Active Driving Behavior Analysis System based on Single Camera

Chenmou Wu [a], Zhijun Fang [b]

School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201600, China

[a]747519479@qq.com, [b]qxr2015@yeah.net

Abstract

One of the main reason which cause traffic accidents is the drivers' bad habits. For this reason, paper proposes a system which can recognize drivers' behaviors and monitor theirs by machine vision. Our system uses the single camera to obtain scene. By comparing the similarity between the 2D image feature and the 3D model projection,the driver's upper body posture can be estimated in real time and the three-dimensional coordinates of the skeleton node can be obtained to achieve the recognition behaviors and monitoring conditions of the driver. Finally, we have tested the effect of human pose estimation on the driving behavior and its feasibility can be proved.

Keywords

Driving Condition Monitoring, Machine Vision, Human Pose Estimation, Behavior Recognition.

1. Introduction

With the increase in car ownership,the traffic safety problem has been paid more and more attention. According to the study,the vast majority of traffic accidents are caused by improper operation of the driver [1].If the driver's emergency response time increases by 0.5s,then the possibility of a traffic accident can be reduced by about 60% [2]. Therefore,it is necessary to study the auxiliary system of active driving behavior analysis,which can promptly remind the driver by correctly identify and forecast the driver's behavior,for reducing the accident.

At present, there are three main methods to study the driver's behavior state:

(1) based on the vehicle information method, with the sensor to detect the steering wheel movement, lane departure position and speed, acceleration and other factors, when these thresholds beyond the specified range after any change, indicating that the driver in varying degrees of driving state. Paper [3], In the simulated driving environment, the chaos theory is used to determine the change rule of the steering angle signal of the driver in different mental states by obtaining the steering angle signal of the steering wheel during the driving process of the driver, so as to judge the driver's behavior state .But this method has a certain risk of error, and the vehicle state change is not entirely due to the driver's operation caused by [4-6].

(2) based on the physiological signal method, the sensor to detect the driver's physiological signal changes to determine the driver's state. This approach is to study the correlation between physiological signals and driver mental states [7]. These signals include EEG, ECG, EMG, EG, skin reaction EDA and pulse signals. This method is superior accuracy, but the driver needs to wear a large number of sensing equipment, to a certain extent, the method itself affects the driver's operation, and cost higher.

(3)Based on the driver's behavior,which use the camera to get the driver's facial features,including blink,yawn,frown,expression,head posture and so on. For example,the PERCLOS algorithm used by the NHTSA. It uses a short time to cover the pupil of the eyelids as a percentage of time,which reflecting the eyelid closure or sagging.The algorithm can effectively check the driver's mental state is tired [8]. However,due to the characteristics of human facial areas will be affected by brightness,angles and many other factors interfere,resulting in robustness is not high. And cannot handle the facial part of the block,the driver called,with sunglasses and other special circumstances.

In this paper,we propose a method based on the three-dimensional attitude estimation of the upper body of the human body,and determine the driver's behavior by detecting the driver's behavior.Finally,we set up an auxiliary system for active driving behavior analysis.

2. Organization of the Text

In this paper,we improve the accuracy and speed of the method proposed by Shian-Ru Ke [9] and Ruizhi Sun [10],and apply it on our auxiliary driving behavior analysis of the auxiliary system. The block diagram of the system is shown in Fig.1.

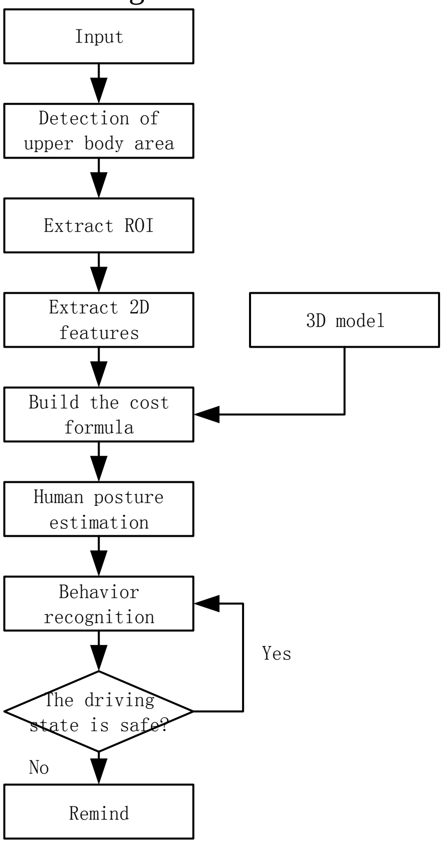

Fig. 3 Active driving behavior analysis system flow chart

In this paper ,We optimizes the method of extracting the 2D features of the image,partially corrects its cost formula,and uses the simulated annealing algorithm instead of APSO to minimize the cost function,thus realizing the 3D model tracking to estimate Predict the driver's posture and use OpenGL to get real three-dimensional coordinates of the 8 skeletal nodes of the human body. And finally the behavior of the driver to identify,so as to determine the current state of the driver's state.

2.1 Determine the Location of the Human Body

Our system environment is in the car, the camera has always been the picture of the front of the upper body, so with two conditions to determine the general position of the human body to improve efficiency.

The first condition is to use the face to roughly determine the human area. We uses Cascade CNN [11] instead of haar-like and Adaboost [12, 13] to determine the face area. Cascade CNN is greater than speed and accuracy relative to haar algorithm. Another condition is to determine the face and hands. In all areas with skin color characteristics, we look for the three largest areas that meet this feature as the position of head and hands. Then, we calculate the distance between them, based on the three positions in most cases can constitute the characteristics of the triangle to establish the relationship between the function 1.If $r_d \le 1$ and $d_3 \le d_1 + d_2$, we think it maybe exists human, and then determine the general position. Eventually, the human body region is extracted from each frame in real time.

$$r_d = \frac{d_3}{d_1 + d_2} \quad (1)$$

Where d_1 is the distance from the head to the left, d_2 is the distance from the head to the right, d_3 is the distance between the hands, and r_d is the distance between the hands and the head.

2.2 Extraction of 2D Features.

We use the system in the two-dimensional features are Edge, color, Projection and so on. The human body has a very obvious edge feature, which often used in human posture estimation. In this paper, we use the classical Sobel operator [14] and add the inter-frame variation to enhance the edge detection.

Color as an intuitive visual feature, can well reflect the appearance of the target image and the scene in which it is one of the widely used features in the field of computer vision. HSV is visually closer to human visual perception of color and is superior to RGB in confrontation noise [15, 16]. By setting the skin color detection model in a specific area in the HSV color space, the region where the hue (H) satisfies [5, 25] ,the saturation (S) satisfies [50,200],and the luminance (V) satisfies [20,200]. The skin color image template is defined as Equation 2.

$$Im^t_{skin} = Im^t_{skin}{}' \cap Im^{t-1}_{skin} \cup (Im^{t-1}_{Skin-motion} \cup Im^t_{skin\_subtraction}) \quad (2)$$

Where $Im^t_{skin}{}'$ is the skin image detected by the current frame, Im^{t-1}_{skin} is the skin image detected by the previous frame, $Im^{t-1}_{Skin-motion}$ is the area where the skin color of the previous frame changes, and $Im^t_{skin\_subtraction}$ is the image of the skin image between the two frames , Im^t_{skin} is the final frame of the current skin image template.

In Fig.2, the 2D projection feature template is defined as

$Im_{silhouette} = Im_{subtraction} \cup Im_{edge} \cup Im_{skin}$, where $Im_{subtraction}$ is obtained by the background difference of two frames, and Im_{edge} and Im_{skin} are mentioned above.

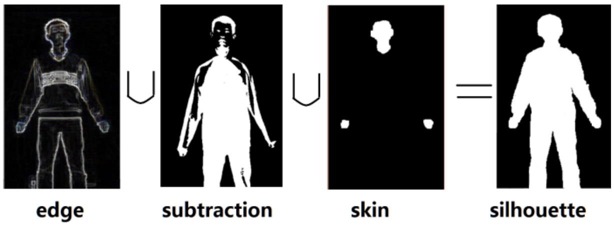

Fig. 3 $Im_{silhouette}$

2.3 Cost Function.

The essence of the human body posture estimation is that each model makes the 3D model in a limited range to do random movement, and then in each movement through the cost function to determine the 3D model and the human upper body of the best match image, the simulation model to track to The process of the human body. In this paper, we construct the cost function, as shown in Equation 3, which includes the edge score, the motion score, the projection score, the scale score, the skin score and the arm score, and the total of the six characteristic scores. $W_n (n = 1,2,3,4,5)$ is the weight value, the scores are determined by Equation 4.

$$Cost = W_0 * S_{edge} + W_1 * S_{motion} - W_2 * S_{silhouette} + W_3 * S_{ratio} + W_4 * S_{skin} + W_5 * S_{arm} \quad (3)$$

$$Score(i, j) = \max_n (\sum I_{xor}(\alpha, \beta) / Area(O_{(i,j,n)})) \quad (4)$$

Where $O_{(i,j,n)}$ is the position of a body part on a 2D image, n is the number of iterations, Area is the selected area, and $(\alpha, \beta) \in O_{(i,j,n)} : (i,j)$ is the coordinate of the selected part. I_{xor} is the intersection of the 2D image projection and the 3D model projection.

2.4 Simulated Annealing Algorithm.

Simulated annealing algorithm is used to minimize the cost function, the smaller of cost function, the higher similarity of matching. Compared to the particle swarm optimization algorithm, simulated annealing is more efficient and achieves similar effects. The algorithm flow is shown in Fig.3.

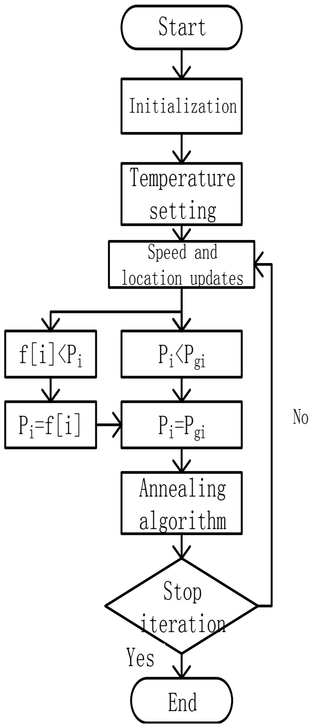

Fig. 3 Simulated annealing algorithm flow chart

3. Experimental Results and Analysis

The auxiliary system of active driving behavior analysis proposed in this paper is based on Windows10 64-bit system, the processor is Inter (R) Core (TM) i7-5500U CPU@2.40GHz,RAM 4.00GB,and OpenCV2.4.9,OpenGL,Visual Studio2013 configuration environment, the use of C + + programming, the algorithm was implemented. The experimental results have an image size of 640 X 320 and 25-35fps.

This paper determines the case of normal driving and one-hand driving by calculating the ratio between the distance between the left and right hands and the shoulder width. For four kinds of driving behavior pose estimation in Fig.4, the initial value of the ratio cutoff point is replaced by the diameter of the steering wheel (35.5-40.5cm) and the distance between the hands and the standard adult man's shoulder width (45-55cm) is 0.85. But because the palm of your hand and the steering wheel after the distance is bound to be greater than the diameter of the steering wheel and the shoulder width of each person should be modified in the initial value after a number of experiments found in the hand away from the steering wheel when the ratio curve to a rapid upward trend, Change the value of the demarcation point to 1, as shown in Fig.5.

In the process of answering the phone, the distance between the hand and the head will be closer. As showed in Fig.6, the blue line is the distance between the right hand and the head, and the red line is the distance between the left and right hands. As showed in the figure, when the distance between the former call will be reduced while the latter will increase rapidly, the program is feasible.

In the case of drunkenness / fatigue driving, the head coordinates will change significantly. Fig.7 blue line for the normal driving coordinate, red line for the drunk driving coordinates. In normal driving, the coordinates of the x-axis will fluctuate, but the whole is in a certain range, and when the intoxicated state, the coordinates will be significantly undulating. Similarly, in the y-axis and z-axis in the ups and downs is particularly prominent, the experimental process, this article in 6 seconds to simulate the three times doze, in Fig.8 and Fig.9, there are three obvious ups and downs. It is shown that this paper through the coordinates of the change to confirm whether the driver drunk / fatigue driving the feasibility.

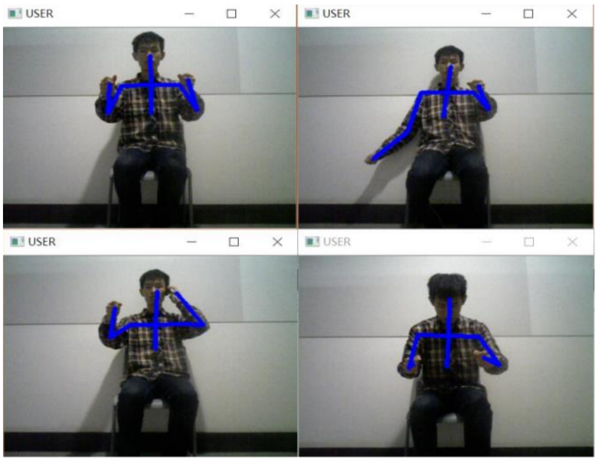

Fig. 4 Four kinds of driving behavior pose estimation

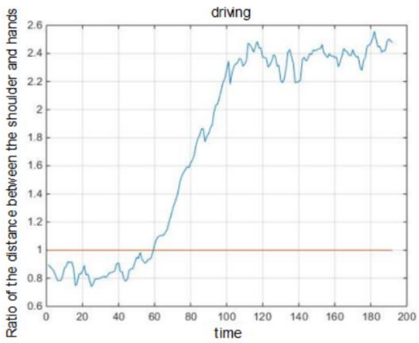

Fig.5 Ratio of the distance between the hand the shoulder and hands.

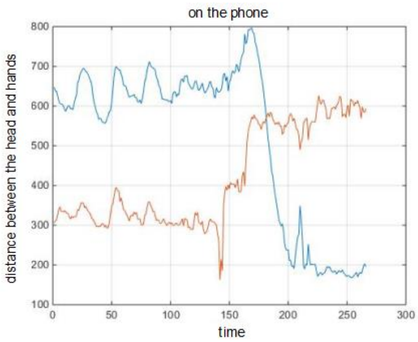

Fig.6 The distance between the hand and /the hand and the head

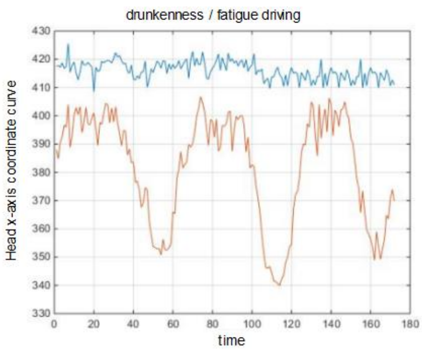

Fig.7 Head x-axis coordinate curve.

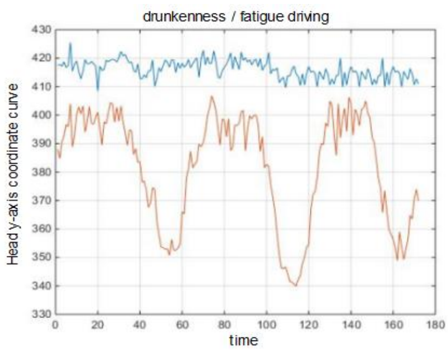

Fig.8 Head y-axis coordinate curve

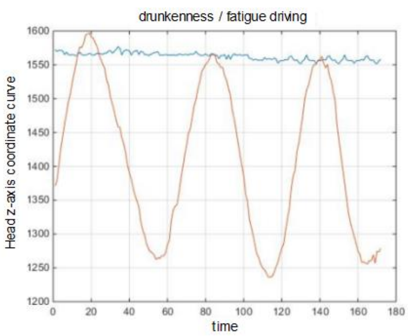

Fig. 9 Head z-axis coordinate curve

4. Conclusion

In view of the bad habit of driving behavior, this paper puts forward the method of 3D human posture estimation to carry on the behavior recognition to the driver. Based on the active driving behavior analysis assistance system realized by this paper, four kinds of driving behavior states were simulated and tested. The experimental results confirm the feasibility and real - time of the system method.

References

[1] Traffic Management Bureau of the Public Security Ministry. Road traffic accident statistics report of China(2014) [R]. Wuxi: Traffic Management Research Institute of the Ministry of Public Security,2015. (in Chinese)
[2] Nobe S A,Wang F Y. An overview of recent developments in automated lateral and longitudinal vehicle controls[C] Systems,Man,and Cybernetics,2001 IEEE International Conference on. IEEE,2001,5: 3447-3452.
[3] Takei Y,Furukawa Y. Estimate of driver's fatigue through steering motion[C] Systems,Man and Cybernetics,2005 IEEE International Conference on. IEEE,2005,2: 1765-1770.
[4] Das D,Zhou S,Lee J D. Differentiating alcohol-induced driving behavior using steering wheel signals[J]. Intelligent Transportation Systems,IEEE Transactions on,2012,13(3): 1355-1368.
[5] Simons R,Martens M,Ramaekers J,et al. Effects of dexamphetamine with and without alcohol on simulated driving[J]. Psychopharmacology,2012,222(3): 391-399.
[6] Mets M A J,Kuipers E,Senerpont Domis L M,et al. Effects of alcohol on highway driving in the STISIM driving simulator[J]. Human Psychopharmacology: Clinical and Experimental, 2011, 26(6): 434-439.
[7] Sahayadhas A,Sundaraj K,Murugappan M.Detecting driver drowsiness based on sensors:a review[J]. Sensors,2012,12(12): 16937-16953.
[8] Ingre M,ÂKerstedt T Ö R,Peters B Ö R,et al. Subjective sleepiness,simulated driving performance and blink duration: examining individual differences[J]. Journal of sleep research, 2006, 15(1): 47-53.
[9] Ke S R,Zhu L J,Hwang J N,et al. Real-Time 3D Human Pose Estimation from Monocular View with Applications to Event Detection and Video Gaming[M]. 2010.
[10] Ruizhi.3D Front-View Human Upper Body Pose Estimation Using Single Camera[D]. University of Washington,2013.
[11] Li H,Lin Z,Shen X,et al. A convolutional neural network cascade for face detection[C]// Computer Vision and Pattern Recognition. IEEE,2015:5325-5334.
[12] Viola P,Jones M. Rapid Object Detection using a Boosted Cascade of Simple Features[C]// Computer Vision and Pattern Recognition,2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. IEEE Xplore,2001:I-511- I-518 vol.1.
[13] Lienhart R,Maydt J. An extended set of Haar-like features for rapid object detection[C]// International Conference on Image Processing. 2002. Proceedings. IEEE Xplore,2002:I- 900 - I-903 vol.1.
[14] Chien Y. Pattern classification and scene analysis[J]. The Library Quarterly: Information, Community,Policy,1974,19(Volume 44,Number 3):462-463.
[15] Yoon I,Kim S,Kim D,et al. Adaptive defogging with color correction in the HSV color space for consumer surveillance system[J]. IEEE Transactions on Consumer Electronics,2012,58(1): 606-607
[16] Schwarz M W,Cowan W B,Beatty J C.An experimental comparison of RGB,YIQ,LAB, HSV, and opponent color models[J]. Acm Transactions on Graphics,1987,6(2):123-158.