VOL. 11, NO. 12, JUNE 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

TRACKING THE EYES USING INTERDEPENDENCE MEAN SHIFT TRACKING ALGORITHM WITH APPROPRIATE INFORMATION PROVIDED

Masrullizam Mat Ibrahim[1], Syafeeza Ahmad Radzi[1] and John Soraghan[2]
[1]Department of Computer Engineering, Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[2]Centre for Excellence in Image and Signal Processing, Faculty of Engineering, University of Strathclyde, Royal College Building, George St, Glasgow, Lanarkshire, United Kingdom
E-Mail: masrullizam@utem.edu.my

ABSTRACT

The human eye tracking algorithm is very important, especially in the facial analysis application. Eyes represent rich of information that able to indicate the reaction, perception, or physiological status of individual. Most of the developed eyes tracking algorithm are not considered the condition of the eyes that would provide the appropriate information to be used in the processes of the facial analysis algorithm. This paper proposes a new technique in how the eyes are tracked and provide the appropriate information for further process. This technique uses Mean Shift tracking algorithm and interdependence scheme to track the eye and stop the tracking when eyes are out of best position and condition to deliver the appropriate information. This proposed algorithm is tested and evaluated using Strathclyde Facial Fatigue (SFF) video footage database, a collection of video footage from the internet and our own facial recording. Based on the experiment carried out, the proposed technique shows the promising result.

Keywords: eye tracking, mean shift and eye detection.

INTRODUCTION

Facial analysis is one of the branches in computer vision research which apply camera to monitor the activities of face. These activities are useful, such as in the facial expression algorithm, human face interaction application, and fatigue detection system. Eyes are a main component that rich with information that able to represent the human awareness level, interaction, and psychological aspect [1,2]. Therefore, to track the eyes simultaneously with extract the appropriate information is crucial in this algorithm.

The existing eye tracking algorithm can be grouped into three methods. Firstly, a shape based method which eyes are detected based on the interior and exterior shape characteristics [3,4]. Then, features shape based method which the set of distinctive features around the eyes such as the edges, the border, the corner and the dark region of the pupil are normally used [5,6]. Lastly, appearance shape based method in which the statistical analysis, probability theory, and the machine learning algorithm is used to train the eyes features in order to track the eye in sequences of video frames [7,8].

However, most the tracking algorithms have not considered the information of eye that crucially required for further process of eye activity analysis. There are situations which the eyes are not providing the appropriate information during tracking such as the situation eye in rigid motion, included, and in which iris area out of view. This proposed technique provides element to skip the tracking process when eye in these three situations. This propose algorithm allow the next stage process of the eye activity algorithm, obtaining the appropriate eye information. In this proposed technique, mean shift tracking algorithm combines with an interdependence scheme to track the eyes. In this technique both side eyes information is used in the tracking algorithm The remainder of this paper is organised as follows. Section II provides an overview of the complete IASMS based eye state analysis system. Section III explains the IASMS algorithm in detail. Section IV discusses the experimental results that have been obtained. The conclusion of the paper is provided in Section V.

INTERDEPENDENCE EYE TRACKING

The proposed interdependence eye tracking algorithm consists of three main processes as shown in Figure-1. This algorithm start with initial process which the eyes locations are searched in the input image. Then, the mean shift tracking algorithm is used to track the eyes, and the interdependence technique is applied to examine the distance of the eyes. If the distance is out of range, the tracking iteration is stopped.

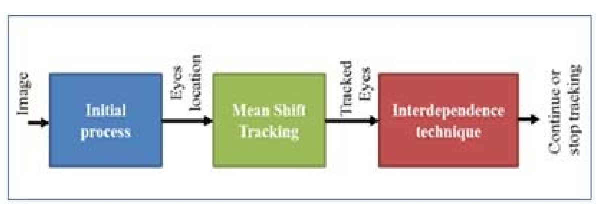

Figure-1. The process flow of interdependence tracking algorithm.

INITIAL PROCESS

The aim of the initial process is to find the location of the eyes and centre of the iris. In this research,

VOL. 11, NO. 12, JUNE 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

the location of the eyes are localized uses the technique that implemented in [9], in which the cascade classifier with Haar-like features [10] is used. The cascade classifier trained the pair of eyes features, and eyes are detected which pairs of eyes are localised together. Prior to find the centre of the iris, the openness of the eye is examined uses Adaptive Thresholding Histogram Enhancement (ATHE) technique to ensure iris can be detected accurately. Then, the prominent iris localization method Integro differential operator which was introduced by Daugman [11] is applied to detect the centre of the iris.

MEAN SHIFT TRACKING ALGORITHM

The Mean Shift algorithm, introduced by Comaniciu *et al* [12], is a popular method to track objects due to its simplicity and efficiency. The Mean Shift algorithm tracks the object based on the probability of colour of the object. In this research, the Mean Shift tracking algorithm is applied to track the movement wherein the iris is tracked based on the probability of the colour histogram of the iris $\hat{q} = \{\hat{q}_u\}_{u=1 \ldots m}$, where m is the number of histogram bins and is computed as follows:

$$\hat{q}_u = c \sum_{i=1}^{n} k\left(\|x_i\|^2\right)\delta\left[b(x_i) - u\right] \tag{1}$$

where x_i is a normalised pixel location from 1 to n with the target iris centred at 0. δ is a Kronecker delta function, while the function $b(x_i)$ determines the bin for pixel x_i. Mean shift tracking uses the Epanechnikov kernel [13] function k that represents the weight function of each pixel. C is a normalisation constant which is defined as follows:

$$C = \frac{1}{\sum_{i=1}^{n} k\left(\|x_i\|^2\right)} \tag{2}$$

In order to track the iris the probability of colour histogram of the target iris candidate model is computed as follows:

$$\hat{p}_u(y) = C_h \sum_{i=1}^{n_k} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)\delta\left[b(x_i) - u\right] \tag{3}$$

where y is the centre position of the current frame and h is the radius of weighting kernel. The normalisation constant for this target model C_h is computed as follows:

$$C_h = \frac{1}{\sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)} \tag{4}$$

The primary task of the Mean Shift tracking algorithm is to find the target location in the current frame. The most probable location of the centre of the iris is obtained by minimizing the distance between two distributions that can be defined as:

$$d(y) = \sqrt{1 - \hat{p}(y)} \tag{5}$$

where the $\hat{p}(y)$ is an estimation of the Bhattacharyya coefficient which can be computed as:

$$\hat{p}(y) = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)\hat{q}_u} \tag{6}$$

In order to find the acceptable centre point of the targeted iris in the next frame, the value $d(y)$ must be minimized while $\hat{p}(y)$ must be maximized. Comaniciu *et al* [12] derived the following method that can be used to compute a new location targeted iris as follows:

$$y_1 = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \tag{7}$$

where the weight w_i is computed as:

$$w_i = \sum_{u=1}^{m} \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}}\delta\left[b(x_i) - u\right] \tag{8}$$

INTERDEPENDENCE TRACKING SCHEME

The common mean shift tracking is based on single tracking, which is independent of other objects. However, in this algorithm, there are pair of eyes and, therefore interrelated. By utilizing information from the pair of eyes, we will incorporate an interdependence technique that aims to be able to track the non-rigid movement of the eyes. There are two primary pieces of information employed in this algorithm, the centre locations of the irises and the distance between them. Initially the eye centres are obtained from the initial processing algorithms. Then the Mean shift algorithm updates these centre positions for every tracking iteration. The C_r and C_l in Equation **Error! Reference source not found.** denote the centre points of the right and left eye respectively.

$$C_r = (x_r, y_r) \qquad C_l = (x_l, y_l) \tag{9}$$

From the obtained centre points of the irises, the distance D_e between them is continually updated during every tracking iteration, and is calculated as follows:

$$D_e(p, q) = \sqrt{(x_r - x_l)^2 + (y_r - y_l)^2} \tag{10}$$

The common drawback of the Mean Shift tracking algorithm is erroneous colour tracking. When dealing with colour, there is the possibility of multiple objects having similar colours. In this case, in order to avoid any erroneous results the following new approach is introduced. First a specific eye region is defined on the

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

image known as the Focus of Eye Region (FER) with the size and location of the region updated in every iteration. Then the Mean Shift algorithm is used to track the iris in the FER region of the image. The size of the FER depends on the centres of the irises C_r and C_l and the distance between them D_e in the following way:

$$I_{FER}(x_R, y_R) = I(x_r, y_r) \qquad x_{Cr,l} - \frac{1}{3}D_e \le x_{r,l} \le x_{Cr,l} + \frac{1}{3}D_e \tag{11}$$

$$I_{FER}(x_L, y_L) = I(x_l, y_l) \qquad y_{C,l} - \frac{1}{3}D_e \le y_{r,l} \le y_{Cr,l} + \frac{1}{3}D_e$$

Occasionally the target object can be occluded; in this case the eye is usually occluded by the hand. To enable the system to evaluate when occlusion or wrong tracking occurs, the Sum of Absolute Difference (SAD) value of FER in between two frames is obtained as indicated next:

$$I_{SAD} = \sum_{i,j} \left| I_1(i,j) - I_2(x+i, y+j) \right| \tag{12}$$

The input image is converted to gray scale to obtain the normalised values. The SAD value is normalized and computed as follows:

$$I_{nSAD} = \frac{I_{SAD}}{255 \times W \times H} \tag{13}$$

where W *and* H denote the width and height of FER respectively. The value of I_{nSAD} is examined in order to ensure that the eyes are always tracked accurately

EXPERIMENTAL RESULTS

The proposed tracking algorithm is intended to track the non-rigid eye movements and skip the tracking operation when the eye is out of region interest or occluded. Figure-2 shows a typical sequence of frames from a normal eye opening until the eye is occluded. As indicated in Figure-3 the differences of SAD in FER of the eye opened and the eye closed does not indicate significant changes from frame 1 to frame 60 (Figure-2). However, when the eye is totally occluded in frame 68 the normalized SAD value for the FER frame 68 increases rapidly. This significant change is used as an indicator in order to stop the tracking.

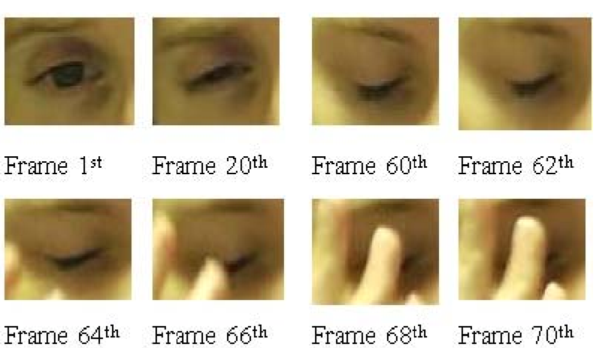

Figure-2. Sequences of frames from eye normally open to eye occluded by hand.

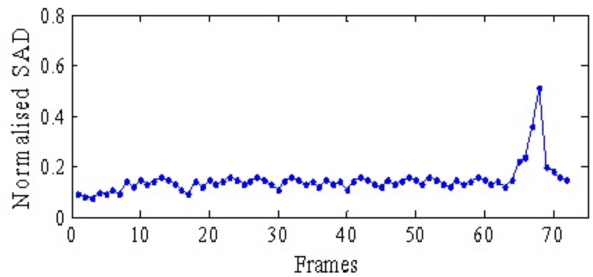

Figure-3. Normalised SAD values for a sequence of frames.

As discussed above, the eye tracking algorithm can stop the iteration when the FER is occluded or when the FER is out of the tracked region. Similarly, the FER is also sensitive to rigid eye movement which may also cause the algorithm to stop the tracking iteration since, in such cases, facial information may suddenly move outside the camera view field. Figure-4(a) shows the tracked eyes images which indicate the algorithm's ability to track the eyes in multi-pose of face as long as the FER region is not occluded or out of the search region. In Figure-4(b) the FER is out of the search region when the face is turned too much to the right or left (first and last views). This eye tracking algorithm emphasizes that both the eyes must be seen clearly in order to read the iris information, otherwise the algorithm stops and the initialisation scheme needs to be re-initiated to find out the new location of the eyes.

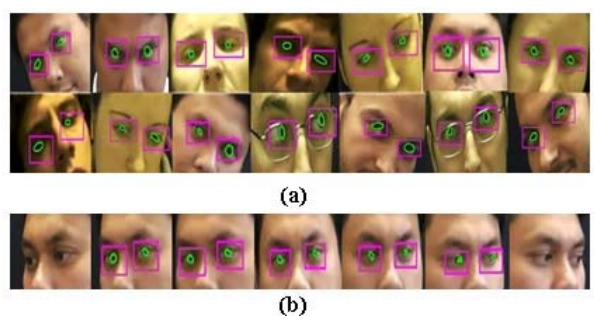

Figure-4. (a) Eyes tracking for multi-poses; (b) Eye tracking stops when FER is out of the search region.

In tracking the eyes, very important to keep the eyes information is whether the eye close or open explicitly. The proposed algorithm has been compared with others algorithms to evaluate how the algorithm able to keep all appropriate information that required. In this experiment, there are three situations of eyes occluded are considered, namely, eye occluded by hand, head turns right of left, and head looks up and down. There are randomly selected video footage from SFF, our own recorded and from internet are used for this experiment. The total number of frames that have been used are 12,050 frames for all situations. Table-1 show the comparison results between the proposed algorithms with another two

ARPN Journal of Engineering and Applied Sciences

algorithms. From the table, the proposed algorithm have achieved the best detection result for detection the appropriate eye information during the three eye situations compared with another two algorithms.

Table-1. Experimental results based on three situations of eyes out of location.

Eye occluded	Proposed Algorithm	Algorithm [2]	Algorithm [3]
Head turn right & left	98.4%	97.2%	98%
Head look up & down	99.2%	96.4%	98.2%
Occluded by hand	99.2%	95.8%	96.6%

CONCLUSIONS

A new algorithm to track eyes was presented. This algorithm is a combination of Mean Shift tracking algorithm and interdependence tracking scheme. In this algorithm eyes are tracked using irises location. The distance between irises are applied in interdependence tracking scheme in order to examine the eyes region position. There are a specific region which called Focus Eye Region (FER) are formed to enhance the performance of Mean Shift tracking algorithm in tracking the region of iris. The eye is detected occluded and out of tracking interest when Sum of Absolute Difference (SAD) of FER is changes rapidly. Based on the experimental conducted using several video footage from Strathclyde Facial Fatigue (SFF) video footage database [14], a collection of video footage from internet and our own facial recording, the proposed algorithm shows the promising results.

ACKNOWLEDGEMENTS

This work was supported in part by UTeM under Grant PJP_2013_FKEKK(47B)_S01274.

REFERENCES

[1] W. W. Wierwille, S. S. Wreggit, and R. R. Knipling. 1994. Developement of Improved Algorithm For On-line Detection of Driver Drowsiness. International Congress on Transportation Electronics, pp. p. 331-340.

[2] J. A. Stern, D. Boyer, and D. Schroeder.M 1994. Blink Rate: A Possible Measure of Fatigue. Human Factors. The Journal of the Human Factors and Ergonomics vol. 36, pp. 285-297.

[3] M. Chau and M. Betke. 2005. Real time eye tracking and blink detection with USB cameras. Boston University, Boston, USA., Technical Report 2005-012.

[4] I. Bacivarov, M. Ionita, and P. Corcoran. 2008. Statistical models of appearance for eye tracking and eye-blink detection and measurement. Consumer Electronics, IEEE Transactions on, vol. 54, pp. 1312-1320.

[5] F. Sukno, S.-K. Pavani, C. Butakoff, and A. Frangi. 2009. Automatic Assessment of Eye Blinking Patterns through Statistical Shape Model Computer Vision Systems. Vol. 5815, M. Fritz, B. Schiele, and J. Piater, Eds., ed: Springer Berlin / Heidelberg, pp. 33-42.

[6] T. Morris, P. Blenkhorn, and F. Zaidi. 2002. Blink detection for real-time eye tracking. J. Netw. Comput. Appl., vol. 25, pp. 129-143.

[7] P. Gang, S. Lin, W. Zhaohui, and L. Shihong. 2007. Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Webcamera. Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on, pp. 1-8.

[8] Z. Sanyuan and S. Tingzhi. 2010. Driver Fatigue Detection Based on Eye Status. Multimedia Technology (ICMT), 2010 International Conference on, pp. 1-4.

[9] M. M. Ibrahim, A.A.M. Isa, and A. M. 2014. Eye Closure and Open Detection Using Adaptive Thresholding Histogram Enhancement (ATHE) Technique and Connected Components Utilisation. Intelligent and Advanced Systems (ICIAS), pp.1,5.

[10] P. Viola and M. J. JONES. 2004. Robust Real-Time Face Detection. International Journal of Computer Vision. Vol. 57(2), pp. 137–154.

[11] J. Daugman. 2004. How iris recognition works. Circuits and Systems for Video Technology, IEEE Transactions . Vol. 14, pp. 21-30.

[12] D. Comaniciu, V. Ramesh, and P. Meer. 2003. Kernel-based object tracking. Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 25, pp. 564-577.

[13] B. W. Silverman. 1986. Density estimation for statistics and data analysis vol. 26. Chapman & Hall/CRC.

[14] M. M. Ibrahim, J. J. Soraghan, and L. Petropoulakis. 2012. Non-rigid eye movement tracking and eye state quantification. Systems, Signals and Image Processing (IWSSIP), 2012 19th International Conference, pp. 280-283.