Multi-feature fatigue driving detection based on computer vision

Juan Huang1a*, Zihui Lin1

1School of Information Engineering, JiangMen Polytechnic, JiangMen, GuangDong, 529000, China

aym1814@wyu.edu.cn *Corresponding author’s e-mail: lovehj23@126.com

Abstract. Fatigue driving is one of the main causes of traffic accidents. This paper proposes a fatigue detection method based on computer vision. The first is the introduction of an optimized algorithm, based on AdaBoost, to detect the face area, and then the ERT algorithm is used to achieve precise localization of the facial landmarks. Finally, a variety of fatigue features of eyes and mouth state associated with driving fatigue are extracted, and after the fusion of all these features, the fatigue driving detection is performed. The experimental results show that multi-feature detection is more accurate than single feature detection.

1. Introduction
As the number of automobiles increases, Road traffic safety issues have become increasingly prominent, and fatigue driving is one of the leading causes of accidents in transportation sector. Therefore, in order to reduce fatigue driving accidents, it is of great practical significance to study a fatigue driving early warning system with good real-time performance, high accuracy and strong reliability.

Many methods have been proposed to fatigue driving detection: Several methods are based on the driving behavior detection by machine learning algorithm trained through data extracted from Steering Wheel Movements(SWM)[1] and brakes monitoring[2]. However, this type of method requires a large amount of data on the state of driver drowsiness. Differences in road scenes, vehicle models, and driver behavior habits will cause large deviations in the detection results. A more reliable approach relies on psychophysiological state monitoring such as Electroencephalogram (EEG)[3] measurement by identifying informative brain region with specific frequency bands and forehead electrooculograms (EOGs)[4] acquired by wearable dry electrodes. Huo[5] fused EEG and forehead EOG to detect drivers’ fatigue level. While the psychophysiological approach has high accuracy, the needed detection equipment will hinder the operation of the driver to a certain extent which reduces the safety of driving. Moreover, the equipment needed for detection is relatively expensive and low cost performance. Y. S. Wu et al.[6] proposed a computer vision-based detection approach which judges the fatigue state by analyzing the eye feature of the driver. This method uses a face detection algorithm based on the AdaBoost, then uses the SVM classifier in the face area to determine the position of the eyes, and identify fatigue driving with the eye features. MC Catalbas et al.[7] proposed a driver fatigue level determination system based on detection of saccadic eye movements. Acceleration, speed, and size of pupils at various traffic scenarios were compared to determine the fatigue level of the driver. The computer vision based approach is more practical as it has no interference to the driver.
In summary, the fatigue detection based on computer vision has become the current leading approach, but its real-time performance and accuracy of single fatigue feature recognition needs to be improved. In this paper, we proposed a method of fast fatigue driving detection. We first used AdaBoost to detect human face, optimized for the specific vehicle use environment, then located and analyzed eyes and mouth state using ERT to determine fatigue driving. The flowchart of fatigue detection is shown in figure 1.

![Flowchart of fatigue detection](image)

**Figure 1. Flowchart of fatigue detection**

2. **Face detection**

Face detection can be affected by the illumination, the reflectivity of the face surface and the face geometric characteristics, which may cause noise pollution and brightness distortion in the image. In order to improve the detection efficiency and accuracy, the original image is grayscaled by weighted average of RGB, each color is weighted differently[8].

\[
\text{gray}(i, j) = 0.1140B(i, j) + 0.5870G(i, j) + 0.2989R(i, j)
\]

The illumination in the vehicle is usually complicated, and will affect the face detection accuracy. The histogram equalization can reduce such effects. As shown in figure 2, the concentrated image grayscale distribution can be corrected by the histogram equalization, thus improves the image quality.

![Comparison of original image and histogram equalized Image](image)

**(a). Original image**  **(b). Image after histogram equalization**

**Figure 2. Comparison of original image and histogram equalized Image.**

First target of face detection is to judge whether the input image contains a human face, and give the area of the face if it exists. Viola and Jones proposed a face detection algorithm [9-10] based on AdaBoost, with a new image representation called the “Integral Image”, a set of features with reminiscent of Haar Basis functions and a cascade classifier. By training a large number of face and non-face samples, AdaBoost obtains strong classifiers from weak classifiers, then the strong classifiers are cascaded to generate a cascaded detector, which greatly improves the accuracy and speed of detection.

When applying the sliding window to detect human faces, because the size of the face in the image is different, the image needs to be zoomed in different proportions. Because a large number of non-face windows will slow the detection, we optimize the algorithm for this specific use environment in the vehicle. The optimized algorithm is shown in figure 3.
In the specific application of fatigue driving detection, the driver is the only detection target with relatively fixed posture. The window of the face will be in a relatively small region, therefore we set the region of interest (ROI) in the next frame by the appropriately enlarged window of the face in the previous frame, this can remove a large number of background areas in the image, and accelerates the algorithm. Through the analysis of the window of face’s displacement changes between consecutive frames, the face window of the previous frame is enlarged 1.3 times horizontally and 1.2 times vertically as the ROI of the current frame. Because the fatigue detection is for the driver alone, and the driver is closer to the acquisition device, which occupies a largest proportion in the image, we improve the performance by increasing the minimum detectable window of face to reduce the number of detection. This can also reduce interference of the rear passengers and improve the accuracy. The face detection result is shown in figure 4, where the outer border is the ROI, and the inner border is the window of face.

![Figure 4. ROI of face detection](image)

### 3. Facial landmarks localization and facial state recognition

**3.1. Facial landmarks localization**

The purpose of facial landmarks localization is to locate the facial feature (such as eyes, nose, mouth, and outer contour of the face) on the basis of face detection. In order to reduce the interference of external environmental factors on the detection and localization of facial landmarks, we uses the ensemble of regression trees (ERT) to locate the facial landmarks, and then according to the position of the landmarks, the eyes and mouth are quickly and accurately located.

Proposed by Kazemi and Sullivan[11], ERT can be used to estimate the facial landmarks’ positions directly from a sparse subset of pixel intensities. We use an ERT model with 68 key landmarks for localizing human faces by establishing a cascaded gradient boosting decision tree (GBDT), as shown in figure 5, and achieves real-time performance of facial landmarks localization, as shown in figure 6.

![Figure 3. Flow chart of the optimized face detection algorithm](image)
3.2. Eye state recognition
Among 68 facial landmarks obtained by ERT, there are 12 for eyes, 6 for left eye and 6 for right eye respectively. As shown in figure 7, landmarks for the left eye and right eye are numbered 36 to 41 and 42 to 47. In order to accurately and quickly identify the open and closed state of the eyes, we calculate the eye aspect ratio (EAR) [12] as in equation (2):

$$\text{EAR} = \frac{||P_{36}-P_{42}||+||P_{38}-P_{40}||}{2||P_{36}-P_{40}||}$$  \hspace{1cm} (2)

When the eyes are opened, the values of EAR change in a small range. If the value of EAR drops rapidly to close to 0 and then resumes, it indicates that a blink has occurred. If the value of EAR drops suddenly and does not resume within a period of time, this indicates that the driver’s eyes have closed. EAR was calculated and plotted in 200 consecutive frames, which include blink, as shown in figure 8.

![Eye landmarks](image)

**Figure 7. Eye landmarks**

![EAR value of 200 frames](image)

**Figure 8. EAR value of 200 frames**
3.3. Mouth state recognition
One of the common reactions of fatigue driving is yawning, which is determined by the state of the mouth. As shown in figure 9, the landmarks of the mouth are number 48 to 67. We locate the mouth and recognize its state by calculating the mouth aspect ratio (MAR).

\[
\text{MAR} = \frac{||P_{61}-P_{67}|| + ||P_{62}-P_{66}|| + ||P_{63}-P_{65}||}{3||P_{60}-P_{64}||}
\]  

(a). Mouth open  
(b). Mouth closed  

Figure 9. Mouth landmarks

Under normal driving conditions, the driver’s mouth is closed. Yawning is detected by measuring both the rate and the amount of the changes in the mouth [13]. In order to judge the state of the mouth, such as speaking, yawning, etc., MAR was calculated and plotted in 200 consecutive frames, which include talking and yawning, as shown in figure 10. When MAR < 0.2, the mouth is closed. When talking or laughing, the MAR value is larger, when MAR > 8.0 and the duration is 3 to 5 seconds, indicates that the driver is yawning.

Figure 10. MAR value of 200 frames

4. Fatigue state recognition

4.1. Multi-feature weighted sum for fatigue state recognition
Fatigued drivers have multiple features such as blink, closed eyes, and yawning. By fusing multiple features, the accuracy of fatigue detection could be improved. To establish a fatigue state recognition model based on information fusion, we calculate the fatigue features of eyes and mouth, and use multi-feature weighted sum to judge the fatigue state of the driver. Features extracted are the percentage of eye close ratio (ECR), blink frequency (BF), max eye closed time (MECT), and yawn frequency (YF). The equation of multi-feature weighted sum is:

\[
F = \sum W_i \times Value_i, \sum W_i = 1, Value_i \in \{1, 0\}, i = (ECR, BF, MECT, YF)
\]

PERCLOS is a reliable criterion for fatigue driving detection. It calculates the percentage of human eye closure time in the total time within a time period [14]. P80 criteria in PERCLOS is the most suitable for identifying fatigue driving. It means the percentage of total time that eye is closed at least
80%. EAR can make a good judgment on the state of closed eyes, we use the ECR as the eye feature parameter as in equation (5).

\[
ecr = \frac{\text{the sum of frames when eye is closed}}{\text{the total number of frames in the time period}} \times 100\%
\] (5)

Normally, driver will blink 10 to 25 times per minute on average, while in a fatigue state, the number of blinks will increase. Blink frequency can be judged by EAR. Set the threshold as \(a\) and the number of frames as \(n\), we count the number \(N\) of consecutive frames when \(\text{EAR} < a\). A blink is detected when \(\text{EAR} > a\) and \(N > n\). Through experiments, it is found that the accuracy of judging blink is high when the threshold \(a\) is 0.2 and the frame number \(n\) is 3.

When the driver is drowsy, the eye closure time often exceeds 1.3 seconds. By examining the number of eye closure frames, we calculate the max eye close time (MECT). If the frame rate is \(f\), the number of frames for keeping eyes closed is \(K\), then we have:

\[
mect = \frac{K}{f}
\] (6)

Fatigued driver may yawn. We use MAR to detect yawn, by judging if \(\text{MAR} > 0.8\) and the duration exceeds 3 seconds. Then the frequency of yawning (YF) is calculated.

| Value | Condition | ECR | BF | MECT | YF |
|-------|-----------|-----|-----|------|-----|
| 1     | ≥ 0.2, <0.2 | 0   | 0   | 1    | 1   |
| 1     | > 25, ≤ 25 | 2   | 2   | 1    | 1   |

4.2. Experimental results

In this section the performance of multi-feature weighted sum for fatigue state recognition is evaluated. Through the experimental analysis of the fatigue driving video, the optimal weights of the eye and mouth fatigue features are determined, as shown in table 2.

| Fatigue feature value | Weight |
|----------------------|--------|
| ECR                  | 0.3    |
| BF                   | 0.2    |
| MECT                 | 0.3    |
| YF                   | 0.2    |

According to the equation (4), when \(F\) exceeds 0.3, it is considered as mild fatigue, when it exceeds 0.5, it is considered as moderate fatigue, and when it exceeds 0.8, it is considered as severe fatigue. We divide the experiment into 5 groups and conduct 100 experiments. The accuracy values of the proposed multi-feature weighted fatigue state recognition method and the accuracy values of single feature are compared in table 3.

| Fatigue detection feature | Correct determination count | False determination count | Accuracy |
|---------------------------|-----------------------------|---------------------------|----------|
| ECR                       | 83                          | 17                        | 83%      |
| BF                        | 74                          | 26                        | 74%      |
| MECT                      | 57                          | 43                        | 57%      |
| YF                        | 69                          | 31                        | 69%      |
| Fusion of multi-feature   | 96                          | 4                         | 96%      |

As shown in the experimental results, the result of experiments using weighted sum of multi-feature fatigue detection achieved highest accuracy.
5. Conclusion
In this paper a multi-feature fatigue detection based on machine vision is proposed. The proposed method can detect the driver’s face efficiently by optimizing an Adaboost face detection algorithm. The ERT algorithm is used to realize the precise positioning of facial landmarks. We calculated multiple features of eye and mouth fatigue state, and fused the weighted features for fatigue driving detection. The experimental results show that recognition accuracy has been greatly improved by the proposed method.

Acknowledgments
This work is supported by JiangMen Polytechnic and the department of education of GuangDong province, China, under Grant No. 2017GkQNCX087. The authors would also like to thank the paper’s anonymous reviewers for their insightful comments.

References
[1] Lawoyin, S., Liu, X., Fei, D.Y., Ou, B. (2014) Detection methods for a low-cost accelerometer-based approach for driver drowsiness detection. In: IEEE International Conference on Systems, Man, and Cybernetics (SMC). San Diego. pp. 1635-1641.
[2] Mühlbacher-Karrer, J., Mosa, A.H., Faller, L.M., et al. (2017) A Driver State Detection System—Combining a Capacitive Hand Detection Sensor With Physiological Sensors. J. IEEE Transactions on Instrumentation and Measurement., 66:624-636.
[3] Chuang, C.H., Huang, C.S., Ko, L.W., et al. (2015) An EEG-based perceptual function integration network for application to drowsy driving. J. Knowledge-Based Systems., 80:143-152.
[4] Zheng, W.L., Gao, K.P., Li, G., et al. (2020) Vigilance Estimation Using a Wearable EOG Device in Real Driving Environment. J. IEEE Transactions on Intelligent Transportation Systems., 21:170-184.
[5] Huo, X.Q., Zheng, W.L., Lu, B.L. (2016) Driving fatigue detection with fusion of EEG and forehead EOG. In: International Joint Conference on Neural Networks (IJCNN). Vancouver. pp. 897-904.
[6] Wu, Y.S., Lee, T.W., Liu, H.S. (2010) An Eye State Recognition Method for Drowsiness Detection. In: Vehicular Technology Conference. Taipei. pp. 1-5.
[7] Catalbas, M.C., Cegovnik, T., Sodnik, J., Gulten, A. (2017) Driver fatigue detection based on saccadic eye movements. In: International Conference on Electrical and Electronics Engineering (ELECO). Bursa. pp. 913-917.
[8] Guo, H.L., Wang, N., Guo, H. (2018) Research on fatigue driving early warning system based on multiple signal characteristics. J. Journal on Communications., 39:22-29.
[9] Viola, P., Jones, M. (2001) Rapid object detection using a boosted cascade of simple features. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai. pp. 511-518.
[10] Viola, P., Jones, M.J. (2004) Robust Real-Time Face Detection. J. International Journal of Computer Vision., 57, 137–154.
[11] Kazemi, V., Sullivan, J. (2014) One Millisecond Face Alignment with an Ensemble of Regression Trees. In: IEEE Conference on Computer Vision and Pattern Recognition. Columbus. pp. 1867-1874.
[12] Soukupová, T., Čech, J. (2016) Real-Time eye blink detection using facial landmarks. In: 21st Computer Vision Winter Workshop. Rimske terme. pp. 1-8.
[13] Abtahi, S., Shirmohammadi, S., Hariri, B., et al. (2013) A yawning measurement method using embedded smart cameras. In: IEEE International Instrumentation and Measurement Technology Conference (I2MTC). Minneapolis. pp. 1605-1608.
[14] Mandal, B., Li, L., Wang, G.S., et al. (2017) Towards Detection of Bus Driver Fatigue Based on Robust Visual Analysis of Eye State. J. IEEE Transactions on Intelligent Transportation Systems., 18: 1-13.