TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 18/18 pp354-362
Volume 17, Number 3, June 2012

# Driver Drowsiness Recognition Based on Computer Vision Technology*

Wei Zhang, Bo Cheng**, Yingzi Lin†

**State Key Laboratory for Automotive Safety and Energy, Department of Automotive Engineering, Tsinghua University, Beijing 100084, China;**
**† Mechanical and Industrial Engineering Department, Northeastern University, Boston, MA 02115, USA**

**Abstract:** Driver drowsiness is one of the major causes of traffic accidents. This paper presents a nonintrusive drowsiness recognition method using eye-tracking and image processing. A robust eye detection algorithm is introduced to address the problems caused by changes in illumination and driver posture. Six measures are calculated with percentage of eyelid closure, maximum closure duration, blink frequency, average opening level of the eyes, opening velocity of the eyes, and closing velocity of the eyes. These measures are combined using Fisher's linear discriminant functions using a stepwise method to reduce the correlations and extract an independent index. Results with six participants in driving simulator experiments demonstrate the feasibility of this video-based drowsiness recognition method that provided 86% accuracy.

**Key words:** driving safety; drowsy driving; drowsiness recognition; computer vision

## Introduction

Driver drowsiness resulting in reduced vehicle control is one of the major causes of road accidents [1]. Driving performance deteriorates with increased drowsiness with resulting crashes constituting 20%-23% of all vehicle accidents[2]. The National Highway Traffic Safety Administration (NHTSA) conservatively estimates that 100 000 reported crashes are caused by drowsy drivers each year in the U.S. alone. These crashes result in more than 1500 fatalities, 71 000 injuries, and an estimated $12.5 billion in diminished productivity and property loss[3].

Many efforts have been made recently to develop on-board detection of driver drowsiness. A number of approaches have been investigated and applied to characterize driver drowsiness using physiological measures, ocular measures, and performance measures.

Physiological measures are based on physiological signals such as brain waves[4], heart rate[5], pulse rate[6], and respiration[7]. These measures are believed to be the most accurate for determining drowsiness. However, these techniques are intrusive since they require that electrodes be attached to the drivers.

Various studies have also shown the possibility of drowsiness detection by means of ocular measures. These methods generally capture the driver's eye state using computer vision systems to indicate his/her drowsiness level. The PERcentage of eyelid CLOSure (PERCLOS) has been used by many researchers as a valid ocular measure to indicate drowsiness[2]. Other measures include the duration of eye closure[8] and blink rate[9].

A driver state of drowsiness can also be characterized by the resulting vehicle behavior such as the lateral position[10], steering wheel movements[11], and time-to-line crossing[12]. Although these techniques are

Received: 2011-05-05; revised: 2012-04-23
* Supported by the National High-Tech Research and Development (863) Program of China (No. 2009AA11Z214) and Independent Scientific Research Program of Tsinghua University (No. 20101081763)
** To whom correspondence should be addressed.
E-mail: chengbo@tsinghua.edu.cn; Tel: 86-10-62780355

not intrusive, they are subject to several limitations related to the vehicle type, driver experience, and geometric characteristics and condition of the road.

Among these various possibilities, the monitoring of a driver's eye state by a camera is considered to be the most promising application due to its accuracy and non-intrusiveness. The driver's symptoms can be monitored to determine the driver's drowsiness early enough to take preventive actions to avoid an accident.

Though many studies have developed image-based driver alertness recognition systems using computer vision techniques, many problems still remain. First, eye detection remains a challenging problem with no inexpensive or commercial solutions. For some applications, eye feature detection can be satisfactory, but these only used frontal face images taken with controlled lighting conditions. In a car, the constantly changing lighting conditions cause dark shadows and illumination changes, such that effective techniques in stable lighting often do not work in this challenging environment. The performance of current algorithms degrades significantly when tested across different postures and illumination conditions, as documented in a number of evaluations[13-15]. A second problem is that current systems do not use identification and correlation analysis of various visual measures. Typical visual characteristics of a driver with a reduced alertness level include longer blink duration, slow eyelid movement, smaller eye openings, and frequent nodding. Most interest reports have only concerned PERCLOS, which is defined as the proportion of time in a specified time period that the participant's eyes are closed. Few studies have focused on the effectiveness of other measures for drowsiness recognition. Moreover, correlation analysis of all the visual measures has not been widely used to extract a more accurate independent index.

# 1 Eye Movement Detection

When the driver is directly opposite the camera, an AdaBoost-based face detector can be used to determine the position and scale of the face[16]. An Active Shape Model (ASM) algorithm is used to precisely locate the human eyes from the front-view images.

The ASM[17] is a popular statistical approach to represent deformable objects where shapes are represented by a set of feature points. Feature points are found by searching gray-level profiles with Principal Component Analysis (PCA) used to analyze the modes of the shape variation so that the object shape can only deform in specific ways that are found in the training data. The local appearance model, which represents the local statistics around each landmark, allows for an efficient search to be conducted to find the best candidate point for each landmark. Figure 1a illustrates an example of an alignment result using the ASM algorithm.

ASM is a powerful statistical tool to identify face alignment; however, the results are degraded by changes in illumination and posture. Figure 1b illustrates the influence of lighting conditions on the robustness and accuracy of the face alignment. Figure 1c shows the result when the face is turned sideways.

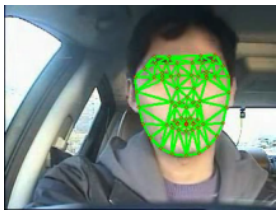
The tolerance of the ASM algorithm to illumination changes is improved by a modified ASM method with a local texture model learned from the self-quotient image instead of the original image.

The self-quotient image, $Q$, of the original image, $I$, is defined as

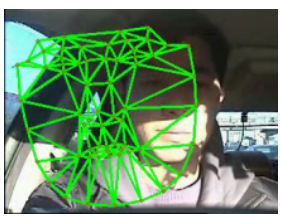
$$Q = \frac{I}{\hat{I}} = \frac{I}{F * I} \tag{1}$$

where $*$ is the convolution operation, $\hat{I}$ is the smoothed $I$ and $F$ is the smoothing kernel. The result is shown in Fig. 2.

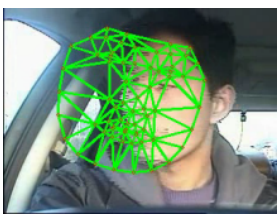
The self-quotient image has illumination invariant properties since it depends only on the albedo information and, therefore, is illumination-free. Figure 3 shows some results of self-quotient images for various illumination conditions.



(a)



(b)



(c)

**Fig. 1 ASM results**

(a) Original image

(b) Smoothed image

(c) Self-quotient image

**Fig. 2 Definition of self-quotient image**

(a)

(b)

(c)

**Fig. 3 Self-quotient images for various illumination conditions**

The results show that the illumination conditions have little effect in the self-quotient images. Moreover, the facial contour is then highlighted in the self-quotient images, which improves the speed and accuracy of the ASM algorithm.

The process of building the statistical texture model for the ASM algorithm using the self-quotient images begins by sampling the profile to get $k$ pixels on either side of the model point in the $i$-th training image as shown in Fig. 4 to get $(2k+1)$ samples, which are put

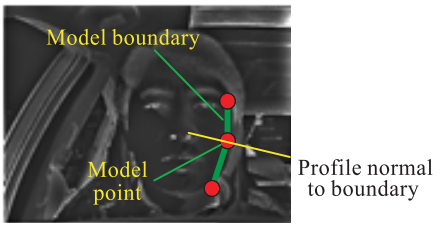



**Fig. 4 ASM texture model of the self-quotient image**

into a vector $\boldsymbol{g}_i$.

This is repeated for each training image to obtain a set of normalized samples, $\{\boldsymbol{g}_i\}$, for the given model point. These are assumed to have a multivariate Gaussian distribution with a mean $\bar{\boldsymbol{g}}$ and covariance $\phi_g$.

$$\boldsymbol{g} \sim N(\bar{\boldsymbol{g}}, \phi_g) \tag{2}$$

This gives a statistical model for the gray-level profile about the point. This is repeated for every model point, giving one gray-level model for each point.

The quality of fit of a new sample, $\boldsymbol{g}'$, to the model is given by

$$f(\boldsymbol{g}') = (\boldsymbol{g}' - \bar{\boldsymbol{g}})^{\mathrm{T}} \phi_g^{-1} (\boldsymbol{g}' - \bar{\boldsymbol{g}}) \tag{3}$$

This is the Mahalanobis distance of the sample from the model mean. Minimizing $f(\boldsymbol{g}')$ is equivalent to maximizing the probability that $\boldsymbol{g}'$ comes from the distribution. This is repeated for each model point, giving a suggested position for each point. An alignment result of the ASM algorithm is shown in Fig. 5.

After the eye region is determined by the ASM algorithm, deformable templates[18] are used to characterize the eyes. An example of the deformable template detection result is shown in Fig. 6.

The ASM algorithm is not used with each frame, but the robustness to the driver's posture and eye detection speed is improved using a mean-shift algorithm as the eye tracker. During the tracking phase, the search space can be reduced since the system has an estimate of the eye's position from the previous frame. Thus, the tracking is relatively fast computationally.

The mean-shift algorithm uses a self-quotient image as the input for the density distribution to be analyzed.

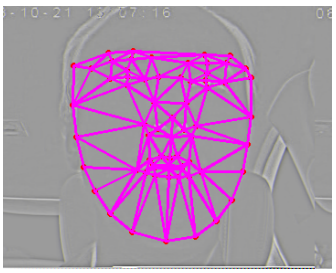


**Fig. 5 ASM result on self-quotient image**

**Fig. 6 Result of deformable templates**

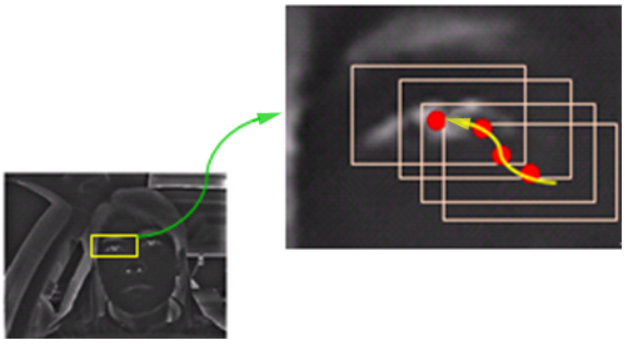


**Fig. 7 Eye tracking by the mean-shift algorithm**

The mean-shift algorithm runs as follows.

**Step 1** Get the initial search window from ASM;

**Step 2** Compute the window's center of mass;

**Step 3** Center the window at the center of mass;

**Step 4** Return to Step 2 until the window stops moving.

The location converges to a local maximum (peak) of the distribution under the window. Figure 7 gives an example of the tracking process. The arrow indicates the convergence process on a local peak (eye) in the self-quotient image.

Every time the driver is directly opposite the camera, the ASM algorithm is used to reset the eye region.

## 2 Drowsiness Features

Tests show that the eye movements differ greatly for various drowsiness levels, as shown in Fig. 8. The fully awake state is characterized by low frequency, high speed movements with larger mean amplitude. When drivers become drowsy, the mean amplitude of the eye movements decreases, and the frequency increases significantly. As the drowsiness increases, the mean amplitude of the eye movements decreases further, with long eye closures in the very drowsy state.

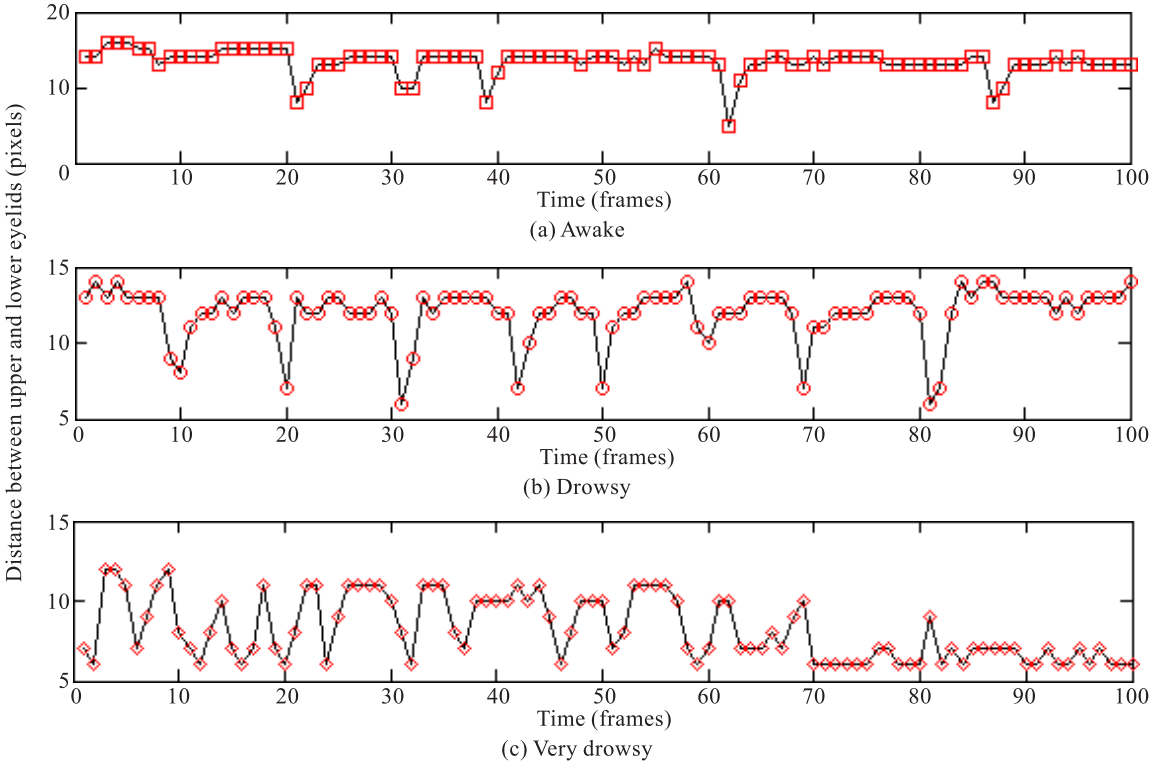



**Fig. 8 Typical movement features of eyes in various states of drowsiness**

These features can be used to define a number of measures such as PERCLOS, Blink Frequency (BF), Maximum Close Duration (MCD), Average Opening Level (AOL), Opening Velocity (OV), and Closing Velocity (CV) which are used in the paper. The definitions of these measures are shown in Table 1 and Fig. 9. In Fig. 9, $T$ is the time window and $t_1$-$t_n$ are the durations of each eye closure.

# 3 Test Design

## 3.1 High fidelity driving simulator

The tests were performed on the moving base simulator of Tsinghua University, Beijing, China. An approximately 360° visualization of the vehicle's surroundings is achieved by five screens as shown in Fig. 10.

Figure 11 illustrates the driving simulator functions, which consisted of control and emulation, vehicle movement control, scenario generation, audio generation, meter control, and tactile sense simulation.

The visual display technology and high-fidelity audio system enhance the driving experience. The test subject is immersed in sight, sound, and movement environments so real that impending crash scenarios can be convincingly presented with no danger to the subject. These are necessary for the drowsy driving experiments. The vehicle and driver data is accurately gathered and stored by the system.

**Table 1 Definition of measures to characterize eyelid motion**

| Measure | Definition |
|---|---|
| PERCLOS | Percentage of eyelid closure, $(t_1+t_2+\cdots+t_n)/T$ |
| MCD | Maximum duration of eye closure, Max $(t_1, t_2, \cdots, t_n)$ |
| BF | The number of blinks per minute, $n/T$ |
| AOL | Average eye opening level |
| OV | Opening velocity |
| CV | Closing velocity |

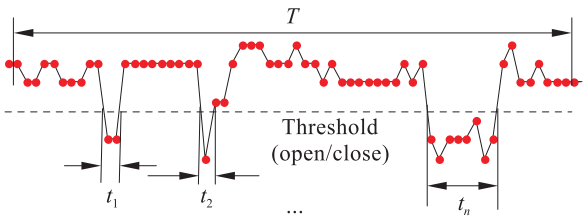



**Fig. 9 Parameters in the definitions of PERCLOS, MCD, and BF**

**Fig. 10 Driving simulator**

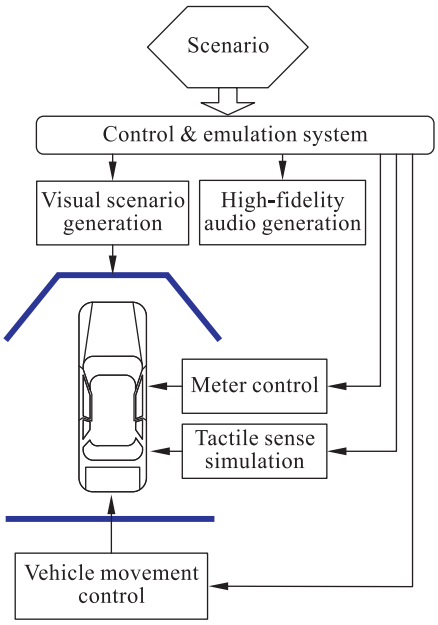



**Fig. 11 Driving simulator block diagram**

## 3.2 Test setup

A camera was mounted on the dashboard inside the vehicle. The camera is stationary with no position or zoom adjustments during operation. The grabbed frames are represented in Red-Green-Blue (RGB) space with 8 bit pixels (256 colors).

In the tests, the driver drowsiness was manipulated by increasing the duration of the driving task. The driving scenario was particularly monotonous and the driver was asked to follow another car on four long, straight roads that provided little visual stimulation. An

example of the scenario is shown in Fig. 12.

Six participants were involved in the tests (all men, 24 to 46 years old, with an average driving experience of 6.5 years).

### 3.3 Criteria

Various criteria were used in the training and testing sets to evaluate the drowsiness using self-rating measures and video-based rating measures.

The self-rating drowsiness measures involved some form of introspective assessment by the driver. The 7-point Stanford Sleepiness Scale (SSS) and the 9-point Karolinska Sleepiness Scale (KSS) are the two most commonly used sleepiness scales in the literature[19]. Their limitation is that the process of asking for ratings may stimulate the driver and, therefore, reduce the level of drowsiness that is being measured. Video-based rating measures, known as the Johns Drowsiness Scale (JDS), are performed by experts off-line using a number of variables relating to eyelid movements[20].

For the video-based rating measures used here, the video was split into non-overlapping intervals 20 s in duration with the sleepiness evaluated on each recording sample by three experts using the scale described in Table 2. The drowsiness scale involves 3 levels with awake (0 point), drowsy (1 point), and very drowsy (2 points), as shown in Table 2. The average of the three experts' scores for the same sample was used as the driver's actual state of drowsiness.

## 4 Test Results

The tests gave 225 samples of level 0, 181 samples of level 1, and 158 samples of level 2. Statistical analyses of each index are shown in Fig. 13.

The classification results of each index are shown in Table 3.

**Fig. 12 Example of the driving scenario**

**Table 2 Video-based drowsiness rating scale**

| Drowsiness level | Drowsiness state | Video image indicators |
|---|---|---|
| 0 | Awake | Normal fast eye blinks; active eyeball movement; apparent focus on driving with occasional fast sideways glances; normal facial tone |
| 1 | Drowsy | Increase in duration of eye blinks; appearance of "glazed-eye" look; abrupt irregular movements-rubbing face/eyes; moving restlessly on the seat; occasional yawning |
| 2 | Very drowsy | Occasional disruption of eye focus; significant increase in eye blink duration; reduction of the degree of eye opening; occasional disappearance of facial tone; episodes without any body movements |

**Table 3 Classification accuracy of each index separately (%)**

| | Actual drowsiness level | Predicted drowsiness level 0 | 1 | 2 |
|---|---|---|---|---|
| PERCLOS | 0 | 94.2 | 5.3 | 4.0 |
| | 1 | 16.0 | 76.2 | 7.7 |
| | 2 | 0 | 26.6 | 73.4 |
| MCD | 0 | 31.6 | 66.2 | 2.2 |
| | 1 | 1.7 | 89.5 | 8.8 |
| | 2 | 1.3 | 50.0 | 48.7 |
| AOL | 0 | 52.4 | 40.9 | 6.7 |
| | 1 | 40.3 | 54.7 | 5.0 |
| | 2 | 36.1 | 10.8 | 53.2 |
| BF | 0 | 88.9 | 11.1 | 0 |
| | 1 | 14.4 | 77.9 | 7.7 |
| | 2 | 0 | 21.5 | 78.5 |
| OV | 0 | 24.9 | 35.1 | 40.0 |
| | 1 | 16.6 | 68.5 | 14.9 |
| | 2 | 36.1 | 14.6 | 49.4 |
| CV | 0 | 18.7 | 47.6 | 33.8 |
| | 1 | 13.8 | 58.6 | 27.6 |
| | 2 | 17.7 | 14.6 | 67.7 |

Fisher's linear discriminant functions (Function 0, Function 1, and Function 2) were then used to combine the indices. However, use of all the measures may not give the best classifying performance due to redundancy.

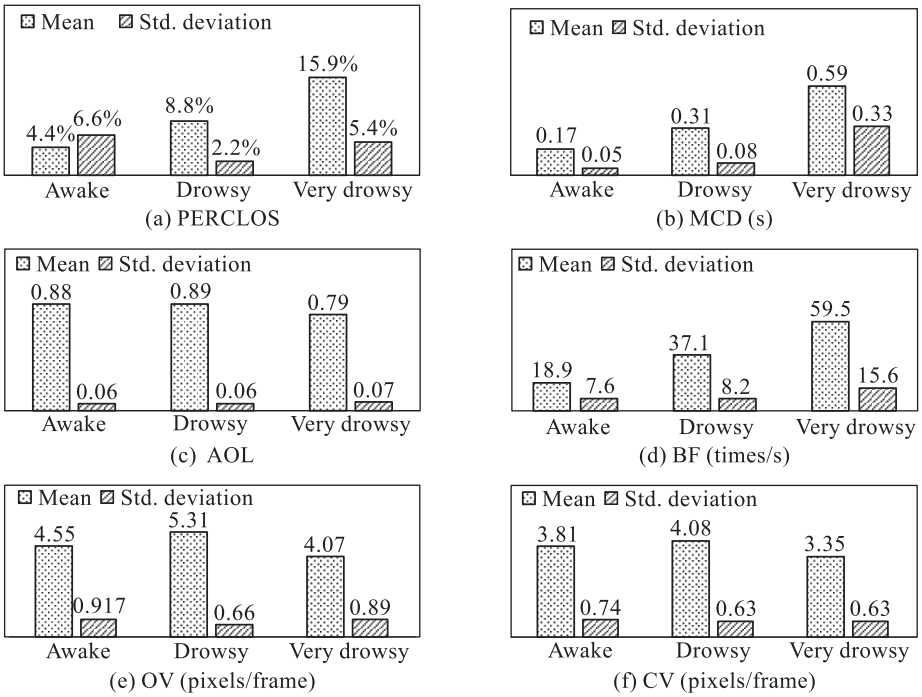



**Fig. 13 Statistical analysis of each drowsiness index**

Thus, the classification capability of each measure in the functions was quantified using a stepwise method using the Mahalanobis distance as the criterion. The analysis finds the variable that maximizes the Mahalanobis distance between the two closest drowsiness levels in each step. Table 4 shows the screening process of the measures using the software SPSS 16.0 and five steps are included.

**Table 4 Variables in the analysis**

| Step | Variables | *F* value | Drowsiness level |
|---|---|---|---|
| 1 | BF | 674.877 | |
| 2 | BF | 662.513 | 0 and 2 |
| | OV | 64.465 | 0 and 1 |
| 3 | BF | 650.399 | 0 and 2 |
| | OV | 37.514 | 0 and 1 |
| | CV | 8.143 | 0 and 1 |
| 4 | BF | 598.486 | 0 and 1 |
| | OV | 35.925 | 0 and 1 |
| | CV | 11.042 | 0 and 1 |
| | MCD | 14.730 | 0 and 1 |
| 5 | BF | 462.479 | 0 and 1 |
| | OV | 24.180 | 0 and 1 |
| | CV | 9.986 | 0 and 1 |
| | MCD | 10.624 | 0 and 1 |
| | AOL | 8.722 | 0 and 1 |

BF, OV, CV, MCD, and AOL were each introduced into the discriminant functions (Function 0, Function 1, and Function 2) in order. Although PERCLOS has a better classification capability than any of the other parameters, it was not introduced into the functions because of the correlations with other parameters. Thus, introducing PERCLOS will not further improve the classifier performance when the other parameters are used.

The discriminant analysis finds functions of the data that optimally discriminate between the three drowsiness levels. A canonical analysis is used to automatically determine the combinations of variables such that the first function provides the best overall discrimination between groups and the second provides the second best discrimination. The standardized canonical discriminant function coefficients are shown in Table 5.

Figure 14 gives a picture of the sample distribution in the canonical discriminant function coordinates with Functions 1 and 2 as the canonical discriminant functions.

**Table 5 Standardized canonical discriminant function coefficients**

| | MCD | AOL | BF | OV | CV |
|---|---|---|---|---|---|
| Function 1 | 0.234 | –0.132 | 0.938 | 0.105 | –0.274 |
| Function 2 | 0.031 | 0.389 | 0.223 | 0.864 | –0.186 |

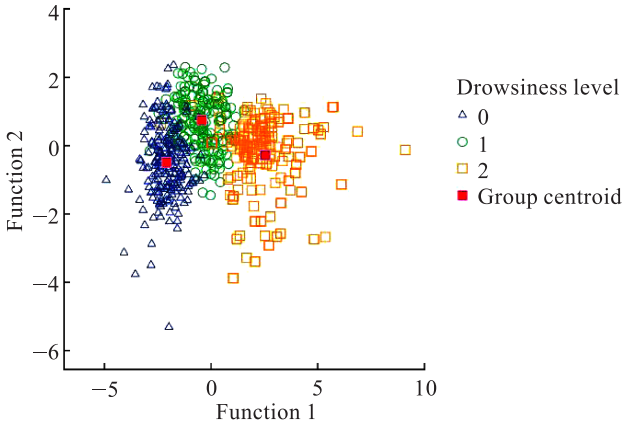



**Fig. 14 Sample distributions**

Fisher's linear discriminant functions were then used as the classification function. The discriminant analysis was performed using SPSS, which reports a linear discriminant function for each group. Their coefficients are listed in Table 6 with the classification results listed in Table 7.

The drowsiness Recognition Rate (RR) of 86% was calculated by

$$\mathrm{RR} = (N_{00} + N_{11} + N_{22})/N \times 100\% \quad (4)$$

The False Alarm Rate (FAR) of no more than 5% was calculated by

$$\mathrm{FAR} = (N_{01} + N_{02})/N \times 100\% \quad (5)$$

where $N$ is the total number of the samples and $N_{ij}$ is the number of samples predicted as $j$ which have the actual state $i$. Here, each sample is a video clip of 20 s in duration with the sleepiness evaluated using the method introduced in Section 3.3.

## 5 Conclusions and Future Work

Eye-tracking and image processing were used in a nonintrusive drowsiness recognition method. The problems caused by changes of illumination and driver posture that influence the eye detection algorithm were reduced by using a self-quotient image as input to the ASM statistical texture model and to provide a density distribution image to the mean-shift eye-tracking algorithm. The six measures used to evaluate the drowsiness are the PERCLOS, MCD, BF, AOL, OV of eyes, and CV of the eyes. The quantitative relationships between drowsiness and these measures were evaluated using Fisher's linear discriminant functions to reduce their correlations to compare the classification capability. Tests with six participants in a high fidelity driving simulator gave a driver drowsiness recognition accuracy greater than 86%.

**Table 6 Classification function coefficients**

| | MCD | AOL | BF | OV | CV | Constant |
|---|---|---|---|---|---|---|
| Function 0 | 20.300 | 262.488 | 0.405 | −2.714 | 1.239 | −118.268 |
| Function 1 | 22.200 | 265.618 | 0.571 | −1.395 | 0.280 | −128.720 |
| Function 2 | 24.696 | 253.952 | 0.766 | −2.169 | −0.436 | −122.806 |

**Table 7 Classification results**

| Actual drowsiness level | Predicted drowsiness level | | |
|---|---|---|---|
| | 0 | 1 | 2 |
| 0 | 198 | 27 | 0 |
| 1 | 17 | 151 | 13 |
| 2 | 1 | 19 | 138 |

Drowsiness characteristics differ for different individuals with people having different blink frequencies, the introduction of personal features for a specific driver in detecting his/her drowsiness could improve the drowsiness recognition rate. An individual model using data during the beginning of the driving task (e.g., in the first 20 min) when the driver is alert could be used to develop individualized evaluations. Combining general criteria developed from many participants can also improve the drowsiness recognition accuracy.

## References

[1] Brown I. Driver fatigue. *Human Factors*, 1994, **36**(2): 298-314.

[2] Dinges D, Mallis M, Maislin G, et al. Final report: Evaluation of techniques for ocular measurement as an index of fatigue and the basis for alertness management. NHTSA, Washington DC, Tech. Rep. DOT HS 808762, 1998.

[3] Yang G, Lin Y. Using ECG bio-signal to quantify mental workload based on wavelet transform and competitive neural network. *International Journal of Biomedical Soft Computing and Human Sciences*, 2009, **14**(2): 17-25.

[4] Lal S K, Craig A. Reproducibility of the spectral components of the electroencephalogram during driver fatigue. *Int. J. Psychophysiol.*, 2005, **55**(2): 137-143.

[5] Lal S K, Craig A. Driver fatigue: Electroencephalography and psychological assessment. *Psychophysiology*, 2002, **39**(3): 313-321.

[6] Hayashi K, Ishihara K, Hashimoto H, et al. Individualized drowsiness detection during driving by pulse wave analysis with neural network. In: Proceedings of the 8th International IEEE Conference on Intelligent Transportation

Systems. Vienna, Austria, 2005: 901-906.

[7] Yang G, Lin Y, Bhattacharya P. A driver fatigue recognition model based on information fusion and dynamic bayesian network. *Information Sciences*，2010, **180**(10): 1942-1954.

[8] Wierwille W, Ellsworth L. Evaluation of driver drowsiness by trained raters. *Accid. Anal. Prev.*, 1994, **26**(5): 571-581.

[9] Ji Q, Zhu Z, Lan P. Real-time nonintrusive monitoring and prediction of driver fatigue. *IEEE Transactions on Vehicular Technology*, 2004, **53**(4): 1052-1068.

[10] Lenne M, Triggs T, Redman J. Time of day variations in driving performance. *Accident Analysis & Prevention*, 1997, **29**(4): 431-437.

[11] Thiffault P, Bergeron J. Fatigue and individual differences in monotonous simulated driving. *Personality and Individual Differences*, 2003, **34**(1): 159-176.

[12] Stephen H. Impairment of driving performance caused by sleep deprivation or alcohol: A comparative study. *Human Factors*, 1999, **41**(1): 118-128.

[13] Lee M, Ranganath S. Pose-invariant face recognition using a 3D deformable model. *Pattern Recognition*, 2003, **36**(8): 1835-1846.

[14] Phillips P, Moon H, Rizvi S, et al. The FERET evaluation methodology for face-recognition algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2000, **22**(10): 1090-1104.

[15] Yang M H, Kriegman D, Ahuja N. Detecting faces in images: A survey. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2002, **24**(1): 34-58.

[16] Viola P, Jones M. Rapid objects detection using a boosted cascade of simple features. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Hawaii, USA, 2001: 511-518.

[17] Cootes T F, Taylor C J, Cooper D H, et al. Active shape models–their training and application. *Computer Vision and Image Understanding*, 1995, **61**(1): 38-59.

[18] Yuille A, Hallinan P, Cohen D. Feature extraction from faces using deformable templates. *International Journal of Computer Vision*, 1992, **8**(2): 99-111.

[19] Liu C C, Hosking S G, Lenné M G. Predicting driver drowsiness using vehicle measures: Recent insights and future challenges. *Journal of Safety Research*, 2009, **40**(4): 239-245.

[20] Johns M W, Tucker A, Chapman R, et al. Monitoring eye and eyelid movements by infrared reflectance oculography to measure drowsiness in drivers. *Somnologie*, 2007, **11**(4): 234-242.