

Train driver fatigue detection based on facial multi-information fusion

HAO Zhengqing¹, WANG Ying^{1,2}, CHEN Xiaoqiang^{1,2}, XIONG Ye¹

(1. School of Automation & Electrical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;

2. Key Lab of Opt-Electronic Technology and Intelligent Control, Ministry of Education, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: In order to improve the accuracy of train driver fatigue detection, a method of train driver fatigue detection based on facial multi-information fusion is proposed. Firstly, low-light enhancement is used for image preprocessing, and human faces are detected by local binary patterns (LBP) feature. Secondly, the driver's facial feature points are obtained by ensemble of regression trees (ERT) algorithm, and face model matching is used to obtain the driver's head posture angle. Finally, according to the special driving environment of the train driver, adaptive threshold correction and eye gaze correction are carried out for the eye characteristic quantities that best show fatigue. The fuzzy inference system is used as a fusion tool, and the features of eyes, mouth and head are used as the input of the fuzzy inference system, and the driver's fatigue value is used as the detection results. Experiment results show that the detection method can distinguish driver fatigue levels with accuracy rates of 95% in normal environments and 86.8% in low-light environments.

Key words: train driver; fatigue detection; feature point detection; head posture; facial multi-information fusion

0 Introduction

With the rapid development of China's railway industry, train mileage has been increased and working conditions have become complex, resulting in more and more fatigue driving situations for train drivers. Therefore, it is important to study the fatigue state of train drivers for driving safety^[1].

The driver fatigue detection is mainly divided into detection methods based on physiological information^[2], based on driving behavior^[3] and based on facial expressions^[4-5]. The ease of operation, low cost of the facial expression detection method and the increase in computing power with modern computers have made the method become mainstream for studying driver fatigue.

The fatigue detection of drivers mainly has the following steps. (1) Face detection; commonly used methods such as Adaboost algorithm^[6], convolutional neural network (CNN)^[7] and MTCNN^[8]; (2) Feature point localization such as CNN^[10], face sequence^[12], template matching method^[13], and Haar-like^[14]; (3) Fatigue feature extraction such as PERCOLS value, blink frequency,

yawn time, continuous eye closure time, eye opening and closing, mouth opening and closing, head posture, etc.; (4) Fatigue determination, CNN, feature fusion^[9], plain Bayesian classifier, fuzzy algorithm, etc. Driving a train is different from driving a car. In addition to looking straight ahead during the driving process, at the same time, various dashboard information have been paid attention to. And the train will also operate in a low-light environment with alternating light and dark conditions as well as long periods of darkness.

A train driver fatigue detection method is proposed that considers the multi-information fusion of the face in the low-light environment for the characteristics of low illumination of the train traveling at night. The low-light enhancement of the images collected in the low-light working condition of the train reduces the influence of environmental factors on the subsequent detection. The eye characteristics are corrected when the train driver has to pay attention to the dashboard information. The establishment of a fuzzy inference system is useful to realize the train driver fatigue detection, improve the detection accuracy and increase the reliability of

Received date: 2023-02-25

Foundation items: National Natural Science Foundation of China (No. 51767013)

Corresponding author: WANG Ying (wangying01@mail.lzjtu.cn)

fatigue determination.

1 Face detection and feature point localization

In order to improve the adaptability of fatigue detection methods to different driving environments, the low-light enhanced local binary patterns (LBP) algorithm for face detection is used, and the ensemble of regression trees (ERT) algorithm is used in the face detection region to achieve 68 feature points localization, which is more conducive to the extraction of fatigue information.

1.1 Low light enhancement based LBP algorithm

The light in the driver's room keeps changing, and the skin color of the face will be affected by the light leading to uneven skin color brightness and darkness, which affects the accuracy of detection. After the images are captured by using the camera, they are processed with LIME algorithm, histogram equalization (HE) algorithm, and gamma correction (GC) algorithm for low light enhancement. The face feature points do not be detected from the original images. The three algorithms are able to locate feature points. The image details are clearer after processed by LIME algorithm. The image is smoother after processed by HE algorithm. Compared to HE algorithm, exposure processing of the GC algorithm is stronger, while the GC algorithm takes the least time. So low light of the GC algorithm is enhanced, and the three algorithms processing results are shown in Fig. 1.

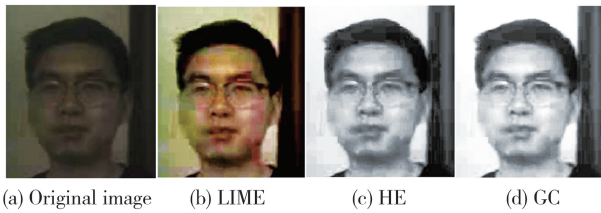


Fig. 1 Processing effect of three low-light enhancement algorithms

LBP^[15] has the advantages of gray invariance and rotation invariance. The center pixels of the neighborhood are taken as the threshold and compared with the neighboring pixels to generate a string of binary numbers as the LBP value of the center pixel. And the LBP value of the center pixels reflects the texture information of the region around the pixel. Due to the high discriminative feature texture of the face and the fast computation speed of

LBP algorithm, it is more appropriate to use LBP to detect the face, and the specific process is shown in Fig. 2. The LBP algorithm is implemented in 3 steps. (1) LBP feature extraction. In the neighborhood of size 3×3 , the center pixel of the neighborhood is used as the threshold, and the grayscale values of 8 adjacent pixels are compared with it. If the surrounding pixel value is greater than the center pixel value, the position of the pixel point is marked as 1, otherwise it is 0. That is to say, the LBP value of the center pixel point of the neighborhood is obtained, and the LBP value of the whole image is traversed sequentially. (2) Histogram of statistical face subregions. The image is divided into subregions of size 7×7 , and the histogram is counted in the subregion according to the LBP value. (3) LBP feature matching. The histogram is used as its discriminative feature to compare with the standard face to get the face region.

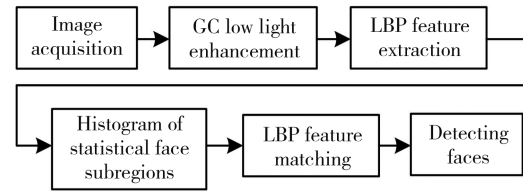


Fig. 2 Face detection process

1.2 Feature point location

ERT^[16] is a regression forest-based algorithm for face feature point localization, which first estimates an approximate feature point location and then uses a gradient boosting algorithm to reduce the sum of squared errors between the original shape and the true shape until the iteration requirement is reached.

Step 1: The training data is $[(I_{\pi i}, \hat{S}_i^{(t)}, \Delta S_i^{(t)})]_{i=1}^N$, the learning rate is $0 < \gamma < 1$, and the initialization $f_i(I, \hat{S}^{(t)})$ is shown as

$$f_0(I, \hat{S}^{(t)}) = \arg \min_{\gamma \in R^{2p}} \sum_{i=1}^N \|\Delta S_i^{(t)} - \gamma\|^2, \quad (1)$$

where I is the image; $\hat{S}^{(t)}$ is the feature point coordinates, and R is the initialization multiplier.

Step 2: Iterative update r_i is expressed as

$$r_{ik} = \Delta S_i^{(t)} - f_{k-1}(I_{\pi i}, \hat{S}_i^{(t)}). \quad (2)$$

Step 3: Get the regression equation $g_k(I, \hat{S}^{(t)})$ by r_{ik} , and update $f_k(I, \hat{S}^{(t)})$.

$$f_k(I, \hat{S}^{(t)}) = f_{k-1}(I, \hat{S}^{(t)}) + \nu g_k(I, \hat{S}^{(t)}), \quad (3)$$

$$r_t(I, \hat{S}^{(t)}) = f_k(I, \hat{S}^{(t)}). \quad (4)$$

Step 4: Repeat steps 2 and 3 until convergence or reaching the number of iterations, the final output is

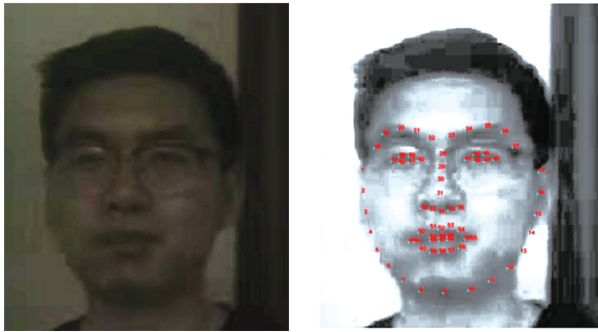
$$f_{k+1}(I, \hat{S}^{(t)}) = f_k(I, \hat{S}^{(t)}). \quad (5)$$

Step 5: Update the location.

$$\hat{S}^{(t+1)} = \hat{S}^{(t)} + r_t(I, \hat{S}^{(t)}), \quad (6)$$

$$\Delta \hat{S}_i^{(t+1)} = \hat{S}_{\pi_i}^{(t)} - \hat{S}^{(t+1)}, \quad (7)$$

where $\hat{S}^{(t+1)}$ denotes the predicted feature point coordinates and $r_t(I, \hat{S}^{(t)})$ denotes the residual coordinates calculated by the current layer. The face feature points localized by ERT algorithm are shown in Fig. 3.



(a) Original image (b) Feature point positioning

Fig. 3 Feature point positioning effect

2 Fatigue characterization

2.1 Eye feature extraction

The ocular feature parameters change most significantly when the driver feel fatigued. To increase the detection robustness, the eye aspect ratio (EAR) is used to define the opening and closing as shown in Fig. 4, which is more accurate compared to the monocular EAR^[17].

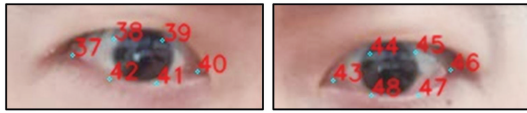


Fig. 4 Eye feature point coordinates

$$EAR = \frac{\|P_{38} - P_{42}\| + \|P_{39} - P_{41}\|}{4\|P_{37} - P_{40}\|} + \frac{\|P_{44} - P_{48}\| + \|P_{45} - P_{47}\|}{4\|P_{43} - P_{46}\|}. \quad (8)$$

Each drivers have different eye sizes, so a fixed EAR threshold will cause false detection. The adaptive threshold method is proposed to determine the EAR threshold for each driver, and the two state

EAR sequences of three drivers with different eye sizes are selected by using the k -means++ method for adaptive threshold, and the experimental results are shown in Table 1.

Table 1 Clustering data of eye opening and closing degree

Group	Normal eye clustering center		Fatigue eye clustering center	
	Open	Close	Open	Close
1	0.282	0.266	0.273	0.152
2	0.303	0.294	0.259	0.195
3	0.223	0.206	0.202	0.137

Considering that the driver's attention to various information on the dashboard will affect the ears, the three drivers' normal open eyes are 0° gaze, the fixed camera and face position are 0.7 m, and the height of the camera is equal to the height of the human eye. The open and closed data of three drivers of the six fixations are measured by sliding the camera position up and down according to the degree of fixation. The 0° open and closed data is used to derive the closed eye data for each angle and fit the data, and the experimental data are shown in Table 2.

Table 2 Eye gaze and eye closure data

Group	Eye closure/ $^\circ$					
	0	5	10	15	20	25
1	0	0.022	0.026	0.068	0.085	0.121
2	0	0.028	0.034	0.067	0.088	0.132
3	0	0.025	0.033	0.072	0.086	0.135
4	0	0.018	0.024	0.070	0.092	0.126
5	0	0.015	0.023	0.066	0.079	0.125
6	0	0.015	0.023	0.068	0.081	0.124

Let the eye gaze correction coefficient be $p = 1 - f(\theta)$. The least squares method is used for fitting. $f(\theta)$ is the fitting function, and the corrected EAR is defined as

$$EAR' = \frac{EAR}{p}. \quad (9)$$

As shown in Fig. 5, the corrected normal condition is maintained at around 0.3. When the eyes are closed, it will rapidly decrease from 0.3 and then rapidly increase to near 0.3.

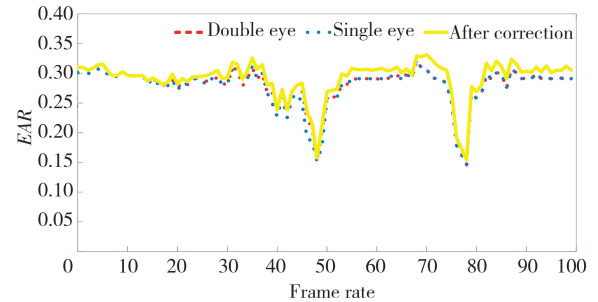


Fig. 5 Comparison of EAR before and after correction

When the human eyes are basically closed during fatigue, the value is close to 0, and the corrected second driver EAR is used as the threshold value.

2.2 Mouth feature extraction

The mouth aspect ration (MAR) is used to determine the mouth opening and closing degree as shown in Fig. 6.

$$MAR = \frac{\|P_{62} - P_{68}\| + \|P_{64} - P_{66}\|}{2\|P_{49} - P_{55}\|}. \quad (10)$$



Fig. 6 Mouth feature point coordinates

When the train driver is driving normally, the MAR is between 0 and 0.1. When yawning, MAR rises from near 0 to near 0.8 and then falls. 0.4 is used as the threshold value to determine whether the mouth is in a normal state and a fatigue state.

2.3 Head posture extraction

When fatigue determination of eyes and mouth is influenced by the environment, head attitude detection can be used to improve reliability. Head

pose is composed of 3 superimposed poses: pitch angle (α), roll angle (β), and yaw angle (γ).

The camera model can be represented as

$$s \begin{bmatrix} u \\ v \\ l \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}, \quad (11)$$

where s is the scale factor; u and v are 2D coordinates in the image coordinate system; f_x , f_y , c_x and c_y are the internal reference matrices of the camera; t_{ij} is the translation vector; and X , Y , and Z are 3D coordinates in the world coordinate system, all of which are known quantities; r_{ij} is the rotation matrix, which is the quantity to be solved. The rotation matrix in terms of Euler angles can be expressed as

$$R = R_x(\alpha)R_y(\beta)R_z(\gamma), \quad (12)$$

$$\text{where } R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{bmatrix}, R_y(\beta) = \begin{bmatrix} \cos \beta & 0 & -\sin \beta \\ 0 & 1 & 0 \\ \sin \beta & 0 & \cos \beta \end{bmatrix}, R_z(\gamma) = \begin{bmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

The rotation matrix is represented as a sine and cosine matrix composed of Euler angles, that is

$$R = \begin{bmatrix} \cos \beta \cos \gamma & \cos \beta \sin \gamma & -\sin \beta \\ \sin \alpha \sin \beta \cos \gamma - \cos \alpha \sin \gamma & \sin \alpha \sin \beta \sin \gamma + \cos \alpha \cos \gamma \cos \gamma & \sin \alpha \cos \beta \\ \cos \alpha \sin \beta \cos \gamma & \cos \alpha \sin \beta \sin \gamma & \cos \alpha \cos \beta \end{bmatrix}. \quad (13)$$

The head pitch angle (α), roll angle (β), and yaw angle (γ) can be expressed as

$$\begin{aligned} \alpha &= \arctan(r_{23}/r_{33}), \\ \beta &= -\arcsin(r_{13}), \\ \gamma &= \arctan(r_{21}/r_{11}). \end{aligned} \quad (14)$$

By counting the head posture of train drivers, it is found that the changes in roll and yaw angles are less pronounced when in a fatigued state, while the changes in pitch angle are the most pronounced with head nodding movements. A 20% change in pitch angle is used as a criterion for fatigue judgment, and the range of adult head pitch angle is from -60° to 70° [18]. That is to say, when the pitch angle exceeds -18° to 18° , it is judged as nodding and considered to be in a fatigue state at that moment. It is specified that the front-to-camera is 0° , and 0.2 is used as the threshold for determining head fatigue.

3 Fatigue feature fusion

Since fatigue is a gradual process and each person has a vague concept of fatigue, it is impossible to establish mathematical expressions through precise mathematical models. The fuzzy inference system is used as a fusion tool for fatigue detection features, and the driver's "experience" is converted into a control strategy through fuzzy inference, and the control strategy rules are shown in Table 3. The fuzzy inference system is established, taking the eye opening and closing degree, mouth opening and closing degree and head posture angle as input. The driver's state is divided into four levels: no fatigue, mild fatigue, moderate fatigue, and severe fatigue, and the four fatigue levels are taken as the output of the fuzzy inference system, as shown in Fig. 7, in which the affiliation function is chosen as the

triangular affiliation function.

Table 3 Fuzzy inference rules

	Eye state	Mouth state	Head posture	Fatigue level
1	open	open	nod	high
2	open	open	normal	mod
3	open	talk	nod	mild
4	open	talk	normal	none
5	open	closed	nod	mild
6	open	closed	normal	none
7	closed	open	nod	high
8	closed	open	normal	mod
9	closed	talk	nod	high
10	closed	talk	normal	mod
11	closed	closed	nod	high
12	closed	closed	normal	mod

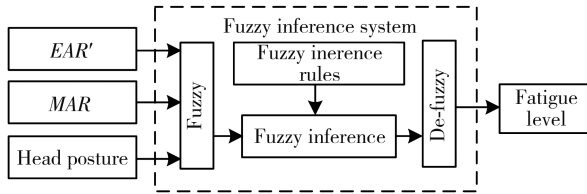


Fig. 7 Fuzzy inference process

The eye state fuzzy set is {open, closed} and the domain of the argument is $[0, 1]$. The threshold cut-off point is the data of the 2nd driver in Table 2. Its affiliation function is

$$\mu_e = \begin{cases} \mu_{\text{closed}}(x) = \begin{cases} 1, & 0 \leq x \leq 0.2, \\ -10x + 3, & 0.2 < x < 0.3, \end{cases} \\ \mu_{\text{open}}(x) = \begin{cases} \frac{20}{3}x - \frac{5}{3}, & 0.25 \leq x < 0.4, \\ 1, & x \geq 0.4. \end{cases} \end{cases} \quad (15)$$

The mouth state fuzzy set is {closed, talk, open} and the domain of the argument is $[0, 1]$. The threshold cut-off point is based on the P80 criterion of PERCLOS, and its affiliation function is

$$\mu_m = \begin{cases} \mu_{\text{closed}}(y) = -5y + 1, & 0 \leq y < 0.2, \\ \mu_{\text{talk}}(y) = \begin{cases} 10y, & 0 \leq y < 0.1, \\ -5y + 1.5, & 0.1 \leq y < 0.3, \end{cases} \\ \mu_{\text{open}}(y) = \begin{cases} 5y - 0.5, & 0.1 \leq y \leq 0.3, \\ 1, & y > 0.3. \end{cases} \end{cases} \quad (16)$$

The head pose fuzzy set is {nod, normal} and the domain of the argument is $[0, 1]$. Its affiliation function is

$$\mu_h = \begin{cases} \mu_{\text{normal}}(z) = \begin{cases} 1, & 0 \leq z < 0.2, \\ -5z + 2, & 0.2 \leq z < 0.4, \end{cases} \\ \mu_{\text{nod}}(z) = \begin{cases} 5z - 1.5, & 0.3 \leq z < 0.5, \\ 1, & 0.5 \leq z < 1. \end{cases} \end{cases} \quad (17)$$

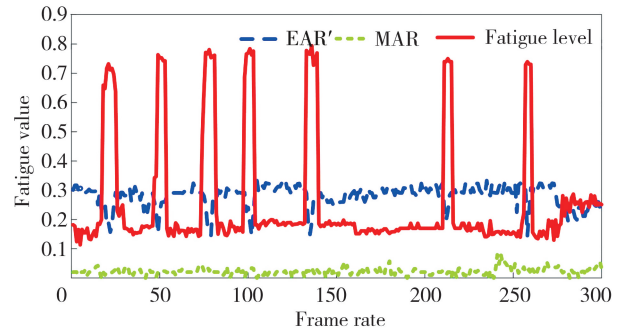
The fuzzy set of fatigue degree is {none, mild, moderate, high} and the domain of the argument is $[0, 1]$. Its affiliation function is

$$\mu_{\text{level}} = \begin{cases} \mu_{\text{none}} = -5l + 1, & 0 < l < 0.2, \\ \mu_{\text{mild}} = \begin{cases} 10l - 1.5, & 0.15 < l < 0.25, \\ -\frac{20}{3}l + \frac{8}{3}, & 0.25 < l < 0.4, \end{cases} \\ \mu_{\text{mod}} = \begin{cases} 5l - 1.5, & 0.3 < l < 0.5, \\ -10l - 4, & 0.5 < l < 0.6, \end{cases} \\ \mu_{\text{high}} = \frac{20}{9}l - \frac{11}{9}, & 0.55 < l < 1. \end{cases} \quad (18)$$

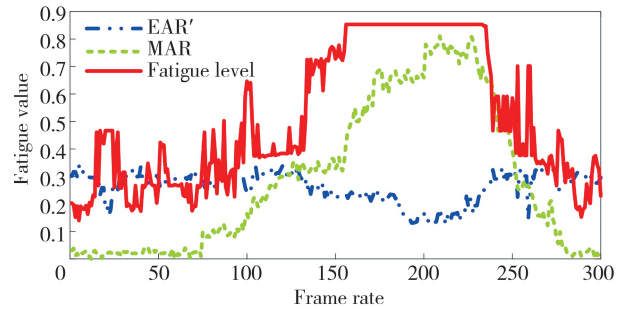
4 Result and discussion

In order to verify whether the proposed algorithm can achieve the expected results, data acquisition is performed in a laboratory environment by simulating a real driving environment with 300 frames of data as a group. Experiments and analysis are conducted to verify the feasibility of the algorithm.

Experiment 1: The eye mouth feature quantity is used to determine the driver's state when normal and fatigued, as shown in Fig. 8.



(a) Normal state



(b) Fatigue state

Fig. 8 Fatigue determination results of eye and mouth feature volume

The MAR value fluctuates around 0 in the normal state, indicating that the mouth is not open for yawning or talking. For EAR values, they remained around 0.3 except for blinks that decreased and

increased rapidly at certain intervals. The fatigue level becomes higher rapidly when blinking, and the fatigue level value is basically around 0.2 when not blinking, which is in the non-fatigue or mild fatigue range. From the overall point of view, non-fatigue accounts for 226 frames, mild fatigue accounts for 38 frames, moderate fatigue accounts for 2 frames, and severe fatigue accounts for 34 frames. The non-fatigue grade accounts for the most that is 75.3%, so the fatigue grade during the cycle is non-fatigue.

The MAR value increases from 0 to about 0.85 in the fatigue state, and is yawning, while yawning is accompanied by the decrease of EAR value, which is consistent with the decrease of eye opening and closing when yawning in life. From the overall point of view, non-fatigue accounts for 79 frames, mild fatigue accounts for 81 frames, moderate fatigue accounts for 44 frames, and severe fatigue accounts for 96 frames. Fatigue accounts for 73.7%, and severe fatigue accounts for the most that is 32%, so the fatigue level in this cycle is severe fatigue.

Experiment 2: The normal and fatigue states of the driver after adding the head posture are shown in Fig. 9.

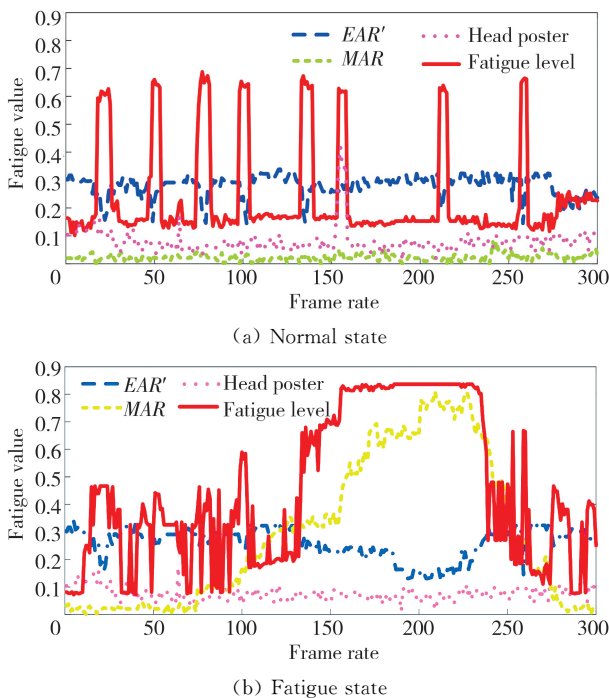


Fig. 9 Fatigue determination results by adding head posture

After adding the head posture, the fatigue degree value decreases at the blink. The head posture changes near the 150th frame in Fig. 9(a), causing the fatigue degree to increase. And the overall is still in the non-fatigue state, which indicates that the head posture in the normal state has less effect on the

fatigue degree. It is derived that the head posture in this cycle does not exceed the threshold value, which is in the normal state from Fig. 9(b). From the overall view, non-fatigue accounts for 72 frames, mild fatigue accounts for 70 frames, moderate fatigue accounts for 44 frames, and severe fatigue accounts for 104 frames. Fatigue accounts for 76%, and severe fatigue accounts for the most that is 34.7%, so the fatigue level in this cycle is judged as severe fatigue.

Experiment 3: Comparison of fatigue values after adding head posture is shown in Fig. 10.

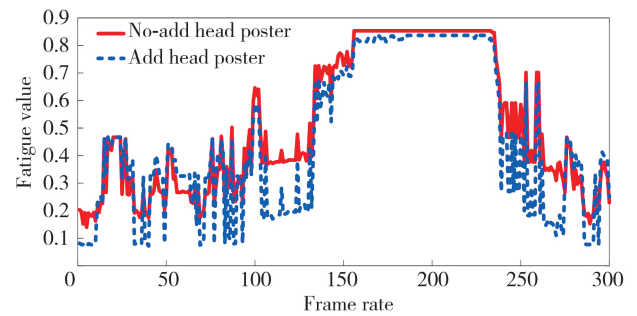


Fig. 10 Comparison before and after adding head posture

The before-and-after comparison of the head posture during fatigue is shown in Fig. 10. Compared with the time when the head posture is not added, the overall fatigue level curve is lower when the head posture is added, which has a slight effect on the overall curve because the head posture is in the normal range.

The comparison between the proposed method and the PERCLOS method is shown in Table 4. The PERCLOS method only determines the fatigue and non-fatigue states, while the proposed method divides the fatigue state into four levels to refine the fatigue degree and reduce the detection error.

Table 4 Comparison of fatigue test results

Fatigue level	PERCLOS method		Proposed method	
	Normal	Fatigue	Normal	Fatigue
None	286	48	226	72
Mild	—	—	38	70
Mod	—	—	2	44
High	—	—	34	104
Fatigue	14	252	—	—

In order to reduce the possible error caused by a single group of data, the accuracy of five groups of data is collected as the average accuracy. Table 5 shows the comparison of fatigue detection accuracy under normal environment, where excessive head deflection or driver not in the detectable range are judged as detection failure. Table 6 shows the

accuracy comparison between adding low-light enhancement processing and without adding low-light enhancement processing in the dark environment. With the addition of low-light enhancement, the image exposure is enhanced, which makes it easier to detect the driver's face and improve the accuracy of detection.

Table 5 Comparison of fatigue detection accuracy

Group	Accuracy/%
Proposed method	95
Eye and mouth fusion method	94.87
Ref. [16]	92

Table 6 Accuracy comparison in low light environment

Group	Accuracy/%
Adding low light enhancement	86.8
Without adding low light enhancement	62.4

By fusing three feature quantities of eyes, mouth and head posture angle for driver fatigue determination, it not only corrects the eye opening and closing degree of train drivers when they pay attention to various dashboard information, but also realizes fatigue classification, improves the sensitivity of detection and achieves accuracy rates of 95% and 86.8% in the dark.

5 Conclusions

A multi-information fusion fatigue detection method is proposed to detect train driver fatigue under low light.

1) Low-light enhancement processing of driver images based on low-light environment, detection and localization of driver face images by using LBP and ERT algorithms are progressed.

2) Adaptive threshold method and eye gaze correction are used to reduce the detection error due to individual differences.

3) A fuzzy inference system with three feature quantities of eyes, mouth and head posture angle as input and driver fatigue level as output is fused to determine the driver fatigue status.

References

- [1] YANG H, CHEN X Q, LEI J Y, et al. A fatigue detection method for train drivers based on inverse projection correction and eye gaze correction. *Journal of the China Railway Society*, 2018, 40(4): 83-89.
- [2] WANG F, WU S C, PING J Y, et al. EEG driving fatigue detection with PDC-based brain functional network. *IEEE Sensors Journal*, 2021, 21(9): 10811-10823.
- [3] CAI S X, DU C K, ZHOU S Y, et al. Fatigue driving state detection based on vehicle running data. *Journal of Transportation Systems Engineering and Information Technology*, 2020, 20(4): 77-82.
- [4] LI K N, GONG Y B, REN Z L. A fatigue driving detection algorithm based on facial multi-feature fusion. *IEEE Access*, 2020(8): 101244-101259.
- [5] FANG B, XU S, FENG X F. Self-calibration driver fatigue detection for advanced RISC machine(ARM) platform. *Journal Automotive Safety and Energy*, 2020, 11(1): 71-78.
- [6] LIU Y Y, ZHU Y, SUN J B, et al. Haar-like feature dual-threshold Adaboost face detection. *Chinese Journal of Graphics*, 2020, 25(8): 1618-1626.
- [7] SONG Y, LU L Y. A fatigue warning algorithm based on spatiotemporal feature extraction of facial motion. *Computer Engineering & Science*, 2019, 41(10): 1763-1770.
- [8] JI Y Y, WANG S G, ZHAO Y, et al. Fatigue state detection based on multi-index fusion and state recognition network. *IEEE Access*, 2019(7): 64136-64147.
- [9] LI X, HONG L, WANG J C, et al. Fatigue driving detection model based on multi-feature fusion and semi-supervised active learning. *IET Intelligent Transport Systems*, 2019, 13(9): 1401-1409.
- [10] ZHAO X P, MENG C N, Mingkui FENG M K. Fatigue detection based on cascade convolutional neural network. *Journal of Optoelectronics Laser*, 2017, 28(5): 497-502.
- [11] PAN Z G, LIU R F, ZHANG M M. Research on fatigue detection algorithm based on fuzzy comprehensive evaluation. *Journal of Software*, 2019, 30(10): 2954-2963.
- [12] SIMA Y, YI J Z, CHEN A B, et al. Localization and recognition of expression-complete frames in dynamic face image sequences. *Journal of Applied Sciences*, 2021, 39(3): 357-366.
- [13] WU M J, MU P G, ZHANG C Y. Driver fatigue detection algorithm based on the states of eyes and mouth. *Computer Application and Software*, 2013, 30(3): 25-27, 54.
- [14] QIANG L. Research on driving fatigue detection of train drivers based on PERCLOS. Beijing: Beijing Jiaotong University, 2014: 61-68.
- [15] ZHENG J W, LIU X M, YIN J L. Pain expression recognition based on LBP and SVM. *Computer System Applications*, 2021, 30(4): 111-117.
- [16] KAZEMI V, SULLIVAN J. One millisecond face alignment with an ensemble of regression trees//2014 IEEE Conference on Computer Vision and Pattern Recognition, Jun. 24-29, 2014, Columbus, USA. New York: IEEE, 2014: 1867-1874.
- [17] SOUKUPOVÁ T. Real-time eye blink detection using facial landmarks. Czech: Czech Technical University, 2016.
- [18] WU K L. Fatigue driving detection based on facial features and head posture. Shanghai: Donghua University, 2020.

基于面部多信息融合的列车司机疲劳检测

郝正清¹, 王 英^{1,2}, 陈小强^{1,2}, 熊 烨¹

(1. 兰州交通大学 自动化与电气工程学院, 甘肃 兰州 730070;

2. 兰州交通大学 光电技术与智能控制教育部重点实验室, 甘肃 兰州 730070)

摘 要: 为了提高列车司机行车疲劳检测的准确性, 提出一种面部多信息融合的列车司机疲劳检测方法。首先, 采用低光增强进行图像预处理, 以基于局部二值模式(LBP)特征检测人脸。其次, 利用回归树集合(ERT)算法获取司机人脸特征点, 并通过人脸模型匹配得到司机头部姿态角。最后, 根据列车司机的特殊驾驶环境, 对最能表现疲劳的眼睛特征量进行自适应阈值修正和眼睛凝视修正, 以模糊推理系统为融合工具, 以眼睛、嘴巴、头部姿态特征量作为模糊推理系统输入, 得到司机的疲劳值作为检测结果。实验结果表明, 该检测方法可以区分司机疲劳等级, 在正常环境下准确率为 95%, 在低光照环境下的准确率为 86.8%。

关键词: 列车司机; 疲劳检测; 特征点检测; 头部姿态; 面部多信息融合

引用格式: HAO Zhengqing, WANG Ying, CHEN Xiaoqiang, et al. Train driver fatigue detection based on facial multi-information fusion. Journal of Measurement Science and Instrumentation, 2023, 14 (4): 493-500. DOI: 10.3969/j.issn.1674-8042.2023.04.012