Driver Distraction Detection and Identity Recognition in Real-time

Jinhua Zeng*†, Yaoru Sun*, Li Jiang*
*Department of Computer Science and Technology
Tongji University
Shanghai, China
Email: 0mrzeng@tongji.edu.cn (Jinhua Zeng),
yaoru@tongji.edu.cn (Yaoru Sun),
09jiangli@tongji.edu.cn (Li Jiang)
†Corresponding author

***Abstract*—Drivers attending to primary driving tasks show specific eye and head movement behaviours, while the distracted drive generally covers the states including drivers' eyes off the road and long-term eye closure. This paper presents a distraction detection system by using the strategy of "attention budget". The states of eyes off the road and face with closed eyes are used to lessen the "attention budget" while the reversed conditions gain it. Drivers' gaze estimation is derived from the head motion, and the stage classifiers working with haar-like features are used to detect head movements and eye states. With regard to the factors of drivers' personal characteristics in distraction detection, the recognition of drivers is implemented by extraction and matching of scale invariant feature transform features in detected frontal face. The results of experiments validate the effectiveness and robustness of the system.**

***Index Terms*—driver distraction detection; identity recognition; attention budget; scale invariant feature transform;**

I. Introduction

Driver distraction is a priority issue in road safety which grabs the drivers' attention and is detrimental for driving [1]. The factors in definition of driver distraction include personal characteristics, driving conditions and situations, and risk of consequences [2]. Drivers are susceptible to distraction, and are frequently engaged in multiple tasks, which impair driving performance and add risks of crash involvement [3] [4]. Currently, in-vehicle information systems (IVIS) provide the drivers with excellent entertainment qualities and risks of accidents caused by distraction as well. Out-vehicle traffic, passengers, and others all have the potential to direct the drivers' attention away from the driving task. Driver distraction related accidents have become a great socio-economic concern.

Drivers in distraction exhibit certain physiological features, vehicle behaviours, and visual driving behaviours. The physiological features include brain activity, heart rate, eye movement, etc., which can be acquired by electroencephalograph (EEG), electrocardiograph signals (ECG), electro-oculogram signals (EOG), and etc. The vehicle behaviours provide many valuable indicators to estimate the drivers' distraction. The non-intrusive feature of the information acquisition techniques for the visual cues makes visual driving behaviours more practically feasible. The cues in visual driving behaviours have been demonstrated to be valuable to detect distraction, e.g., eye-related parameters, nodding, yawing, head orientation [5] [6] [7] [8], and etc. In some works [8] [9], vehicle surrounding analysis was assembled in the system for driver distraction detection. In [8], driver gaze estimation and vehicle surrounding analysis were used to predict intentions and estimate the driver distraction. Driver gaze estimation was derived from the head motion, while the vehicle surrounding analysis was implemented by the motion-based visual saliency map of the scene. The gaze patterns were validated by the saliency map to determine whether the gaze shift was caused by visual distraction or due to goal-oriented visual search.

The work in this paper aims to develop a system for driver distraction detection which can be served as an automated co-drivers in-vehicle. Detection of drivers' frontal face with robustness to head offset and slight tolerance of head rotation was carried out by using the tree of stage classifiers of frontal face working with haar-like features. The states of the two eyes were independently detected through a cascade of boosted classifiers of the monocular in limited searching areas. The recognition of drivers' identity was accomplished by extracting scale invariant feature transform (SIFT) keypoints from detected frontal face and matching the features with a database in our system. The flow of driver distraction detection is shown in Fig. 1.

II. Head movement and states of eyes detection

A. Head movement detection

The strategy of measuring distraction in the work used the so-called "attention budget" [10]. Each driver will be assigned a certain "attention budget" according to personal characteristics and driving environments, e.g., age, medical conditions, time of day, and etc. The "attention budget" can be depleted either by events of eyes off the road or faces with closed eyes. And it will be exhausted if the first event lasts longer than 2 seconds or the second exceeds 1 second, but the losing "attention budget" can be compensated if the drivers look back at the road or open their eyes. In our system, drivers' gaze estimation is derived from the head motion [8], and

978-0-7695-4304-8/10 $26.00 © 2010 IEEE
DOI 10.1109/GCIS.2010.83

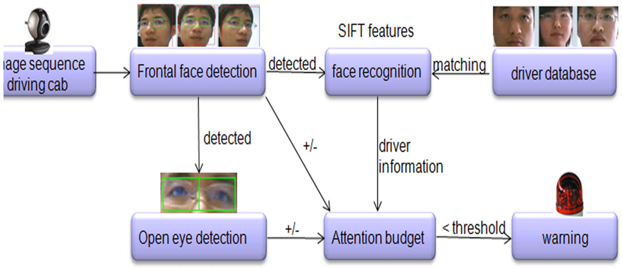

Fig. 1. The system flow of driver distraction detection. The size of the input image is $320*240$ pixels, and the frame rate is 25 frames per second. The size of the face in the database is $100*100$ pixels. The sign "$+/-$" indicates that the system increases/decreases the "attention budget". If the "attention budget" is less than given threshold, the system will trigger the warning.

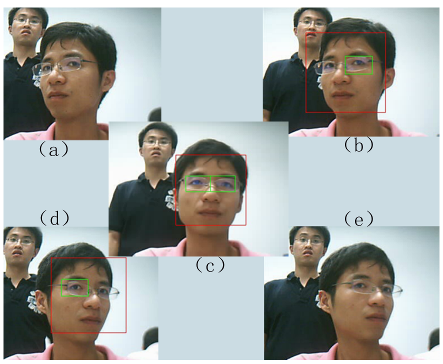

Fig. 2. The results of the face detection. Head was rotated too right and too left to be detected in the sub-figures a and e respectively. Head offset and slight head rotation could be detected in the system, and the interference of the multiple faces was excluded through the different face sizes.

detection of the event of eyes off the road was implemented by detecting frontal face using tree of stage classifiers [12] working with haar-like features. The detection of frontal face must be robust to head offset and the interference of multiple faces, and have a certain tolerance to head rotation. The interference of multiple faces in driving environments can be excluded by resorting to different face sizes, because the driver is generally closer to the in-vehicle webcam than any other passengers. Our system limits the minimum size of detection window in the classifiers so as to only detect the face with adequate size. Head offset and slight head rotation can be robustly detected in tree of stage classifiers because of their searching strategy and non-strict training samples of frontal faces. If the result of the detected face is only one, then we think that the driver's eyes are keeping in the road, or else the opposite condition will be considered, i.e., eyes off the road. The results of the frontal face detection are shown in Fig. 2 and Fig. 3.

B. States of eyes detection

Open eye detection in the system was implemented by a cascade of boosted classifiers of the monocular [13]. Because the locations of two eyes have a relative fixed relation with

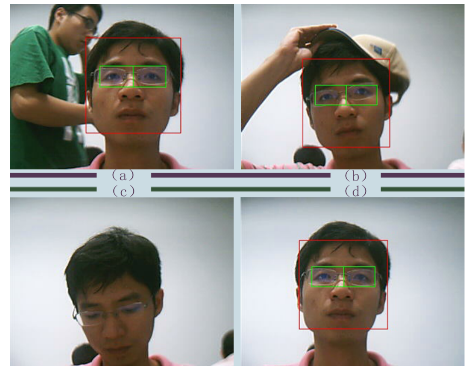

Fig. 3. The results of the face detection. Face detection could tolerate some ornaments, e.g., glasses and hat, as shown in sub-figures a, b, and d. The system considered the nodding as the eyes off the road as well, as shown in sub-figure c.

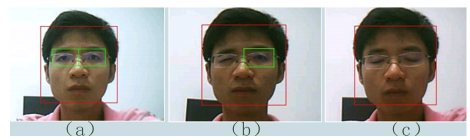

Fig. 4. The results of the eye detection. Only open eye could be detected, as shown in sub-figures a, b, and c.

the area of the detected face, we minimize the searching areas of two eyes as follows.

Suppose that the area of the detected face is set as $[x, y, width, height]$ (x and y is 2D coordinates of the upper left corner of the face in the image, and $width, height$ is the width and height of the face.), the searching area of the left eye and right eye will be $[x\_l, y\_l, width\_l, height\_l]$ and $[x\_r, y\_r, width\_r, height\_r]$ respectively.

$$
\begin{aligned}
&\text{For the right eye:} \\
x\_r &= x + width/2 * 0.3 \\
y\_r &= y + height/2 * 0.6 \\
width\_r &= width - width/2 * 0.3 \\
height\_r &= height - height/2 * 0.55 \\
&\text{For the left eye:} \\
x\_l &= x\_r + width\_r \\
y\_l &= y\_r \\
width\_l &= width\_r \\
height\_l &= height\_r
\end{aligned}
$$

Our system detects two eyes independently . If results of the detection in two searching areas are both none, the driver's eyes are thought to be closed, or else we consider that the driver's eyes are open. The results of the open eye detection are shown in Fig. 4.

The frame rate of the webcam is 25 frames per second. The "attention budget" was updated with the time interval of 40 ms. In each time, head movement and states of eyes were

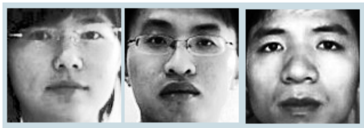

Fig. 5. Image samples of the frontal faces in the database. The color images were firstly converted to greyscale, and then the histogram equalization was applied to them.

detected, and the results were used to update the budget in real-time.

III. SIFT FEATURE-BASED DRIVERS' IDENTITY RECOGNITION

Personal characteristics, e.g., age and medical conditions, constitute a potential part in the definition of distraction [2]. Intelligent system for driver distraction detection must dynamically adjust system parameters, e.g., warning threshold of "attention budget", according to each driver's characteristics and their driving surroundings. The recognition of the driver's identity is an essential part of distraction detection system in practice.

The recognition of the drivers' identity in our system was implemented by face recognition by using the feature points of SIFT [14] [15]. The SIFT feature descriptor is invariant to scale, orientation, affine distortion, and even partially robust to illumination changes [14]. The process of the recognition can be divided into two general steps. The first step is to extract the SIFT key-points in detected frontal face and the second is to match the key-points with the database where sets of all drivers' SIFT features are stored independently in advance. In our system, the number of the drivers is 10. The approach of feature matching is a modification of the k-d tree algorithm proposed in [16]. Suppose that the number of the SIFT features extracted from the detected frontal face is Nd, and each feature will be compared with all sets of features in the database separately.

In order to be robust to illumination changes, our database keeps three sets of SIFT features of each driver, and these three sets are extracted from the driver's frontal face in three different illumination conditions respectively, i.e., weak, medium, and bright illumination conditions. The image pre-processing technique of histogram equalization is applied to automatically standardize the brightness and contrast of frontal face images. Partial frontal face images of drivers in the database are shown in Fig. 5. The result of the matching is set as a array Am i.e.,

$$Am = [(m_{11}, m_{12}, m_{13}), \ldots, (m_{ij}, \ldots), \ldots] \qquad i = 1, \ldots, 10$$

and m_{ij} is the matching number of feature sets between detected frontal face and the $i\text{-}th$ driver's set of SIFT features under the $j\text{-}th$ illumination condition. Suppose that m_{ij} is the maximum of the elements in the array Am, if m_{ij} is greater than a given threshold (the threshold was set to 15 in our system), then we think that the identity of current driver is the one with index of i and the corresponding parameters of this driver will be loaded into the system, otherwise we reject the recognition result and consider the current driver as a new one and the default parameters of the driver will be loaded into the system.

A. Data Analysis of Experiments

We test the recognition system by testing the same driver field ten times at different moments. The amounts of SIFT features extracted from ten subjects' frontal faces in the database under bright illumination condition are shown in table I. The size of each frontal face image in the database is $100 * 100$ pixels. Mean and standard deviation of the amounts of SIFT features extracted from the database are 72.7 and 9.09 respectively.

TABLE I
THE AMOUNTS OF SIFT FEATURES EXTRACTED FROM THE DATABASE UNDER BRIGHT ILLUMINATION CONDITION.

Subject	01	02	03	04	05	06	07	08	09	10
Features Numbers	82	77	67	90	66	71	63	61	72	78

The amounts of SIFT features matching between the detected frontal face (driver's index is 01 in the experiments) and ten subjects' frontal faces in the database under bright illumination condition are shown in table II. We use the

TABLE II
THE MATCHING AMOUNTS OF SIFT FEATURES BETWEEN THE DETECTED FRONTAL FACE (SUBJECT'S INDEX IS 01 IN THE EXPERIMENTS) AND TEN SUBJECTS' FRONTAL FACES IN THE DATABASE UNDER BRIGHT ILLUMINATION CONDITION.

Subject	Sample Number	Mean	Std. Deviation
01	10	38.5	4.28
02	10	0.2	0.42
03	10	0.0	0.0
04	10	1.5	0.71
05	10	1.2	0.63
06	10	1.1	0.57
07	10	1.4	1.17
08	10	0.8	0.63
09	10	0.9	0.57
10	10	2.3	1.06

statistics analysis of one-way ANOVA to analyze the mean difference between those matching numbers. The result shows significant difference.

$$F(9, 90) = 614.48 \text{ and Sig. = 0.000}$$

The system successfully recognized the driver in all tests. The matching numbers in ten tests were shown in table III.

TABLE III
THE MATCHING NUMBERS BETWEEN THE SIFT FEATURES EXTRACTED FROM THE DETECTED FRONTAL FACE AND THE SET OF SIFT FEATURES WITH DRIVER'S INDEX OF 01 IN THE DATABASE UNDER THE BRIGHT ILLUMINATION CONDITION.

Test No.	01	02	03	04	05	06	07	08	09	10
Matching Number	34	31	39	41	36	39	41	43	45	36

IV. Conclusion

The strategy of "attention budget" in our system can effectively detect drivers' distraction. We use the events of eyes off the road or face with closed eyes to lessen the "attention budget", and use reversed conditions to gain it. The face and eye state detectors were efficiently integrated in the vigilance detection system. The frontal face and eyes are the distinct patterns with respect to the backgrounds because of their internal components. The detectors have a robust detection rate and a low false alarm rate, which makes the system perform amazingly well. Because of the stability of the identity recognition, the function of the drivers' identity recognition can serve as an access control system, and only the identified drivers can access to driving.

Acknowledgment

This work was supported by Grants from the National Natural Science Foundation of China (60775019), Science and Technology Commission of Shanghai Municipality (09511502500), and Shanghai Pujiang Program (09PJ1410200).

References

[1] K. Young and M. Regan, "Driver distraction: A review of the literature," in: I. J. Faulks, M. Regan, M. Stevenson, J. Brown, A. Porter, and J. D. Irwin, Eds. Distracted driving, Sydney, NSW: Australasian College of Road Safety, 2007, pp. 379-405.

[2] K. Kircher, "Driver distraction: A review of the literature," VTI report 594A, 2007.

[3] T. A. Dingus, S. G. Klauer, V. L. Neale, A. Petersen, S. E. Lee, J. Sudweeks, M. A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z. R. Doerzaph, J. Jermeland, and R. R. Knipling, "The 100-car naturalistic driving study, phase II - Results of the 100-car field experiment," Report DOT HS 810593, NHTSA, U.S. Department of Transportation, 2006.

[4] S. G. Klauer, T. A. Dingus, V. L. Neale, J. D. Sudweeks, and D. J. Ramsey, "The Impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data," Report DOT HS 810594, NHTSA, U.S. Department of Transportation, 2006.

[5] Q. Ji and X. J. Yang, "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance," Real-time imaging, vol. 8, pp. 357-377, 2002.

[6] T. D. Orazio, M. Leo, C. Guaragnella, and A. Distante, "A visual approach for driver inattention detection," Pattern recognition, vol. 40, pp. 2341-2355, 2007.

[7] F. Xiao, C. Y. Bao, and F. S. Yan, "Yawning detection for monitoring driver fatigue," Proceedings of the Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, pp. 664-668, 2007.

[8] A. Doshi and M. Trivedi, "Investigating the relationships between gaze patterns, dynamic vehicle surround aalysis, and driver intentions," 2009 IEEE Intelligent Vehicles Symposium, 2009.

[9] L. Fletcher and A. Zelinsky, "Driver inattention detection based on eye gaze road event correlation," The international journal of robotics research, vol. 28, pp. 774-801, 2009.

[10] R. Karlsson, "Evaluating driver distraction countermeasures," VTI notat 38A, 2005.

[11] T. Dukic, L. Hanson, K. Holmqvist, and C. Wartenberg, "Effect of button location on driver's visual behaviour and safety perception," Ergonomics, vol. 48, no. 4, pp. 399-410, 2005.

[12] R. Lienhart and J. Maydt, "An extended set of haar-like features for rapid object detection," IEEE ICIP 2002, Vol. 1, pp. 900-903, 2002.

[13] M. Castrillón, O. Déniz, C. Guerra, and M. Hernández, "ENCARA2: Real-time detection of multiple faces at different resolutions in video streams," Journal of Visual Communication and Image Representation, vol. 18, no. 2, pp. 130-140, 2007.

[14] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[15] R. Hess and A. Fern, "Improved video registration using non-distinctive local image features," in Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2007.

[16] J. Beis and D. G. Lowe, "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces," Conference on Computer Vision and Pattern Recognition,Puerto Rico, pp. 1000-1006, 1997.