IET Intelligent Transport Systems

Research Article

Journals
The Institution of Engineering and Technology

Smartphone-based fatigue detection system using progressive locating method

ISSN 1751-956X
Received on 17th April 2015
Revised on 24th July 2015
Accepted on 18th September 2015
doi: 10.1049/iet-its.2015.0076
www.ietdl.org

Zibo Li[1], Guangmin Sun[1,2] ✉, Fan Zhang[1], Linan Jia[1], Kun Zheng[1], Dequn Zhao[1]

[1]Department of Electronic Engineering, Beijing University of Technology, Beijing 100124, People's Republic of China
[2]Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing University of Technology, Beijing 100124, People's Republic of China
✉ *E-mail: gmsun@bjut.edu.cn*

Abstract: Smartphone applications became very popular nowadays as they provide useful functionalities to our daily lives over ordinary voice services. They offer a small but powerful computing platform where intelligent algorithms can be coded into some live-saving products which profoundly impact our daily lifestyles. A mobile fatigue detection system which is proposed in this study is an important life-saving application running on smartphone. Mobile detection system faces technological challenges such as: (i) embracing relatively low-resolution images in image recognition; (ii) supporting fast response time given a low-power CPU in comparison to a desktop computer; and (iii) demanding for high prediction accuracy by light-weight machine learning algorithms, as the software programs embedded in smartphones is resource constrained. Solutions with respect to indoor facial profiling system which are mainly based on progressive locating method for eye detection are discussed in this study. Acceptable experimental results in terms of eye detection rate and driver fatigue detection in different situations are presented.

1 Introduction

In the past two decades, car-ownership rate escalated almost exponentially in urban cities; driving a vehicle has become a part of our daily life. With the increasing number of vehicles, the rate of road accidents reached an alarming level as one of the major societal problems. According to the data of Sleep Research Center in UK, the 20% of road accidents are happened due to driver fatigue [1]. In perspective of sociologist [2, 3], the cause of driver fatigue is the combination of endogenous and exogenous factors. On one hand, drivers who spend long hours on driving are subjected to the 24 h physiological circadian rhythm; on the other hand, the monotony road environment makes the contribution to the decrease of arousal and drivers' attention. Therefore, drivers' fatigue is a major cause of incidence of traffic accidents.

Driver fatigue can be detected in different ways. Based on the exogenous factors, some studies focused on the environment of highway. Thiffault, P *et al.* [2] discussed the factors that may lead to driver fatigue and pointed out that highway hypnosis plays an important role in the cause of driver fatigue. Natasha Merat and Hamish Jamson [4] designed three engineering treatments (chevron, rumble strips, and variable message signs (VMS) for keeping the attention of driver. However, the analysis on highway lies mainly on the assistance of government and the application of the idea will be expensive. Compared with these analyses, monitoring the conscious state of car driver dynamically using computer program could be a feasible approach to avoid traffic accident by alerting the driver when he/she starts to feel sleepy. Conventionally, driver fatigue detection system is designed by processing different types of signal feeds: driving performance parameter, and driver's physiological signal, and visual cues.

Driving performance parameter includes the vehicle speed, lateral deviation, and steering wheel movements [5, 6]. Various systems [7, 8] have been designed with these parameters; however, the performance of these methods may be linked to some factors, such as different road conditions or the driver experience [9].

Apart from the driving performance parameter, physiological signals that involve electroencephalogram (EEG) [10, 11] and electrocardiogram [12] also show a good accuracy when detecting driver fatigue. However, the limitation of this measurement is that the data is obtained by an annoying and invasive way.

Compared with the other types of signals, acquiring the drivers' mental state in a non-invasive and direct way can be treated as an advantage of visual cues [1, 9, 10, 13–19]. Monitoring driver fatigue by yawn, eye state, and tilted head is used as a common way. Lamiaa Fattouh Ibrahim *et al.* [9] employed template matching method and edge extraction to detect eye and its state, respectively. Chiliang Xiong *et al.* [18] adopted percentage of eyes closed (PERCLOS) measurement to detect eye state. Du Yong *et al.* [19] depicted the yawn by combining the feature of statistic and texture with kernelised-fuzzy set. Then yawn state was detected by support vector machine (SVM). Bogusław Cyganek and Gruszczyński [15] utilised the integrated infrared (IR) image, and made use of the shape of eye to detect the eye region. The eye state was subsequently classified through the higher order singular value decomposition (HOSVD) algorithm. However, these systems were designed by sole factor. To enhance the accuracy of driver fatigue detection, other systems were built by analysing the combination of different types of signals. For example, eye state with yawn [14, 17, 37–39]; eye state with several driving performance parameters [1, 16]; titled head detection with eye detection [13]; and titled head detection with EEG [10]. Our method is also included in this category, but applied in different platform.

According to the relevant article, desktop PC is always being used as essential monitoring equipment because the algorithm often relies on its high-speed and sophisticated computing environment. However, smartphone is a handier in size than a PC if the monitoring would be carried out in the cabin of a car. In addition to compact size, smartphones have other advantages over a PC such as simple user-interface, portability, and easy to configure as an App. Moreover, smartphone as a small accessory incurs low operation cost and it should interfere driving as little as possible. On the other hand, driver fatigue detection system based on smartphone also suffers from some technical shortcomings. First, low-quality images which are acquired through camera of smartphone may lack of high image resolution that may be necessary for accurate prediction. Second, low-speed processor on the phone generally affects the operational performance of the system. To compromise with these limitations, several systems [20, 21] have been built on the basis of smartphone. However, these methods may not perform well under different conditions like changes in illumination or glasses. Our proposed system is designed by a series of basic algorithms with the aim of reducing the system workloads and deal with different conditions. The driver fatigue detection system based on smartphone is made up of modules of face recognition, eye detection, and fatigue state classification. Among these three parts, eye detection offers the key to address the whole problem.

Recalling different strategies of eye detection, there are four instinct categories that have been actively discussed by researchers [22]: near IR illumination-based methods, characteristic-based methods, statistical-based approaches, and structure-based methods. Wei [23] designed external digital signal processing (DSP)-controlled IR source, and detect the eye by exploiting IR illumination method to keep their system being low-computational costs. Receiving a lot of research interests in recent years, statistical-based methods have been widely analysed [22, 24–26] as reported in the literature. Among these methods, Yan Ren *et al.* [22] proposed a two-class sparse representation classifier (SRC) in place of SRC by using feature of scale invariant feature transform (SIFT) to keep scale invariant and rotation invariant. Stan Z. Li *et al.* [25] discussed a multi-Adaboost algorithm in different conditions of with/without eyeglasses, left/right eyes to achieve a high recognised rate. Everingham and Zisserman [26] introduced Bayesian method to localise eye position. Regardless of hardware or machine learning, structure-based methods study the geometric relationship among features extracted on eyes and the spatial association between eyes and other face features, like active shape model [27] and isophote-based method [28]. Without the influence of hardware or machine learning, characteristic-based methods are on the basis of simple structure and basic algorithm [29–33]. As common way, grey-level projection is usually adopted as an eye detector. The authors in [29, 30] used this detector after face recognition separately. Except for projection methods, circular Hough transform (CHT) is also employed for eyes detection. The authors in [30–32] employed CHT to analyse the edge information of face image. As another contribution, Alan *et al.* [33] proposed deformable template method.

Two contributions of this paper are described in the following:

(i) Designing a driver fatigue detection system in smartphone.
(ii) Proposing a novel eye detection method named progressive locating method (PLM).

The main contribution of the paper is proposing a successful driver fatigue detection application in smartphone. Taking consideration of the background, the advantage of our method is saving the memory of smartphone and keeping the robustness of system at the same time. Especially, our method aims at analysing facial profile rather than frontal portrait for avoiding driver from sight interfering.

In this paper, PLM that is treated as the secondary contribution is designed as a novel method. The aim of designing PLM is getting rid of the restriction of statistic method and trying to focus on apparent and geometric relationship within facial region. To accomplish these goals, an idea of progressive location is employed in this paper since the eye region might not be simply discriminated within the whole face. The idea of progressive location indicates that the procedure of detecting eye is divided into several parts. In each part, different and simple method is used for reducing the analysing region. Therefore, a more accurate location of eye can be obtained than the previous step. Based on facial image, our PLM algorithm is made up of three parts. First, a mouth detector is adopted to reduce the analysing field from whole profile to coarse eye region. Second, grey-level projection is used to secondarily discriminate the right eye region after the mouth detection. Third, an accurate eye coordinate is measured by connected analysis. The main idea of PLM is reducing the uncertainty of eye detection part, which is brought by the unrelated facial region and avoiding the high-computation cost in case of adopting several traditional eye detection algorithms. Consequently, driver's right eye is measured within a single object field by PLM in the procedure of constantly reducing the analysing region.

The rest of the paper is organised as follows. Section 2 introduces the proposed system. Section 3 illustrates details about our experiments and presents some discussions about result of these experiments. Section 4 concludes the paper.

2 Methodology

By incorporating PLM algorithm, our driver fatigue detection system integrates three major parts (see Fig. 1). First, skin-colour model is employed to detect the field of the driver's face. Second, PLM enhances the accuracy of eye detection rate by constantly eliminating various interferences via image segmentation. The processing flow of PLM algorithm is emphasised according to the first box. In the facial region, the dotted rectangle represents coarsely the specific region, where the eye is located. After the field of result becomes more precise than whole facial region, the secondary step eliminates some useless content and focus on the eye position. The dual steps in the first box are affected by the concept of 'progressively locating'. The concept of 'progressively

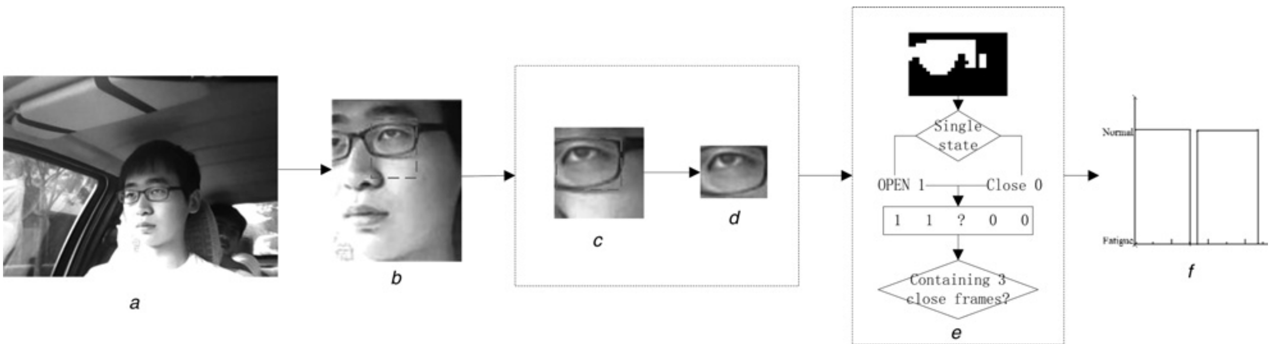

Fig. 1 *System is divided into three parts: face recognition, PLM algorithm, and state classification. Our system applies to human fatigue detection*

a Input image captured by smartphone contains vehicle background and test subjects
b Cropped face is obtained by a face detector
c Locating result, corresponding to dotted rectangle in Fig. 1*b*
d Reducing region, corresponding to dotted rectangle in Fig. 1*c*
e Image segmentation result of eye region
f System output diagram in time series

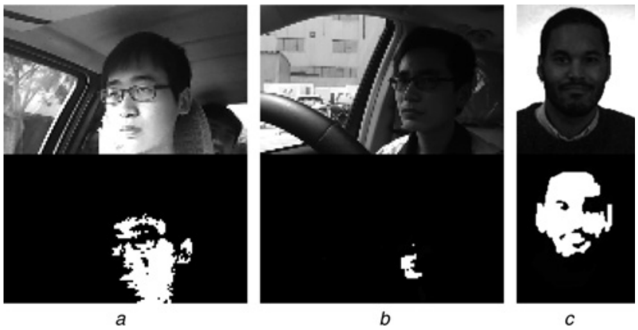

Fig. 2 *Examples of skin-colour model under different conditions*

a Normal condition
b Low illumination condition
c Dark coloured skin subject

locating' means that an eye region is progressively tracked from a complex and full facial image to a simple region. Third, the driver's wakefulness state which is in opposite of the fatigue state is measured by counting over the different combinations of states, which is represented by the second box. On the basis of eye detection result, height and dimension are achieved from the binary image (Fig. 1*e*). Then a simple classifier is adopted in order to get the single frame state. As the system's output, the driver's state is measured by counting a five-element queue that includes five single frame states in adjoining frame. As a warning to the driver, 'fatigue' state is sent by the system when it contains at least three close elements. In the following subsections, each part will be introduced with enough intermediate results.

2.1 Face recognition

As aforementioned above, face recognition is a preprocessing step in our driver fatigue detection system. Skin-colour model concentrates on analysing the information of a colour vector (i.e. YC_bC_r) instead of a single grey value. Unlike the complex structure in Adaboost algorithm, skin-colour model method detects face using simple calculation. Although the performance of skin-colour model is lower than Adaboost algorithm, our system can detect face without any external data (i.e. cascade classifier) due to limitation of smartphone. Therefore, the statistic method is kept away. Based on the colour sequences captured by camera of smartphone, skin-colour model is adopted in our system. This model utilises single Gaussian model (1) to determine the similarity between the current colour and skin-colour model in YC_bC_r domain (2)

$$N(x) = \exp\left[-0.5 * (\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})\right] * \det(\boldsymbol{\Sigma}) \tag{1}$$

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.500 & -0.419 & -0.081 \\ -0.169 & -0.031 & -0.500 \end{bmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{2}$$

where $\boldsymbol{x}$ is input vector with $\{C_b, C_r\}$, $\boldsymbol{\mu}$ is average of input vector, and $\boldsymbol{\Sigma}$ represents covariance matrix. In terms of optimisation, expectation maximum is adopted to drive these parameters. Some results with parameters $\boldsymbol{\mu}/\boldsymbol{\Sigma}$ are shown by

$$\begin{gathered} \boldsymbol{\mu} = \begin{bmatrix} 103.0056 & 140.1309 \end{bmatrix} \\ \boldsymbol{\Sigma} = \begin{bmatrix} 160.1301 & 12.1430 \\ 12.1430 & 299.4574 \end{bmatrix} \end{gathered} \tag{3}$$

To propose the performance of skin-colour model, three subjects are shown in Fig. 2. The changes in illumination and races conditions are involved within these three subjects. Besides, due to the deficiency of African subjects, the third subject in Fig. 2 is an example of the colour FERET database [34, 35]. As Fig. 2 shows, binary image indicates the potential facial region clearly. Compared with the original image, binary image fully emphasises the face region, removes the surrounding objects in Fig. 2*a*, and makes contribution to measuring the position of face. Among three examples, the lowest performance is obtained under low illumination because the colour distribution around the facial region is converted as the distribution of shadow. Therefore, the low illumination could be treated as the limitation of the skin-colour model. Moreover, the dark coloured skin subject is considered as a condition under which skin-coloured method may show a low performance. However, as shown in Fig. 3, the skin-coloured method adopted in this paper seems to achieve a well performance. To demonstrate this phenomenon, we randomly selected a subset of the colour FERET database where some dark coloured skin subjects were included. As a result, 96.3% of all faces in this subset can be well detected. Therefore, the variation of the coloured skin could not cause the decrease of face recognition rate.

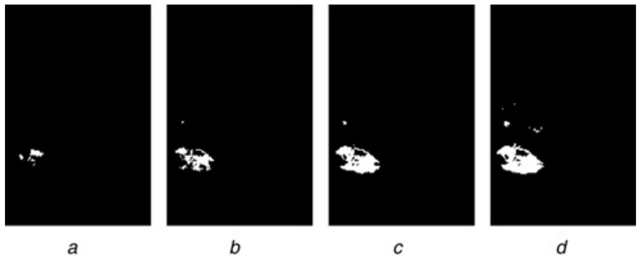

Fig. 3 *Mouth label results under different thresholds*

a Mouth label results by controlling the threshold from the 85% of maximum
b Mouth label results by controlling the threshold from the 80% of maximum
c Mouth label results by controlling the threshold from the 70% of maximum
d Mouth label results by controlling the threshold from the 60% of maximum

2.2 Progressive locating method

Apart from the other eye detection methods, the idea of PLM determines that detection work cannot be carried out in the field of whole face. Consequently, in the first part of PLM, finding a facial object as a reference is focused to get the coarse region where the right eye locates. Containing the most vivid colour within a facial region, mouth is used as a reference object in PLM. To depict the colour distribution of mouth, the *YIQ* colour domain, especially with *Q*-component, is adopted to detect mouth. As *Q*-component can describe the independent saturation range from yellow, green to purple, a better performance is shown by *Q*-component than *R*-component in *RGB* domain. The transformation from *RGB* to *YIQ* colour domain is shown as

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4}$$

Therefore, a threshold of *Q*-component is employed to discriminate potential mouth pixels in a facial image. As the experimental result

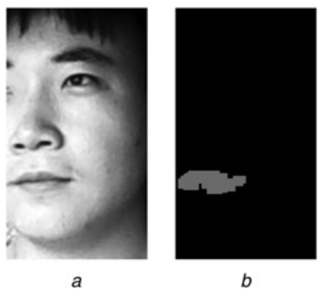

Fig. 4 *Mouth detection result*

a Face images in different subjects
b Grey region indicates mouth location results

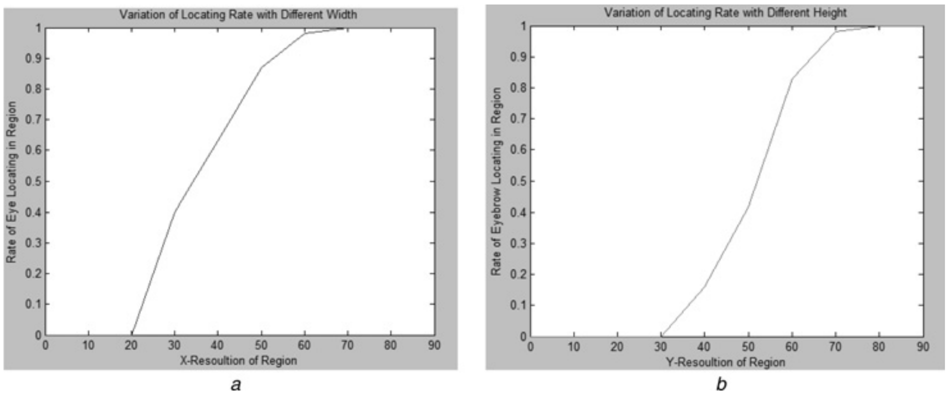

Fig. 5 *Variation curve of locating rate with different height and width*

a Occurrence rate of eye and eyebrow in different *X*-resolution
b Occurrence rate of eye and eyebrow in different *Y*-resolution

(Fig. 3), a pixel is classified as a potential mouth pixel when its *Q*-component value is larger than 70% of the maximum. Among four result images, too much useless information is kept, so Figs. 3*a* and *b* just show a part of the mouth. Compared with 70%, thresholds which are both 60% and lower than 60% of maximum are prone to cause an over-labelled result during mouth detection. Therefore, choosing 70% of maximum as the threshold parameter is sensible in mouth detection.

Fig. 4 shows the mouth detection results. Through analysing *Q*-component and morphological operation, grey region image in Fig. 4*b* are obtained, indicating the position of mouth. Based on the coordinate information of mouth region, the bottom and right points are chosen as the reference points to get the coarse region. Without taking specific detection into consideration, the result after first location just offers a smaller region than the whole facial region. More specifically, the bottom boundary of this coarse region is measured by adding 20 rows of coordinate on the bottom reference points and the left boundary of coarse region is measured by the right reference directly. As shown in Fig. 4*c*, a coarse eye region is located among various subjects. Although each person's eye locates in different position of box, the following detection is affected rarely by this influence unless eye disappears. Owing to the low accuracy generated in the first eye location work, the resolution of coarse region is limited to 60×60. The reason why we choose 60 as the length and height is shown in Fig. 5. In Fig. 5*a*, the occurrence rate of eye reaches 98.56% at 60 *X*-resolution. Although there are higher values within 70–90 *X*-resolution, we choose 60 as the width for the sake of saving computation cost. Likewise, 60 is a proper value for *Y*-resolution by observing Fig. 5*b*.

In terms of the secondary eye location part, our main purpose is to modify the upper and lower boundary information from the coarse coordinate result. Making full use of the projection method, binary information is emphasised in the second part of PLM. After image segmentation with the facial part, horizontal projection is executed within coarse eye region to detect the upper and lower boundary of secondary eye region. For example, Fig. 6 shows condition of wearing glasses. As the result of image segmentation, the shape of eyebrow, glasses frame, and eye are shown in 6*a*. Based on the binary content in the coarse box, the distance between each white region in vertical direction can be observed. Therefore, it is possible to further locate the eye region by discriminating the change rule in horizontal projection. To demonstrate this idea, a projection curve is obtained within the line box (see in Fig. 6*a*) to show the complexity of the analysing region. According to the curve plotted in Fig. 6*b*, several peaks (i.e. up arrows) can be observed in the first-two-thirds row, due to different white regions like eyebrow or glasses frame. At the same time, there is a sharp decrease (i.e. down arrows) in later-one-third half. As a result from the observation, when there are several peaks in the first-two-thirds curve, the last peak always means the centre of eye. On the other hand, the last sharp decrease in the second-one-third curve means the lower boundary of glasses frame or eye. To obtain a reduced region, we choose the position of last sharp decrease as the bottom coordinate and the peak point which contains the nearest distance to bottom coordinate as the top coordinate, as shown in Algorithm I (see Fig. 7). In terms of algorithm, parameter T is measured by the width of eye. Based on paper's background, the variation of T can be neglected. As Fig. 6*b* plots, the distance between the

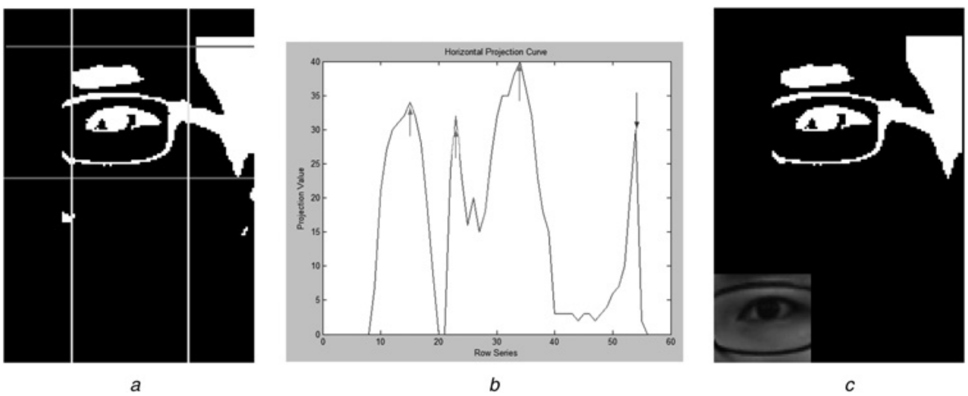

Fig. 6 *Secondary eye detection diagrams*

a Binary image with boundary line
b Horizontal projection curve, containing peak-pointed arrows
c Secondary eye location results. Left-bottom box shows acquired eye region, which is more accurate than the big field in first eye location

Algorithm 1

Input: a binary matrix $A \in B^{H \times W}$ of region contains the eyebrow, eye, or glasses frame.
Output: Accurate Eye Region Vertical Coordinate Information, includes upper boundary Y_u and lower boundary Y_l.

$k = 0$.
FOR $i = H, H-1, \cdots 1$
  $I_i = \sum_{j=1}^{W} A_{i,}$
  If $I_i > W/3$
    FOR $t = i, i-1, \cdots 1$
      $S_k = \arg\max_{1 <= t <= i} \sum_{j=1}^{W} A_{tj}$
      If $I_t <= W/3$
        $k = k+1$
        $i = t$
        BREAK FOR
    END FOR
END FOR
$Y_l = S_{k-1} - 5$
If $S_{k-1} - S_{k-2} > T$
  $Y_u = S_{k-2}$
Else
  $Y_u = S_{k-3}$
END

Fig. 7 *Eye region vertical localisation*

second peak and the third peak which represents the position of upper glasses frame and eye, respectively, is 10. Therefore, 15 is chosen as the value of this parameter. After renewing the height information, the left boundary is modified by sweeping the first white pixel. On the other hand, the right boundary just deducts value to update the width to 50, which is the largest width to store the image of a complete eye. Compared with Fig. 4*c*, a more specific result than coarse location box is shown in Fig. 6*c*.

In the past two steps, the idea of progressive location is implemented by the scheme of PLM. Based on the facial region, the proceeding of our algorithm is divided into coarse location part and specific location part. In terms of the coarse location part, the motivation of this part is converting the focus from the facial region to the region around eye. What we concern in this part is not the accurate position of eye but the reduction of analysing region. Therefore, this idea determines that our algorithm tries to find approximate field where the eye region is located. In this way, the algorithm gets rid of burdening high computation costs. Moreover, more convenience in specific location is offered by this scheme. As far as the aim of secondary part is concerned, the centre of specific region is kept in the accordance with the centre of eye. To achieve a high accuracy result, the algorithm of second part shows a good property in the situation of eliminating redundant information. By the horizontal projection method, the centre of eye can be distinguished among several similar objects (Fig. 6*b*). In conclusion, to cope with the limitation of smartphone, PLM is designed by dual-step scheme (i.e. the two boxes shown in Fig. 1). In terms of our detection algorithm, each part contains a simple and efficient method for eye detection.

2.3 Driver fatigue detection

As the system's output, driver's fatigue detection relies mostly on the detected object and the proper feature. According to previous researches, the eye state is assumed to be a good indication of drowsiness level characterised by micro-sleep. During a short moment, the driver rapidly closes the eyes and dozes off intermittently. As a usual standard, many researchers use the PERCLOS as a drowsiness indicator [36]. In this paper, the binary information with the extracted right eye is discussed (Fig. 8), and height and dimension are calculated as features to characterise the drowsiness level. After the standards is trained through utilising the most influential features within the first 30 frames, the current driver's fatigue is detected when the height is lower than half of the height standard and the dimension is lower than 60% of the dimension standard. To avoid some miss-detection, a data queue is set to orderly store five adjoining states from corresponding

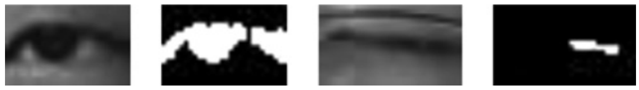

Fig. 8 *Eye state detection results. Each pair contains the original image and image segmentation result. By observing the binary images, height and dimension can play a useful role in status discrimination*

frames. With counting the different driver status within five frames, the output of system follows the constraint (5). Then an alarm is issued to avert the drowsy driver when the output of system is zero

$$\text{state} = \begin{cases} 0 & \text{if there are at least three close frames in five frames} \\ 1 & \text{otherwise} \end{cases} \tag{5}$$

3 Experiment

3.1 Introduction of experimental environment

In this section, we present an experiment on a database of self-made video sequences in various lighting conditions and different subjects to validate the proposed system. All video sequences are acquired by the camera of smartphone, creating images of low resolution (640 × 480 pixels) at 30 frames per second (fps). The arrangement of experimental facilities which is made up of smartphone and supplied frame is shown in Fig. 9. This photograph describes the circumstance of our practical testing under the car cabin environment. The rectangle implies the position of smartphone as a fixed instrument. Under this monitoring condition, the drowsiness can be detected while the view of driver is not obstructed.

The testing platform includes both a PC and a smartphone. To obtain valuable data, the proposed system is evaluated adequately in a realistic driving environment. As shown in Table 1, 15 properly qualified drivers aged between 21 and 40 are volunteered as subjects in this study. All testing subjects are physically fit and free from medication or alcohol. In terms of data analysis, we get the accurate testing results under the environment of PC with a

Fig. 9 *Experimental facilities*

Table 1 Distribution of 15 subjects severed to experiment

Type of subjects	Number of subjects	Total number
male	9	15
female	6	
with glasses	6	15
without glasses	9	

Fig. 10 *Sample of successes and failures on different subjects; left-bottom boxes represent the right eye region*

typical CPU speed of 2.10 GHz. In terms of the speed of our system, our system runs at 110 ms per second in computer environment and achieves the speed at 330 ms per second in smartphone. All codes are initially compiled in C environment. Then they are converted to codes that can cope with running environment of smartphone.

3.2 Procedure and measures

The procedure of experiment is made up of two parts: eye location testing and driver fatigue detection. As the most important part in the system, the success of fatigue detection depends on the efficacy of PLM. Therefore, the eye detection rate is first discussed within different subjects to validate the performance of the proposed system. In terms of detection rate, C_{error} is measured according to (6), which defines the ratio of right frames to the whole ones. Based on the eye detection result, a statistic diagram is shown later in order to verify the integrated system

$$C_{\text{error}} = \frac{\text{Total frames} - \text{Wrong frames}}{\text{Total frames}} \quad (6)$$

3.3 Testing

Our testing database contains over 90 video sequences filmed from the 15 subjects. Each sequence is captured by the smartphone at the speed from 1.5 to 3 fps during our several experiments. The database shows very challenging conditions for the proposed method, like different types of gesture, facial outlines of the subjects, and illumination. In this database, each frame roughly keeps a similar facial scale due to space limitation of car cabin environment. Based on these sequences, the testing results are presented as follows.

3.3.1 Eye detection results: In this part, the detection rate of eye is focused first. As the main idea of our paper, eye detection is more important than driver fatigue detection because fatigue detection rate is mainly determined by the eye detection result. Therefore, we demonstrate our PLM algorithm by two parts: visual diagrams collected by PLM algorithm and statistic results compared with both classic and state-of-art methods. Fig. 10 shows the results quantitatively on the basis of different subjects in our database. As shown in Fig. 10, the success and failure samples are depicted in various subjects. We observe that the method successfully deals with slight changes in pose, scale, and the presence of glasses. By analysing the failure samples (Fig. 10), it can be observed that the algorithm is prone to failure when the eye region is occluded by light reflection on the surfaces of spectacles or the distribution of greyscale around the right eye region is changed by fringe. In these cases, the image segmentation cannot reach a good result, compared with the normal circumstance. Under this condition, eyebrow or glasses frame may replace the importance of eye.

Table 2 Three different groups under realistic driving

Group series	Type of subjects
I	normal
II	with glasses
III	fringe and long haircut

Table 3 Results of eye detection rate under group I

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
total frames	1015	612	417	690	107
failure frames	56	30	5	14	1
detection rate, %	94.48	95.09	98.80	97.97	99.07
average detection rate, %			96.26		

Table 4 Results of eye detection rate under group II

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
total frames	2211	474	722	1268	485
failure frames	125	26	38	126	40
detection rate, %	94.34	94.48	92.38	88.90	93.81
average detection rate, %			93.12		

Table 5 Results of eye detection rate under group III

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5
total frames	1257	373	541	211	753
failure frames	61	22	17	34	14
detection rate, %	95.14	94.10	96.85	83.88	97.61
average detection rate, %			95.27		

Table 6 Comparison of the eye detection rate

Methods	Detection rate, %
Bayesian [27]	92.79
multi-Adaboost [26]	89.07
2-C SRC [23]	90.09
proposed method	94.88

To show a comparable data before comparing our algorithm to other methods, all sequences are divided into three groups, as shown in Table 2. In terms of group I, the subjects with non-glasses and common haircut are selected as member to validate the algorithm in normal condition. Based on the basic testing, we set groups II and III which are not independent because some subjects wear glasses and have a long haircut. Tables 3–5 qualitatively present the accuracy of our eye detection algorithm within different groups. Compared with groups I and III, group II represents subjects who wear glasses where the results achieve worse average detection rate at 93.12% (Table 4). Although abnormal colour detection aiming at glasses reflection is adopted in image segmentation, the distribution of binary pixel in projection step is still affected by glasses reflection. Consequently, with enhanced grey value around the eye region caused by glasses, it is difficult to detect eye with low-quality binary image. By averaging eye detection rate from three groups, our system reaches high detection rate at 94.88%. This detection rate represents the robustness of algorithm when it faces different circumstance in the real.

In terms of comparison part, we compare our eye detection algorithm to classic Bayesian method [27], 2-C SRC method [23], and multi-Adaboost method [26]. To get a convincing comparison result, 840 images are randomly selected in our database. Within all test images, we then randomly choose 80 images for training and 760 images for testing. Table 6 shows the detection rate of proposed method in our own database. Obviously, PLM performs better than other proposed method. In our opinion, although statistic algorithm can reach a high performance in common situation, it is hard to keep detection rate when algorithms are faced with various conditions of image pixel due to the limitation

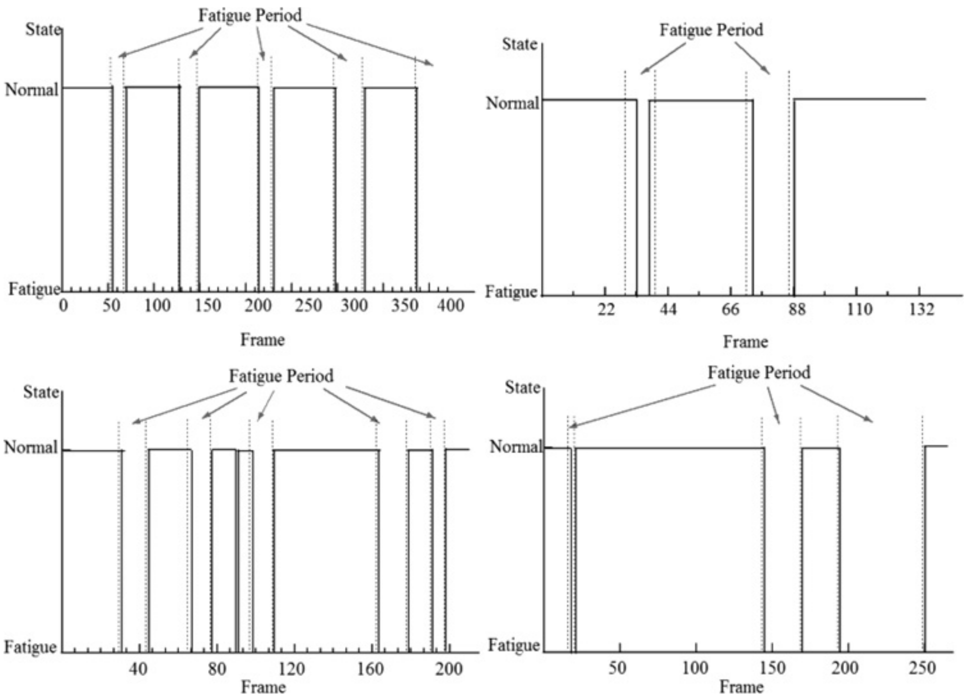

Fig. 11 *Fatigue detection result diagrams*

Four sub-graphs, respectively, depict the changing state between normal and fatigue. Each sub-graph is extracted from the different subjects' sequences. Solid lines are drawn by our system. Dotted lines represent the intervals of actual fatigue period to which arrows point

of training set. Without training data, PLM algorithm pays more attention to colour information, geometric domain, and progressive structure which are combined to help eye detection being more robust.

In conclusion, PLM algorithm achieves a high detection rate at 94.88% of average detection rate and outperforms both the classic and state-of-art method. It can be considered as a well-performing method.

3.3.2 Fatigue detection results: As whole system's output, fatigue detection results are introduced in this section. After classifying different combinations of two features (i.e. dimension and height) into the corresponding state of single frame, the system achieves the output by counting different states within five adjoining frames. Based on accurate eye detection rate, some fatigue detection results collected by our system are presented in Fig. 11. As shown in Fig. 11, four fragments of different subjects' sequences contain, respectively, several fatigue periods. In terms of our experiment, driver fatigue is depicted as the state of closing eye for a period of time by our subjects. So the duration and frequency of fatigue states are under control. In each sub-graph, double dotted lines limit the actual fatigue period to which arrows also point. The fatigue detection results are drawn by the solid line. According to Fig. 11, the start time of actual fatigue period always places in front of the corresponding algorithm-detected fatigue period because the system sends the fatigue state when there are at least three close frames in five adjoining frames. With the comparison within four diagrams, there is an obvious miss in the second diagram. In this sub-graph, the conversion of fatigue state is delayed about five frames during the first fatigue period, while algorithm misses three fatigue frames at the same period. This wrong detection is manifested due to the difficulties in detecting a closed eye. When people close their eyes, the shape of eye converts to a slight line. This phenomenon leads to two usual types of wrong detection under the condition of strong illumination. First, closed eye is prone to be missing in binary face image and PLM algorithm subsequently detects eyebrow instead of the eye. On the other hand, after successfully detecting eye position, it is possible that local eye region is poorly binarised under strong illumination. Compared with similar binary image with closed eye in Fig. 8, the height and dimension of this binarised object cannot offer enough support to classify this frame as an eye closing state. Besides this problem, in Fig. 11, most part of fatigue period can be detected in time. Consequently, the system can timely respond to a driving threat arising from fatigue.

To assess the performance of our system, a qualitative analysis is employed in this paper (Table 7). In terms of the criteria, two groups of factors are introduced. According to [9], accuracy, latency, reliability, and working under severe conditions (i.e. sunglasses or at night) are treated as the convolution factors. In accordance with the background in this paper, two factors (cost in smartphone and application in smartphone) with smartphone are also included in this analysis. As comparison objects, we select the papers [1, 9, 17, 20] that are similar to our proposed system, especially for [20] that had been transplanted into smartphone in the past year. However, this method based on cellular nonlinear networks (CNN)

Table 7 Comparison of our system and other systems

System	Accuracy	Latency	Reliability	Working under severe conditions	Cost in smartphone	Application in smartphone
Fatigue detection system (FDS) [9]	high	low	medium	yes	–	no
Rafi Ahmed *et al.* [1]	high	medium	high	yes	high	yes
Xiaoqing Luo *et al.* [17]	medium	low	high	yes	–	no
CNN [20]	not mentioned	extremely low	uncertainty	no	low	yes
our system	high	low	high	no	low	yes

method could be treated as an attempt because its accuracy and reliability cannot be ensured. On the basis of CNN, the speed of this system is convincing because each part of this algorithm is simple. However, these algorithms that rely on several template and an update equation is so simple that the robustness of the whole system may not be achieved within the whole facial region. On the other hand, the other systems are put forward on PC. Although the papers [9, 17] show a balance performance in four conventional criteria, the input of two systems is IR images. However, the camera that can receive IR light is rarely employed in the platform of smartphone. Among three compared systems, Rafi Ahmed *et al.* [1] proposed an integrated system that detects yawn state, eye state, and lane departed at the same time. There is no doubt that the system can show a high performance in terms of reliability and accuracy. Due to the method employed in [1], this system could be transplanted into smartphone. However, detecting driver fatigue in a variety of ways leads to the decay, especially for smartphone. As the comparison of PC and smartphone mentioned in introduction part, a high storage cost in smartphone can lead to a high computation cost. Consequently, the system proposed in [1] is insufficient to embedding in smartphone. Compared with the other method, our system shows a more balance performance. Although cannot deal with severe conditions due to the limitation of camera, the system is executed with a high eye detection accuracy, low latency and cost in smartphone because the motivation of this system is to focus the driver fatigue detection based on smartphone.

4 Conclusions and discussions

In this paper, we propose a driver fatigue detection system on smartphone platform. Our system is made up of face detection, PLM algorithm, and fatigue detection. Our main contribution in system is on designing a low computation cost system based on PLM algorithm that can be also considered as a contribution in this paper. In terms of PLM algorithm, two steps of mouth detection for coarse eye detection and greyscale projection for further detection are included. On the basis of previous works, the state of single frame is achieved by classifying the different combination of two features (i.e. dimension and height), which are measured in binarised local eye region. With counting the states within five adjoining frames, system responds to the driver fatigue when there are at least three close frames in total five frames.

In terms of experiments, we employ the self-made database for eye detection and fatigue detection test. As the emphasis part, eye detection test divides into three groups: normal, glasses, and different types of haircut. Each group consists of five subjects. As a result, PLM algorithm shows good performance at 94.88% of average accuracy in contrast to the previous methods. In the second part, fatigue detection result is shown by the state conversion diagram, which is achieved by analysing the extracted sub-sequences in same database. It is theoretically and experimentally demonstrated that the system yields good adjustable and robust property in different situations.

On the other hand, the limitation of our PLM algorithm is over-relying on grey distribution in facial region. If the binary image of secondary localisation region cannot show the eye region clearly, PLM cannot detect the eye by projecting binary result. Therefore in the future, the extension of our method based on idea of progress is studying the performance on finding a proper way to replace the way of segmentation. Another improvement of PLM that can be attracted is detecting eye from a rotated face or in more complicated situation. We believe that our effort contributes in advancing fatigue detection especially on smartphone applications.

5 Acknowledgments

This work was supported by National Natural Science Foundation of China (20222201 and 61305026) and Beijing Municipal Commission of Education (KM200710005009, PXM2009_014204_09_000154 and KM201310005006).

6 References

1 Ahmed, R., Emon, K.E.K., Foisal Hossain, Md.: 'Robust driver fatigue recognition using image processing'. Third Int. Conf. on Informatics, Electronics, and Vision, Dhaka, The People's Republic of Bangladesh, May 2014, pp. 1–6

2 Thiffault, P., Bergeron, J.: 'Monotony of road environment and driver fatigue: a simulator study', *Accident Anal. Prev.*, 2003, **35**, pp. 381–391

3 Karrer, K., Roetting, M.: 'Effects of driver fatigue monitoring – an expert survey'. Seventh Int. Conf. on Engineering Psychology and Cognitive Ergonomics, Beijing, China, 22–27 July 2007, pp. 324–330

4 Merat, N., Hamish Jamson, A.: 'The effect of three low-cost engineering treatments on driver fatigue: a driving simulator study', *Accident Anal. Prev.*, 2013, **50**, pp. 8–15

5 Boyle, L., Tippin, J., Paul, A., *et al.*: 'Driver performance declines in the moments surrounding a microsleep', *Transp. Res. F, Traffic Psychol. Behav.*, 2008, **11**, pp. 126–136

6 Ingre, M., Kerstedt, T., Peters, *et al.*: 'Subjective sleepiness and accident risk avoiding the ecological fallacy', *J. Sleep Res.*, 2006, **15**, pp. 142–148

7 Syed, R., Eskandarian, A.: 'Unobtrusive drowsiness detection by neural network learning of driver steering', *Proc. Inst. Mech. Eng. D, J. Automob. Eng.*, 2001, **215**, (9), pp. 969–975

8 Eskandarian, A., Mortazavi, A.: 'Evaluation of a smart algorithm for commercial vehicle driver drowsiness detection'. IEEE Intelligent Vehicles Symp., Istanbul, Turkey, 13–15 June 2007, pp. 553–559

9 Ibrahim, L.F., Abulkhair, M., AlShomrani, A.D., *et al.*: 'Using Haar classifiers to detect driver fatigue and provide alerts', *Multimedia Tools Appl.* **71**, (3), pp. 1857–1877

10 He, Q., Li, W., Fan, X., *et al.*: 'Driver fatigue evaluation model with integration of multi-indicators based on dynamic Bayesian network', *IET Intell. Transp. Syst.*, 2015, **9**, (5), pp. 547–554

11 Lala, S.K.L., Craiga, A., Boorda, P., *et al.*: 'Development of an algorithm for an EEG-based driver fatigue countermeasure', *J. Saf. Res.*, 2003, **34**, pp. 321–328

12 Patel, M., Lal, S.K.L., Kavanagh, D., *et al.*: 'Applying neural network analysis on heart rate variability data to assess driver fatigue', *Expert Syst. Appl.*, 2011, **38**, pp. 7235–7242

13 Zhang, L., Liu, F., Tang, J.: 'Real-time system for driver fatigue detection by RGB-D camera', *ACM Trans. Intell. Syst. Technol.*, 2015, **6**, (2), pp. 1–17

14 Azim, T., Arfan Jaffar, M., Mirzada, A.M.: 'Fully automated real time fatigue detection of drivers through fuzzy expert systems', *Appl. Soft Comput.*, 2014, **18**, pp. 25–38

15 Cyganek, B., Gruszczyński, S.: 'Hybrid computer vision system for drivers' eye recognition and fatigue monitoring', *Neurocomputing*, 2014, **126**, pp. 78–94

16 McCall, J.C., Trivedi, M.M.: 'Visual context capture and analysis for driver attention monitoring'. 2004 IEEE Intelligent Transprtation Systems Conf., Washington, DC, USA, 3–6 October 2004, pp. 332–337

17 Luo, X., Hu, R., Fan, T.: 'The driver fatigue monitoring system based on face recognition technology'. 2013 Fourth Int. Conf. on Intelligent Control and Information Processing, Beijing, China, 9–11 June 2013, pp. 384–388

18 Xiong, C., Xie, M., Wang, L.: 'Driver fatigue detection based on AdaBoost', *Adv. Intell. Soft Comput.*, 2012, **159**, pp. 13–17

19 Yong, D., Degang, C., Qinghua, H., *et al.*: 'Kernelized fuzzy rough sets based yawn detection for driver fatigue monitoring', *Fundam. Inform.*, **111**, (1), pp. 65–79

20 Buscarino, A., Fortuna, L., Frasca, M.: 'Driving assistance using smartdevices'. 2014 IEEE Int. Symp. on Intelligent Control (ISIC), Part of 2014 IEEE Multi-Conf. on Systems and Control, Antibes, France, 8–10 October 2014, pp. 838–842

21 He, J., Roberson, S., Fields, B., *et al.*: 'Fatigue detection using smartphones', *J. Ergon.*, 2013, **3**, pp. 1–7

22 Ren, Y., Wang, S., Hou, B., *et al.*: 'A novel eye localization method with rotation invariance', *IEEE Trans. Image Process.*, 2014, **23**, (1), pp. 226–239

23 Xinjin, Wei.: 'The study of fatigue detection of drivers based on the infrared conditions'. Master thesis, Zhejiang Sci-Tech University, 2011

24 Reinders, M.J.T., Koch, R.W.C., Gerbrands, J.J.: 'Locating facial features in image sequences using neural networks automatic face and gesture recognition'. Proc. Second Int. Conf. on Automatic Face and Gesture Recognition, Killington, USA, October 1996, pp. 230–235

25 Li, S.Z., Chu, R.F., Liao, S.C., *et al.*: 'Illumination invariant face recognition using near-infrared images', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2007, **29**, (4), pp. 627–639

26 Everingham, M., Zisserman, A.: 'Regression and classification approaches to eye localization in face images'. Seventh Int. Conf. on Automatic Face Gesture Recognition, Southampton, UK, April 2006, pp. 2105–2112

27 Gercia, H., Salazar, A., Alvarez, D., *et al.*: 'Driving fatigue detection using active shape models'. Sixth Int. Symp. on Advances in Visual Computing, Las Vegas, NV, USA, November 2010, pp. 171–180

28 Valenti, R., Gevers, T.: 'Accurate eye center location through invariant isocentric pattern', *IEEE Trans. Pattern Anal. Mach. Learn.*, 2012, **34**, (8), pp. 1785–1798

29 Zhang, Z., Zhang, J.: 'A new real-time eye tracking based on nonlinear unscented Kalman filter for monitoring driver fatigue', *J. Control Theory Appl.*, 2010, **8**, (2), pp. 181–188

30 Alioua, N., Amine, A., Rziza, M., *et al.*: 'Driver's fatigue and drowsiness detection to reduce traffic accidents on road'. 14th Int. Conf. on Computer Analysis of Images and Patterns, Seville, Spain, August 2011, pp. 397–404

31 Soylemez, O.F., Ergen, B.: 'Eye location and eye state detection in facial images using circular Hough transform'. 12th IFIP TC8 Int. Conf., Krakow on Computer Information Systems and Industrial Management, Poland, September 2013, pp. 141–147
32 Wang, Q., Yang, J.Y.: 'Eye location and eye state detection in facial images with unconstrained background', *J. Inf. Comput. Sci.*, 2006, **1**, (5), pp. 284–289
33 Yuille, A., Hallinan, P., Cohen, D., *et al*.: 'Feature extraction from faces using deformable templates', *Int. J. Comput. Vis.*, 1992, **8**, (2), pp. 99–111
34 Phillips, P., Wechsler, H., Huang, J., *et al*.: 'The FERET database and evaluation procedure for face-recognition algorithms', *Image Vis. Comput.*, 1998, **16**, (5), pp. 295–306
35 Phillips, P., Moon, H., Rizvi, S., *et al*.: 'The feret evaluation methodology for face-recognition algorithms', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2000, **22**, (10), pp. 1090–1104
36 Wierwille, W., Ellsworth, L., Wreggit, S., *et al*.: 'Research on vehicle-based driver status/performance monitoring: development, validation and refinement of algorithms for detection of driver drowsiness'. DOT HS 808 247, National Highway Traffic Safety Administration, 1994,
37 Li, L., Chen, Y., Li, Z.: 'Yawning detection for monitoring driver fatigue based on two cameras'. Proc. 12th Int. IEEE Conf. on Intelligent Transportation Systems, St. Louis, USA, 3–7 October 2009, pp. 12–17
38 Fan, X., Yin, B.-c., Sun, Y.-f.: 'Yawning detection for monitoring driver fatigue'. Proc. Sixth Int. Conf. on Machine Learning and Cybernetics, Hong Kong, 19–22 August 2007, pp. 664–668
39 Azim, T., Jaffar, A., Mirza, A., *et al*.: 'Automatic fatigue detection of drivers through pupil detection and yawning analysis'. 2009 Fourth Int. Conf. on Innovative Computing, Information and Control, Kaohsiung, Taiwan, 7–9 December 2009, pp. 441–445