Research on Night Driving Fatigue Detection Based on Infrared Active Vision

Cai-ping CHAI1,*, Li-Min GUAN2 and Hui-feng WANG2

1School of Highway & Railway engineering, Shaanxi College of Communication Technology, Xi’an 710018, Shaanxi, China
2School of Electronic &Control Engineering Chang’an University, Xi’an 710064, Shaanxi, China

*Corresponding author

Keywords: Intelligent transportation, Night driving fatigue, Infrared active illumination, Eye localization, KLT tracking, Gabor transform, Parabolic fit.

Abstract. Night is the period of fatigue for drivers and of accidents happened. At night, Ambient lighting is weak, and the full-in light of visible light influence the normal driving. According to these conditions, a infrared active vision fatigue detection design for night driving is proposed. Firstly, the 940nm light source with no red is adopted to perform fill-in light, by which imaging quality is developed and the interference of ambient light is prevented. Secondly, the existing fatigue detection algorithm is improved. Face location and tracing is realized with ST_Adaboost and KLT tracing. Eyes are located by gray-level integration method for Gabor transformation, which increases the accuracy of eye positioning. Finally, the fatigue is judged by PERCLOS algorithm. The experiment results show that the proposed algorithm can detect fatigue driving around the clock, and the accuracy of eye positioning is more than 90%, the accuracy of eye state recognition is more than 87%. The proposed method can resolve the problem about driving fatigue at night, so there is great significance for lowering the risk of traffic accidents especially at night.

Introduction

Fatigue driving[1] has become a serious problem that causes traffic accidents and threatens human life. According to statistics, 40% of the major traffic accidents are caused by fatigue driving, especially for large and medium-sized vans. The number of people carrying and large amount of cargo is the key and difficult for traffic accident prevention, most of them are driving at night, it is the frequent stage of driver fatigue[2]. Therefore, nighttime fatigue driving[3] research is very important.

Due to the weak and unstable illumination of the driving environment at night, the collected driver image has low signal-to-noise ratio and poor data consistency, and the fatigue detection based on eye recognition is very sensitive to the image quality. The literature[4] combines the advantages of wavelet “digital microscope” with Mean Shift’s strong non-parametric probability density estimation and fast template matching to remove unknown noise in night-time images; some identification systems have adopted the method of visible or near-infrared active illumination[5], but at the same time introduced a new problem, that is, the illumination beam cannot uniformly illuminate[6] the recognition objec.

The fatigue warning technology based on the eye state is generally a problem that the human eye positioning takes a long time and tracking has poor robustness. Literature[7] proposed a CSDS detection method, which introduces the verification and confirmation mechanism to effectively control the occurrence of false detection under the condition of ensuring the detection rate, and improves the training speed while ensuring performance; Zhang, Z. utilization Nonlinear unscented Kalman filtering for eye tracking[8]. These improved algorithms can reduce the time of human eye positioning and improve the performance of human eye tracking to some extent.

In view of the shortcomings of nighttime fatigue warning research and practical application requirements, this paper improves the following aspects of fatigue driving detection: the infrared active uniform illumination model is analyzed, and the uniform fill light system is designed to solve
the problem of the poor image quality; The Adaboost algorithm is used for face localization, and the matching size threshold of the human eye image is artificially set by the actual experiment; Using gray level integral of Gabor transform Eye positioning reduces the influence of illumination changes on the detection results and improves the robustness of the system.

Active Infrared Uniform Illumination and Image Acquisition

Infrared Uniform Illumination Field Model

To construct a uniform illumination model using an infrared LED array, it is first necessary to determine a single infrared LED illumination model. A single infrared LED has a small illumination size and can be approximated as a point source relative to its illumination distance. The ideal radiation distribution function of the LED point source \( E(s, \theta) = E_0(s) \cos^m \theta \) is

\[
E(s, \theta) = E_0(s) \cos^m \theta
\]

Among them, \( E(s, \theta) \) is target irradiance, \( s \) is the distance between the infrared LED and the target plane, \( \theta \) is the angle between the light and the optical axis, \( m \) is the value related to the half-life angle \( \theta_{1/2} \). Half-life angle \( \theta_{1/2} \) is the angle between the light and the optical axis when the radiation intensity is half of the radiation intensity in the direction \( \theta = 0 \), which is an inherent characteristic of the LED. The \( m \) of the general LED is greater than 1, (Lambertian \( m \approx 1 \)).

Due to the non-correlation of the infrared LED source, the illumination of a point in the target area is a superposition of all infrared LED illumination. Then the irradiance distribution on the target surface can be expressed as

\[
E(x, y, z) = \sum_{i=1}^{N_i} \sum_{n=1}^{N_{in}} \frac{1}{(x-x_{in})^2 + (y-y_{in})^2 + (R \cos \alpha_i - z)^2 + R^2 - (x^2 + y^2 + z^2)^{m/2}} (2R)^{m/2}
\]

Among them \( N_i \) is the number of LEDs per ring, \( x, y \) is the position coordinate of the calculated LED, and \( \alpha_i \) is the angle between the normal to the infrared LED of the \( i \) ring and the \( z \) axis. The irradiance distribution of the entire region can be obtained by the formula (2).

Uniform Fill Light Illumination System Design

The infrared light active illumination image acquisition system is shown in Fig.1. Mainly include: infrared \(^9\) LED array light source, diffuser plate, low illumination camera, band pass narrow band filter and corresponding drive circuit.

As shown in Fig. 1(a), 1 is a scattering plate (annular); 2 is a 940 nm band pass narrowband filter; 3 is an infrared LED array (three-ring dome) light source; 4 is a low illumination camera; 5 is the target Lighting area. The uniform diffuser and the infrared LED array light source form an active lighting system to supplement the lighting of the cab. The whole system is an integrated structure. The physical object is shown in Fig.1(b).
The infrared light active illumination image acquisition device has the following characteristics:

1. A 940 nm near-infrared light without red exposure to fill light is used, the driver's stimulation is small, and does not affect its normal driving behavior. Moreover, the spectrum, street lamps, sunlight, and other vehicles have small overlapping area of lights, and it is easy to shield other lights.

2. A plurality of infrared LEDs form a three-ring dome-shaped LED array, enclosing the camera in the middle, focusing the light on the center of the target illumination area, expanding the illumination range and increasing the illumination intensity.

Human Eye Tracking and Positioning

Face Localization Based on ST_Adaboost Algorithm

The traditional Adaboost algorithm, when using the strong classifier generated by training to perform face target location detection, first needs to perform image matching normalization processing to adjust the size of the face target image to complete the matching of the human eye. The system installation diagram is shown in Fig. 2. In this study, the Adaboost algorithm for artificially setting the target image matching size threshold is named Size Threshold Adaboost algorithm, namely ST_Adaboost algorithm.

The simulation experiments were carried out with 10 test videos (66 762 frames) loaded in the laboratory and in the field. The effects of human eye detection before and after setting the matching size threshold were simulated and compared. The statistical test results are shown in Tab.1

<table>
<thead>
<tr>
<th></th>
<th>Detection rate/%</th>
<th>False detection Rate/%</th>
<th>Detecting average time/ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional Adaboost</td>
<td>91</td>
<td>5</td>
<td>485.5</td>
</tr>
<tr>
<td>ST_Adaboost</td>
<td>95</td>
<td>3</td>
<td>332.3</td>
</tr>
</tbody>
</table>

It can be found from the statistical results that the ST_Adaboost algorithm that sets the matching size threshold is superior to the traditional Adaboost algorithm in both the detection accuracy and the detection speed, which can significantly improve the detection performance of the system.
KLT Face Tracking Algorithm Based on Harris Corner

In order to improve the face positioning speed, after the face is located by the ST_Adaboost algorithm, the face window is tracked based on the Harris corner KLT tracking algorithm.

The Harris operator detects the corner points through the differentiation operation and the autocorrelation matrix. The steps for solving the Harris corner of the face window to be tracked are as follows:

1) Calculate the correlation matrix $M$ of each pixel in the face window:

$$M = \begin{bmatrix}
G(x, y) \otimes p_x^2 & G(x, y) \otimes p_x p_y \\
G(x, y) \otimes p_x p_y & G(x, y) \otimes p_y^2
\end{bmatrix} = \begin{bmatrix} A & C \\
D & B
\end{bmatrix}$$

(3)

Among them, $G(x, y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$, $P_x$ and $P_y$ are the gradients of the corresponding pixel points in the $x$ and $y$ directions, respectively.

2) Calculate the Harris corner response of each pixel in the face window

$$R = \frac{AB - CD}{A + B}$$

(4)

3) Look for the maximum point in the face window. If the Harris corner response $R$ value is greater than the threshold, it is treated as a corner point. By solving the Harris corner of the face window, a series of tracking points of the KLT face tracking algorithm are obtained.

Let the face area image at time $t$ be $P(x, y)$, and the face area image corresponding to time $t + \Delta t$ be $Q(x, y)$. From time $t$ to time $t + \Delta t$, the displacement of the feature point $X(x, y)$ is $d = (\Delta x, \Delta y)$. It is assumed that the face area in the image is shifted only when compared with the previous frame, and the brightness is unchanged. Then the noise is

$$n(X) = Q(X + d) - P(X)$$

(5)

The tracking of the face area is to obtain the displacement $d$ such that $n(X)$ is minimum. Define the error amount $\varepsilon$ as:

$$\varepsilon = \sum_w [n(X)]^2 = \sum_w [Q(X + d) - P(X)]^2$$

(6)

The $Q(X + d)$ Taylor is expanded and substituted into the formula (6). When the differentiation of $\varepsilon$ is 0, $\varepsilon$ takes the minimum value to obtain a formula of the form $Zd = \varepsilon$. Among them

$$Z = \sum_w \begin{bmatrix} Q_x^2 & Q_x Q_y \\
Q_x Q_y & Q_y^2
\end{bmatrix}$$

$$\varepsilon = \sum_w Q_x [Q_x, Q_y]^T$$

(7)

$$Q_x$$ and $Q_y$ are the derivatives of $Q(x, y)$ in the $x$ and $y$ directions, respectively, and $Q_x$ is the difference between the pixels of the two frames of images, and $d = Z^{-1}\varepsilon$.

Assume that in the video, $P$ and $Q$ are the face regions in the two consecutive frames, and select $W$ as a small window of 3x3. Then, feature point $X(x, y)$ is used as the tracking point, and 8 points around it are calculated points. The difference in brightness is

$$l_i = P_i - Q_i, \ i = 1, 2, \cdots, 8$$

(8)

The sum of horizontal gradient and vertical gradient is
\( g_i = P_i + Q_i, \quad g_{ni} = P_{ni} + Q_{ni}, \quad i = 1, 2, \cdots, 8 \) 

Therefore

\[
Z = \begin{bmatrix}
\sum_{i=1}^{8} g_i^2 & \sum_{i=1}^{8} g_i g_{ni}
\end{bmatrix} = \begin{bmatrix} a & c \\ c & b \end{bmatrix} e = \begin{bmatrix}
\sum_{i=1}^{8} l_i g_i
\end{bmatrix} = \begin{bmatrix} e_x \\ e_y \end{bmatrix}
\]

Get the amount of displacement

\[
d = Z^{-1} e = \frac{1}{ab - c^2} \begin{bmatrix} be_y - ce_x \\ ae_y - ce_x \end{bmatrix} = \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}
\]

After obtaining the displacement of all the tracking points, the displacement of the face region between the two frames can be determined, and the tracking of the face region in the image is realized.

**Gray-integrated Human Eye Localization Combined with Gabor Transform**

After locating the face region, the human eye features are extracted by Gabor filtering to coarsely locate the human eye, and then the gray integration and edge detection are combined to achieve human eye positioning.

Taking the positioning of the left eye as an example, the precise positioning of the human eye is shown in Fig3. Fig. 3(a) shows the face area after Gabor transformation. It can be seen that the image after Gabor filtering mainly highlights the eyebrows, eyes, nose and mouth. After the obtained image is subjected to simple filtering, gray scale scaling and corrosion expansion, the result of Fig. 3(b) is obtained, and the positions of the eyes, the tip of the nose, and the mouth can be clearly determined. According to the structure of the face, the human eye region can be roughly positioned, and the horizontal integral projection is performed on the region to find the \( Y_m \) coordinate of the pupil center. The upper and lower limits of the human eye can be obtained according to the inherent characteristics of the human eye, the size of the captured image, and the shooting angle, so that \( E_u = Y_m - 6, \quad E_d = Y_m + 8 \). The Canny edge detection is performed between the face regions \( E_u \) to \( E_d \), and the detection result is vertically integrated and projected, and the projection is performed from the \( X_m \) column with 3 columns of projections and the moving block is leftward and rightward, respectively, to obtain the left and right limits of the human eye \( E_l \) and \( E_r \). At this point, the precise positioning of the human eye has been completed. As shown in Fig. 3 (c).

**Fatigue Feature Extraction and Fatigue Determination**

**Fatigue Feature Extraction**

Since the precisely positioned human eye region occupies too few pixels in the original image, which is not conducive to the extraction of the fatigue feature, the human eye region is magnified 25 times, and Fig. 4(a) is the enlarged eye-opening state. Fig. 4(b) shows the closed eye state after the enlargement. The Canny edge detection is performed on the enlarged human eye region, and Fig.4(c) is the human eye contour detection result in the eye-opening state, and the upper and lower eyelids can be clearly seen from the figure. Fig. 4(d) shows the result of the human eye contour detection in the closed-eye state, which is affected by the eyelashes, and the detected upper and lower human eye contours are not completely eyelids, but can be approximated as eyelids.
Since the system has a downward shooting angle, the imaging result of the human eye is not an elliptical type in the case of flat shooting, and the upper and lower eyelids are parabolic in the upward direction. Parabolic fitting of the upper eyelid contour curve was performed by least squares method. By calculating the height difference between the upper and lower eyelids at the center of the human eye and the change in the radius of curvature of the upper eyelid fitting parabola at the center of the human eye, the degree of opening of the eye can be determined, which provides a basis for fatigue determination.

**Fatigue Determination Based on PERCLOS**

This paper selects the PERCLOS method (P80 standard) to judge the fatigue state of the driver. The video has 25 frames per second and the unit time is 10 s (that is, 250 frames of effective image). The PERCLOS value $f$ is

$$f = \frac{N_c}{250} \times 100\%$$

In the formula, $N_c$ is the number of closed-eye frames in the time period, and the threshold of the PERCLOS setting fatigue is 0.4, which can determine whether the driver is fatigued.

**Experimental Results**

The camera used in the experiment is WATEC902H3 low illumination camera, $f=3.6$ automatic aperture, illumination source is of our designed light source, USB image acquisition card, the infrared light active illumination image acquisition system is installed in the actual driving environment, at night, 5 videos were recorded and fatigue detection was performed on the video image.

It can be seen that the system can accurately locate the eyes of different drivers, and is not interfered by the rear passengers. The test results are shown in Tab. 2.

**Table 2. Test results of Fatigue.**

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Total number of frames</th>
<th>Correct positioning of eye frames</th>
<th>Human eye positioning accuracy</th>
<th>Correctly determine the number of eye frames</th>
<th>Eye recognition accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5325</td>
<td>4851</td>
<td>91.10%</td>
<td>4726</td>
<td>88.75%</td>
</tr>
<tr>
<td>2</td>
<td>8664</td>
<td>7818</td>
<td>90.24%</td>
<td>7616</td>
<td>87.90%</td>
</tr>
<tr>
<td>3</td>
<td>9316</td>
<td>8513</td>
<td>91.38%</td>
<td>8246</td>
<td>88.51%</td>
</tr>
<tr>
<td>4</td>
<td>12648</td>
<td>11642</td>
<td>92.05%</td>
<td>11362</td>
<td>89.83%</td>
</tr>
<tr>
<td>5</td>
<td>15313</td>
<td>14059</td>
<td>91.81%</td>
<td>13691</td>
<td>89.41%</td>
</tr>
</tbody>
</table>

In the 5-segment detection video, the accuracy of human eye positioning is above 90%, and the accuracy of eye recognition is above 87%, which ensures the accuracy of nighttime fatigue driving.
detection. The system has good robustness when the driver's head rotation angle is less than 30 degrees. Excessive angle of rotation of the driver's head is the main cause of misidentification.

Conclusion
In view of the characteristics of fatigue driving at night, without using the driver's driving behavior, using infrared fill light uniform illumination imaging, improving the imaging quality, effectively filtering out the interference of ambient light, the improved fatigue detection algorithm is significantly improve the robustness of human eye positioning reduces the false positive rate of fatigue determination. Tested in a real driving environment, the night fatigue driving detection system is easy to install, has a low false positive rate, and has excellent detection performance. It has important significance for the development of fatigue driving detection technology and reducing the incidence of night traffic accidents.

Acknowledgements
The authors thank the editors and the anonymous reviewers for their valuable comments that helped to improve the paper. The work was supported the Special scientific research project of education department of shaanxi province (No.19JK0112).

References