Driver Fatigue State Detection Based on Facial Key Points

Nana Zhang
Information Technology Institute, Shanghai Jian Qiao University
SHANGHAI, P.R. CHINA

Hui Zhang
Information Institute, Shanghai Ocean University
SHANGHAI, P.R. CHINA

Jun Huang
Information Institute, Shanghai Ocean University
SHANGHAI, P.R. CHINA

Abstract—**The proportion of drivers causing traffic accidents due to fatigue driving has been increasing year by year, which has become one of the main causes of traffic accidents. Therefore, accurate and effective detection of driver fatigue is a hot research topic at present. In this article, the driver's head posture estimation based on face key points is used to judge fatigue. Firstly, the face image is collected from the camera in real time. The model-based method is used to estimate the head posture of the person. The face key points method based on the Dlib library is used to judge the eye state、the mouth state、the head turning posture of the person. The experimental results show that the head pose estimation method based on face key points can accurately judge the fatigue state, and has good real-time and accuracy.**

Keywords- Head posture; Fatigue driving; Face key points; Model method

I. INTRODUCTION

Fatigue driving is defined as: The driver drives down for a long time or when the body is exhausted, resulting in an imbalance of physiological functions and mental functions, thereby causing a decline in driving skills objectively. Traffic accident statistics in Europe and the United States show that 80% to 90% of traffic accidents are caused by human factors. On U.S. roads, more than 100,000 traffic accidents are caused because of driver fatigue every year, and more than 7 million of people are directly injured or killed[1]. Fatigue drives account for about a quarter of all traffic accidents in Germany that result in fatalities. French accident statistics show that fatigue accounts for about 15 percent of all injuries and 21 percent of deaths[2]. China's Ministry of Communications has published statistics on traffic accidents, which shows that traffic accidents caused by fatigue driving account for about 20% of the total number of accidents, more than 40 percent of major traffic accidents and 83 percent of traffic deaths[3]. The research of Klauuer et al. shows that the probability of traffic accidents caused by fatigue driving is about 4-6 times that of normal driving[4].

Many countries in the world attach great importance to the development and application prospect of fatigue driving detection in traffic accident prediction. Researchers at home and abroad have also started research in this field. Fuwamoto Y and Makikawa M [5] monitor the driver's ECG signal based on various physical sensors through the built-in capacitor electrode in the driver's seat; Wathiq and Ambudkar[6] used the feedforward neural network to detect eye position, head position, and mouth position to confirm the driver's distraction; Baheti et al [7] used a deep convolutional neural network to detect driver dispersion behavior and improved the VGG-16 model. In China, Jianping Li [8] realized the accurate positioning of human eyes based on the AdaBoost algorithm, using Harris strong corner points to detect the central area of the human eye, to obtain the eye line state information, according to the fatigue judgment model to achieve the driver's fatigue state classification warning; Liu et al[9]. proposed an algorithm of Laplace support vector machine to detect the eyes and head of the driver in the semi-supervised mode to evaluate the driver's driving behavior. The above experiments have been carried out by developing hardware or improved algorithms to judge the fatigue of the driver from different angles and different behaviors of fatigue driving, and all experiments have achieved good results. Most of the above test methods were carried out for the driver's eye features, and no research experiments were conducted for the state of yawning and bowing when tired.

In this paper, a model-based method is used to estimate the driver's head posture. The model-based method mainly estimates the head by constructing a geometric model of the head shape by detecting key contour points such as face contour, eye corner, nose tip and lip[10-11]. When the Angle of the user's bow is more than 30 degrees, it is judged that the driver is in a doze state and is prompted. On this basis, the face feature point extraction method based on Dlib library is used to calculate the Euclidean distance between the upper and lower eyelids. When the Euclidean distance is less than a certain value, it is judged as the closed eye state. Similarly, when the mouth state reaches a certain length of time, it is judged to be in a state of yawning. Thus, fatigue judgment is performed by closed eye detection, yawn detection, and low head detection. Figure 1 is a system flow chart.

978-1-7281-5255-1 ©2019 IEEE

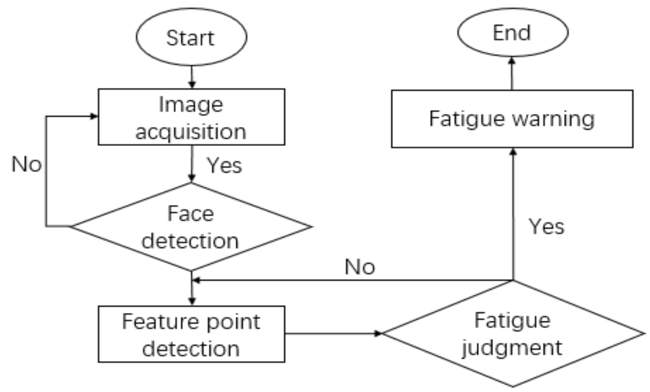


Figure. 1. System flow chart

II. HEAD POSE ESTIMATION BASED ON FACE KEY POINTS

A. Face Image Processing

The method used in the experiment is based on the open source computer vision library OpenCV. The OpenCV library encapsulates a class method for reading file video and camera video. In the class function cv::VideoCapture cap(int device) , you can call the camera or open the video file by passing in the corresponding parameters. In the experiment, the parameter "1" is passed in, and the camera is called to read the video streams.

B. Head Pose Estimation

The head pose estimation refers to the computer determining the position and posture parameters of the head in the three-dimensional space of the person by analyzing and predicting the input image or the video sequence. The attitude parameters include horizontal rotation (yaw), vertical rotation (pitch), and in-plane rotation (roll) [6], which are called the Euler angles..

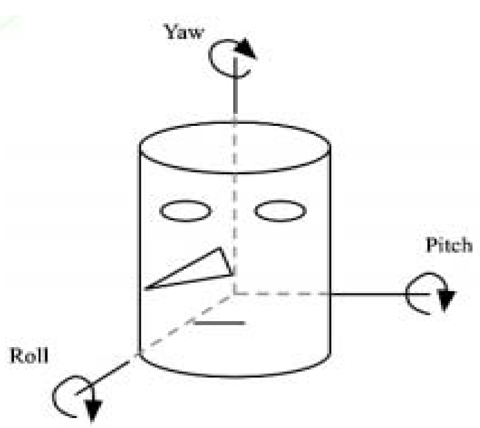


Figure 2 Head posture parameter Map

The steps of the head pose estimation algorithm are: 2D face key point detection; 3D face model matching; solving the conversion relationship between the 3D point and the corresponding 2D point; and solving the Euler angle according to the rotation matrix. We took the tip of the left and right eyebrows, the left and right canthus of the left and right eyes, the tip of the nose, and the left and right corners of the mouth as 2D and 3D positions in the reference frame.

The pose of the object relative to the camera can be represented using a rotation matrix R and a translation matrix T. The 3D coordinates of the facial features are displayed in world coordinates （U,V,W）, which are converted into camera coordinates （X,Y,Z） by translation and rotation. The coordinate points are projected into the image coordinates（x,y）by parameters inherent to the camera (eg. focal length, optical center, etc.). The formula for converting world coordinates to camera coordinates is as follows:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = R \begin{pmatrix} U \\ V \\ W \end{pmatrix} + T = [R|T] \begin{pmatrix} U \\ V \\ W \\ 1 \end{pmatrix} \tag{1}$$

The rotation matrix R is a 3x3 matrix, and the translation matrix T is a 3x1 vector. The camera coordinates are projected to the image coordinates as follows:

$$s \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \tag{2}$$

where f_x and f_y are focal lengths in the x and y directions, and c_x and c_y are optical centers. Finally, the direct linear transformation method (DLT) is used to solve the above equation to obtain the rotation matrix.

The Euler angles θ_x, θ_y and θ_z are rotated around the x-axis, the y-axis, and the z-axis to obtain a rotation matrixR_x, R_y, R_z

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{pmatrix} \tag{3}$$

$$R_y = \begin{pmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{pmatrix} \tag{4}$$

$$R_z = \begin{pmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5}$$

R_x, R_y and $R_z$are multiplied to obtain R, from which we get the Euler angle. When the user's head is rotated relative to the camera, the Euler angles change, as shown in Figures 3 and 4.

Figure 3 Change of y value when the head is turned right

Figure 4 Change of z value when the head is tilted

C. Face Key Point Detection

Like OpenCV, Dlib is an open source library, it has an image processing toolbox, and supports machine learning training functions such as deep learning, which can be used to detect objects in images, including frontal detection and target pose estimation. Dlib provides a face-critical 68-point training tool and marks 68 points that represent key features of the face. Through these 68 points, face alignment and face state estimation can be completed. The point identification and serial number of the face key point are shown in Figure 5, from which the position of each key point in the face is known.

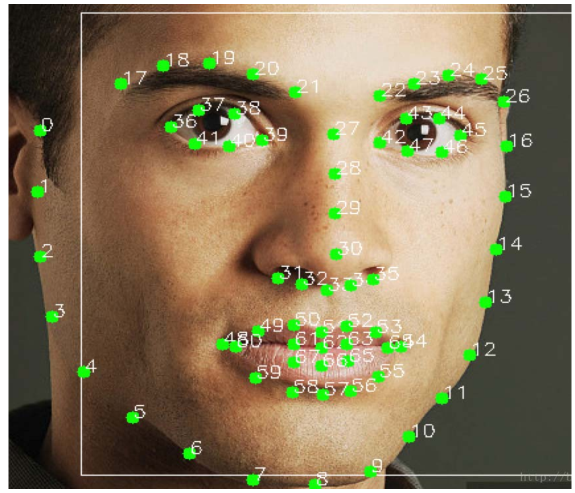


Figure 5 face key points identification sequence diagram

III. FATIGUE JUDGMENT

When people feels tired, the body will subconsciously do something to signal, such as yawning and closing the eyes for an extended period of time. When people starts to doze off, the head will consciously bow down. This experiment tests the above three behaviors to determine whether the user is fatigued.

A. Closed Eye Detection

In general, the distance between the upper and lower feature points of the eye area is larger when the eyes are opened and smaller when the eyes are closed. The EAR value is obtained by calculating the height and width of the eye using the distance of the eye feature points, and calculating the aspect ratio of the eye. It is used to judge the blink of an eye[12]

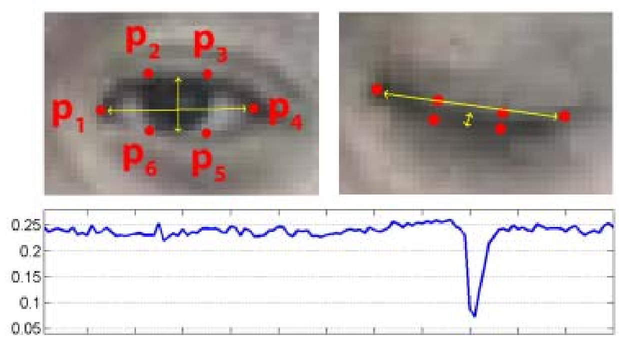


Figure 6 EAR value Curve

The EAR value is calculated as follows:

$$EAR = \frac{\|P_2 - P_6\| + \|P_3 - P_5\|}{2\|P_1 - P_4\|} \quad (6)$$

where, P_1, P_4 are the corners of the eyes, P_2, P_3, P_5, P_6 are the upper and lower eyelids. Corresponding to the 68 feature points of the face, the points in [36, 47] are the feature points of the eye, wherein the points 36, 39, 42, 45 are the corners of the eye, and the remaining points are the key points of the upper eyelid and the lower eyelid. The Euclidean distance is calculated as:

$$Distance(a,b) = \|P_a - P_b\| = \sqrt{(P_a.x - P_b.x)^2 + (P_a.y - P_b.y)^2} \quad (7)$$

where $P_a.x$ is the x coordinate of point a; $P_a.y$is the y coordinate of point a. When the EAR value is below a certain threshold EYE_AR_THRESH, we determine that the state of the eye is in the closed eye state. The threshold is set to 0.3.

B. Yawn Detection

The yawn test is similar to the eye-closing test. The distance between the upper and lower feature points in the mouth area is larger when the mouth is open and smaller when the mouth is shut. The height and width of the outer contour and the inner contour of the mouth are calculated by using the distance of the feature points of the mouth. Corresponding to the 68 feature points of the face, the points in [48, 67] are the feature points of the mouth, wherein the points 48, 54 are the corners of the mouth of the mouth, and the points 60, 64 are the corners of the mouth of the mouth, point 51, 57 is the highest point and the lowest point of the outer contour of the mouth, and points 62, 66 are the highest point and the lowest point of the inner contour of the mouth are the key points of the upper eyelid and the lower eyelid.

$$MOUTH = \left(\frac{\|P_{51} - P_{57}\|}{\|P_{48} - P_{54}\|} + \frac{\|P_{62} - P_{66}\|}{\|P_{60} - P_{64}\|}\right)/2 \quad (8)$$

When the MOUTH value is greater than a certain threshold, we determine that the mouth state is the open mouth state and the threshold is set to 0.3.

Generally, when yawning, the user will open his mouth and the eye will have a closed eye trend. We judge that when the mouth is open for more than 3 seconds and the eye state is half blink or closed eye, it is judged to be yawning..

C. Head Detection

Using the head rotation algorithm in 2.2, we obtain x, y, z, values, where x represents pitch, y represents yaw, and z represents roll. When the Angle of the head is greater than 30 degrees, the value of x is greater than 10, and the head is in the bowed state.

The following table shows the user's state judgment under different actions:

TABLE I. FATIGUE STATE DIAGRAM

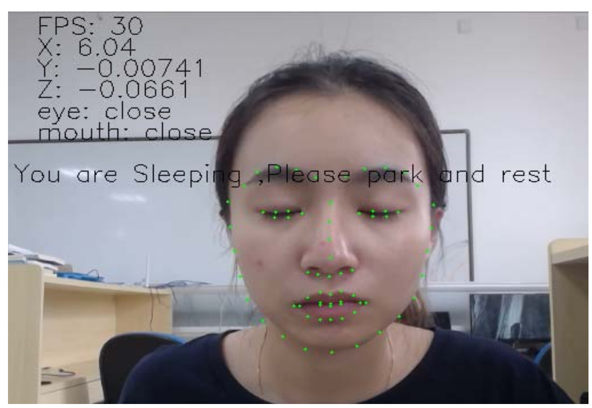
State	Full open	Half open	Closed
Open mouth	Normal	Fatigue	Fatigue
Closed mouth	Normal	Doze	Doze
Bow	Doze	Doze	Doze

IV. EXPERIMENTAL RESULTS AND ANALYSIS

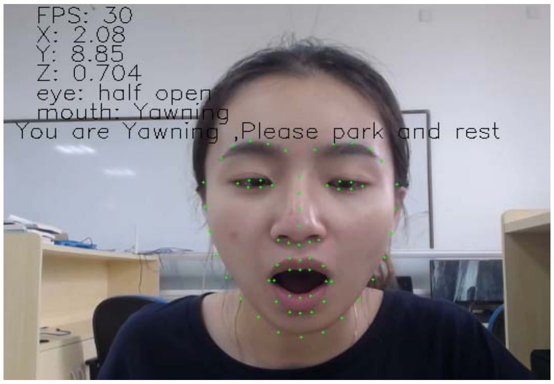
The platform of this experiment is Windows10, the processor is Inter(R) Core(TM) i5-3230m, the main frequency of CPU is 2.6ghz, and the memory is 8G. This experiment uses the C++ programming language, and uses Visual Studio and OpenCV image processing library as the development environment. It uses USB external camera to capture image data. The experimental environment is the laboratory. The testers make a series of actions to simulate the fatigue behavior of the driver. Then the image is captured by the camera for fatigue judgment.

The eye state is judged when people's eyes are fully open, half open or closed. When the eyes are closed, the mouth is yawned or the head is lowered for a certain time, the user is judged to be tired and prompted. Experimental results are shown in figure 7. Figure a is the normal state diagram of the head for horizontal head-up, opening eyes and closing mouth. B is a figure for detecting whether the user closes his eyes or not. If the user closes his eyes for 3 seconds, the user is judged to be asleep and prompted. C is the user's mouth opening detection. If the user's mouth opening time reaches 3 seconds and the eye state is half open, the user is judged to be yawned and prompted. D is an effect diagram for detecting whether the user is bowing. When people is dozing off, the head will subconsciously descend. If the Angle of bending down is more than 30 degrees,, the user's state is judged to be doze and prompt..

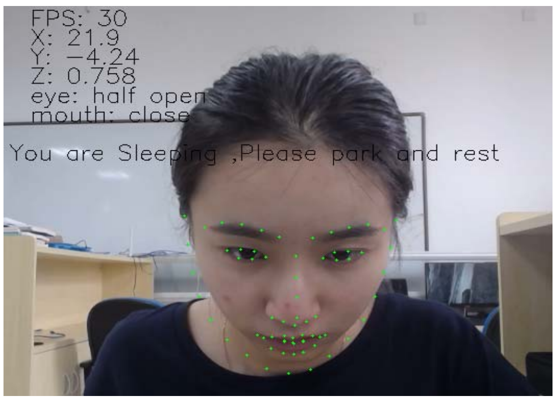
(a) The effect of the normal person



(b) Closed eyes detection



(c)Yawning detection



(d) Head Detection

Figure. 7. Experimental results

In the experiment, testers were invited to simulate the driver's fatigue behavior, and they made fully open, half open and closed eyes in the state of mouth opening, mouth closing and head bowing respectively. According to the state in table 1, the experiment was conducted, The data is shown in table 2:

TABLE II. SYSTEM TEST DATA

Test content	Number of tests	Number of passes	Accuracy
Normal	200	194	0.97
Fatigue	250	250	1.0
Doze	250	232	0.928

For other closed-eye detection algorithms, the testers are all tested in the same state. The data obtained are shown in Table 3:

TABLE III. COMPARISON TEST DATA

Algorithm	Number of tests	Number of passes	Accuracy
Document [13]	250	221	0.885
Document [14]	250	230	0.92
Document [15]	250	237	0.95
This article	250	242	0.97

The experimental results show that the head pose estimation method based on face key points can accurately judge the fatigue state and has better accuracy.

V. CONCLUSION

The driver fatigue driving detection system can reduce the occurrence of many accidents and has great social value. In this paper, the head pose estimation method based on face key points is used for real-time measurement. The yawn detection, closed-eye detection and low-head detection are used to judge whether the user has fatigue tendency. When the user closes eyes, yawns or bows, it will warn . Experiments show that the method has good real-time and accuracy.

ACKNOWLEDGMENT

This work was supported by the Shanghai Municipal Education Commission's "Morning Plan" project (NO. AASH1702).

REFERENCES

[1] Liu C,Subramanian R.Factors related to fatal singlevehicle run off road crashes[R]. DOTHS811232,U.S. Department of Transportation, American National Highway Traffic Safety Administration. Washington:DC,USA,2009.

[2] Li Duhou,Liu Qun, Yuan Wei, et al. Relationship between fatigue driving and traffic accidents.Journal of Traffic and Transportation Engineering, 2010, 10(02): 104-109.

[3] Niu Qingning, Zhou Zhiqiang, Jin Lisheng, et al. Fatigue driving detection method based on eye movement feature. Journal of Harbin Engineering University, 2015, 36(3): 394 - 398

[4] Klauer S G,Dingus T A, Neale V L, et al. The impact of driver inattention on near crash/crash risk:An analysis using the 100 car naturalistic driving study data. Washington: National Highway Traffic Safety Administration,2006.

[5] Fuwamoto Y, Makikawa M, Ito Y, et al. IN-VEHICLE ELECTROCARDIOGRAPH DEVICE AND VEHICLE: US, US20100049068. 2010.

[6] Wathiq O, Ambudkar BD. Driver safety approach using efficient image processing algorithms for driver distraction detection and alerting [M] // Intelligent Engineering InformaSStics. Springer, 2018: 461-469.

[7] Baheti B, Gajre S, Talbar S. Detection of distracted driver using convolutional neural network // IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018: 1145-1151

[8] Li Jianping, Niu Yanxiong, Yang Lu et al. Non-contact fatigue driving monitoring and early warning system based on human eye state information. Advances in Laser and Optoelectronics, 2015, 52(4): 83-88.

[9] Liu TC, Yang Y, Huang GB, et al. Driver distraction detection using semi-supervised machine learning . IEEE Intelligent Transportation Systems Magazine, 2016, 17(4): 1108-1120.

[10] Ahn B, Choi D G, Park J, et al. Real-time head pose estimation using multi-task deep neural network. Robotics and Autonomous Systems, 2018, 01: 1-12.

[11] Tang YQ, Sun Z N, Tan T N, A Survey on Head Pose Estimation. Pattern Recognition and Artificial Intelligence, 2014, 27(3):213-225.

[12] Soukupova T, Cech J. Eye blink detection using facial landmarks[C]//21st Computer Vision Winter Workshop, Rimske Toplice, Slovenia. 2016.

[13] CUIQING ZHANG, LIZHEN WEI, PEI ZHENG. Research on Driving Fatigue Detection Based on PERCLOS. 2017 4th International Conference on Vehicle, Mechanical and Electrical Engineering (ICVMEE 2017

[14] Tang Jie, Chen Renwen, Yu Xiaoqing. Implementation of Vehicle Fatigue Driving Detection System Based on Convolutional Neural Network. Foreign Electronic Measurement Technology, 2018, 37(1): 116-121

[15] Pan Zhigeng, Liu Rongfei, Zhang Mingmin. Fatigue driving detection algorithm based on fuzzy comprehensive evaluation.Journal of Software 2019: 1-12