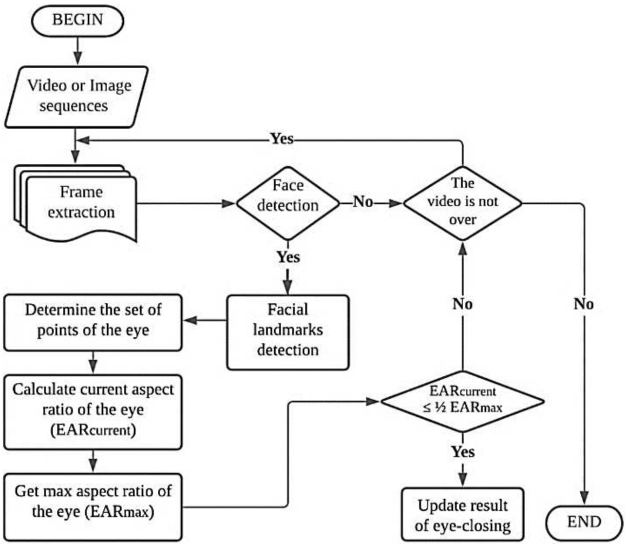


Figure 5. Flowchart to detect eye opening/closing behavior

C. Evaluate the complexity of the algorithm

Algorithm execution time is determined, including extracting frames from the video, determining the eye opening/closing behavior of the objects in that frame and the formulas for calculating the aspect ratio of the eyes are shown as follows:

Let: l_v is the length of the video

n_o is the number of objects in each frame

TABLE 1. COMPLEXITY OF ALGORITHM

No	Process	Execution time
1	Frame Extraction	$O(l_v)$
2	Face detection on each frame	$O(n_o)$
3	Face landmarks detection	O(1)
	Determine the set of points of the eye	O(1)
	Calculate EARcurrent	O(1)
	Get EARmax	O(1)
4	**Total execution time of algorithm**	$O(l_v \times n_o)$

IV. EYE BLINK DETECTION+ EVALUATION

We have installed Blink detection+ algorithm. To identify 68 facial landmarks in the video we used the Dlib library [18], with the iBUG 300-W training dataset (see Figure 6).

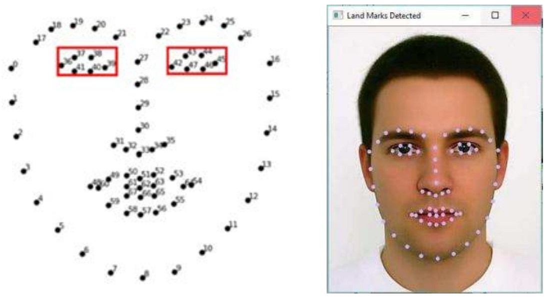


Figure 6. Set of 68 facial landmarks

A. Dataset for the experiment

TABLE 2. BLINK DATASET

Dataset	Length	Resolution	Object
Talking face	5,000 frames	720 × 576	1
Eyeblink8	70,992 frames	640 × 480	4
HUST-LEBW	8,749 frames	1280 × 720 1456 × 600	172

The Talking Face[1] dataset consists of 5,000 frames taken from a video of a person participating in a conversation. Frames were captured at 25 frames per second with a resolution of 720×576.

The Eyeblink8[2] dataset contains 8 videos with 4 subjects (1 person wearing glasses). The video was shot in a home environment. People are sitting in front of the camera and mostly acting spontaneously with live face simulations, similar to the Talking face dataset. There were a total of 408 eye blinks across 70,992 frames with a resolution of 640×480. Both datasets are labeled and annotated by the authors about eye opening and closing behavior at each frame.

Figure 7. Subjects in dataset Talking Face, Eyeblink8 and HUST_LEBW

The HUST_LEBW[3] is a multimodal database collected in the wild. It consists of 673 eye blink video clip samples corresponding to 172 subjects. Each clip sample comprises 13 frames and include a blink sample or un-blink sample. There were a total of 381 eye blink across 8,749 frames with both resolutions of 1280×720 and 1456×600. This dataset is rich in pose, illumination changes, and other naturally occurring factors, published in 2019 by [19] for the research purpose of eye blinks in nature.

[1] https://personalpages.manchester.ac.uk/staff/timothy.f.cootes/data/talking_face/talking_face.html
[2] https://www.blinkingmatters.com/research
[3] https://mega.nz/folder/JQRUTSgS#1uM9jh8Oulw-BGrZUfaCQQ

B. Results

In this section, we evaluate the Blink detection+ method compared with the method of Soukupova [16] based on three standard data sets, the results are presented in Table 2.

The performance of the technique is evaluated by the following formulas:

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - Score = \frac{2 * (Precision * Recall)}{Precision + Recall}$$

Where: TP (True Positives) is the case where the eye is detected correctly, FP (False Positives) is an undetectable case, FN (False Negatives) is the case where the eye is detected incorrectly.

TABLE 3. THE RESULTS OF RUNNING TESTS ON THREE SETS OF DATA

Dataset	SA	EC	TP	FP	FN	Pre	Recall	F1-Score
Talking face	1	61	58	3	4	**95.08**	**93.55**	**0.943**
Eyeblink 8	1	38	35	3	2	92.11	94.59	0.933
	2	88	86	2	2	97.73	97.73	0.977
	3	65	63	2	1	96.92	98.44	0.977
	4	31	30	1	2	96.77	93.75	0.952
	5	30	28	2	2	93.33	93.33	0.933
	6	41	39	2	1	95.12	97.50	0.963
	7	72	70	2	3	97.22	95.89	0.966
	8	43	40	3	2	93.02	95.24	0.941
						95.28	**95.81**	**0.955**
HUST_LEBW		381	305	76	212	**80.05**	**58.99**	**0.679**

Note:

- SA: Sample
- EC: Number of eyes closed
- Pre: Precision

TABLE 4. THE RESULTS COMPARE WITH THE PROPOSED SOLUTION

Solution	Dataset	Precision	Recall	F1-Score
[16]	Talking face	-	-	
Blink detection+	Talking face	95.08	93.55	0.943
[16]	Eyeblink8	94.30	96.02	0.952
Blink detection+	Eyeblink8	95.28	95.81	0.955
[19]	HUST_LEBW	82.95	49.27	0.618
Blink detection+	HUST_LEBW	80.05	58.99	0.679

V. CONCLUSIONS

The study of the opening and closing state of the eye is a topic that is always of interest to researchers because of its wide range of applications. In this paper, we propose a technique to detect the open/closed state of the eyes based on the rectangular aspect ratio containing the eyes. The dynamic technique determines the threshold when eye closure occurs in all subjects.

This is a prerequisite study of research-intensive applications such as: monitoring driver fatigue and sleepiness to prevent auto accidents, or assessing student concentration during class hours., which can help teachers make reasonable adjustments to increase their teaching effectiveness.

Techniques have been installed experimental and proved effective for the first object with a straight posture and gaze at the camera. In further studies, the behaviors related to the nodding and shaking of the head will be of interest to us to increase the accuracy in detecting the open/closed state of the eyes.

ACKNOWLEDGEMENTS

This research is funded by the project NVCC02.06/22-22 “The task of providing funding for scientific research activities to senior researchers in 2022, NVCC02.06/22-22”.

REFERENCES

[1] Qijie Zhao; Junye Jiang; Zhigao Lei; Jingang Yi, "Detection method of eyes opening and closing ratio for driver’s," *IET Intelligent Transport Systems*, pp. 31-42, 2020.

[2] Mohamed Hedi Baccour, Frauke Driewer, Enkelejda Kasneci and Wolfgang Rosenstiel, "Camera-Based Eye Blink Detection Algorithm for Assessing Driver Drowsiness," *IEEE Intelligent Vehicles Symposium*, 2019.

[3] Vidal, M., Turner, J., Bulling, A., and Gellersen, H, "Wearable eye tracking for mental health monitoring," *Comput. Commun*, vol. 35, pp. 1306–1311, 2012.

[4] Lu, Y., Zheng, W., Li, B., and Lu, B, "Combining eye movements and EEG to enhance emotion recognition," *in Proceeding of International Joint Conference on Artificial Intelligence (IJCAI) (Buenos Aires)*, 2015.

[5] Aleksandra Kro´lak, Paweł Strumiłło, "Eye-blink detection system for human–computer interaction," *Springer*, pp. 409-419, 2012.

[6] Marcos-Ramiro, A., Pizarro-Perez, D., Marron-Romera, M., Gatica-Perez., D., "Automatic Blinking Detection Towards Stress Discovery," *In: Proc. of the 16th Int. Conf. on Multimodal Interaction*, pp. 88-95, 2017.

[7] Prabin Sharma, Shubham Joshi, Subash Gautam, Sneha Maharjan, Vitor Filipe, Manuel Cabral Reis, "Student Engagement Detection Using Emotion Analysis, Eye Tracking and Head Movement with Machine Learning," 2019.

[8] Anna Rogalska, Filip Rynkiewicz, Marcin Daszuta, Krzysztof Guzek and Piotr Napieralski, "Blinking Extraction in Eye gaze System for Stereoscopy Movies," *Open Access*, 2019.

[9] Fa-Deng G, Min-Xian H, "Study on the detection of locomotive driver fatigue based on image," *In: 2nd International Conference on Computer Engineering and Technology*, pp. 612-615, 2010.

[10] S. Kawato and N. Tetsutani, "Detection and tracking of eyes for gaze-camera control," *Image and Vision Computing*, vol. 22, no. 12, pp. 1031–1038, 2004.

[11] Yunqi L, Meiling Y, Xiaobing S, Xiuxia L, Jiangfan O, "Recognition of eye states in real time video," *In: International Conference on Computer Engineering and Technology*, pp. 554-559, 2009.

[12] K. Grauman, M. Betke, J. Gips, and G. R. Bradski, "Communication via eye blinks: detection and duration analysis in real time," *in Proceedings of the 2001 IEEE Computer Vision and Pattern Recognition (CVPR)*, 2001.

[13] S. Kawato and J. Ohya, "Real-time detection of nodding and headshaking by directly detecting and tracking the “between-eyes”," *in Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France*, 2000.

[14] Mohanakrishnan, J., Nakashima, S., Odagiri, J., and Yu, S., "A novel blink detection system for user monitoring," *in IEEE/ACM International Conference on Utility and Cloud Computing (UCCV) Workshop (Tampa, FL)*, 2013.

[15] Fogelton, A., and Benesova, W., "Eye blink completeness detection.," *Comput. Vision Image Understand*, pp. 78-85, 2018.

[16] Soukupova, T. & Cech, J., "Real-Time Eye Blink Detection using Facial Landmarks," *21st Computer Vision Winter Workshop*, 2016.

[17] Guo J, Guo X, "Eye state recognition based on shape analysis and fuzzy logic," *In: Intelligent Vehicles Symposium*, pp. 78-82, 2009.

[18] D. King. (2018) Dlib C++ Library. [Online]. http://dlib.net/

[19] Yang Xiao, Zhiguo Cao, Lubin Meng, Zhiwen Fang, Joey Tianyi Zhou and Junsong Yuan Guilei Hu, "Towards Real-Time Eyeblink Detection in the Wild: Dataset, Theory and Practices," *IEEE*, pp. 2194 - 2208, 2019.

Tran Thanh Phuong received his Master of IT in 2008 at Lac Hong University.
Currently, he is a lecturer at Information Technology of Lac Hong University.
His research interests involve computer vision, image processing and machine learning.

Lam Thanh Hien received the MSc. Degree in Applied Informatics Technology in 2004 from INNOTECH Institute, France, and his PhD in 2017 at Vietnam Academy of Science and Technology.
Currently, he is a rector of Lac Hong University. His research interests involve computer vision, image processing and machine learning.

Do Nang Toan is an Associate Professor at the Institute of Information Technology - Vietnam Academy of Science and Technology. He received BSc. Degree in Applied Mathematics and Informatics in 1990 from Hanoi University, and PhD in Computer Science in 2001 from Vietnam Academy of Science and Technology. His main research interests are Pattern recognition, Image processing and Virtual reality.

Ngo Duc Vinh received his Master of IT in 2006 from the Military Technical Academy, and his PhD in 2016 at the Graduate University of Science and Technology, Vietnam Academy of Science and Technology.
Currently, he is a lecturer at Hanoi University of Industry.
His research interests involve computer vision, image processing and machine learning.