Expert Systems with Applications 36 (2009) 7651–7658

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Driver drowsiness detection with eyelid related parameters by Support Vector Machine

Hu Shuyan *, Zheng Gangtie

School of Astronautics, Beijing University of Aeronautics and Astronautics, 37 Xueyuan Road, Haidian District, Beijing 100083, China

ARTICLE INFO

Keywords:
SVM
Physiological signal
Paired *t*-test
Driver drowsiness prediction

ABSTRACT

Various investigations show that drivers' drowsiness is one of the main causes of traffic accidents. Thus, countermeasure device is currently required in many fields for sleepiness related accident prevention. This paper intends to perform the drowsiness prediction by employing Support Vector Machine (SVM) with eyelid related parameters extracted from EOG data collected in a driving simulator provided by EU Project SENSATION. The dataset is firstly divided into three incremental drowsiness levels, and then a paired *t*-test is done to identify how the parameters are associated with drivers' sleepy condition. With all the features, a SVM drowsiness detection model is constructed. The validation results show that the drowsiness detection accuracy is quite high especially when the subjects are very sleepy.

1. Introduction

Safe driving is a major concern of societies all over the world. Thousands of people are killed, or seriously injured due to drivers falling asleep at the wheels each year. Recent studies show that drivers' drowsiness accounts for up to 20% of serious or fatal accidents on motorways and monotonous roads, which impairs the drivers' judgment and their ability of controlling vehicles (Driver fatigue, 2006). Therefore, it is essential to develop a safety system for drowsiness-related road accident prevention.

Many methods have been developed and some of them are currently being used for detecting the driver's drowsiness (Hamada, Ito, Adachi, Nakano, & Yamamoto, 2004; Hayami, Matsunaga, Shidoji, & Matsuki, 2002; Ji & Yang, 2002; Liu, Xu, & Fujimura, 2002; Wierwille, Lewin, & Fairbanks, 1996), including the measurements of physiological features like EEG, heart rate and pulse rate, eyelid movement, gaze, head movement and behaviors of the vehicle, such as lane deviations and steering movements. Among those different technologies, ocular measures, such as eye-blinking and eyelid closure, are considered as promising ways for monitoring alertness. A system was described in Ueno, Kaneda, and Tsukino (1994) for drowsiness detection by recognizing whether a driver's eyes are open or closed, and, if open, computing the degree of openness. The study showed that the system performance is comparable to those of techniques using physiological signals. A system called FaceLAB developed by a company Seeing Machines was presented in Seeing Machines (2004), which could estimate the driver's fatigue level by monitoring eyelids and determining eye opening and blink rates. In Boverie, Leqellec, and Hirl (1998), a system was developed for monitoring driving vigilance by studying the eyelid movement. A preliminary evaluation revealed promising results of using eyelid movement for characterizing a driver's vigilance level. Another system called Copilot was presented in Grace (2001), which used a simple subtraction process for finding the eyes and calculating a validated parameter called the percent eye closure (PERCLOS) to measure a driver's drowsiness.

However, previous systems mainly focus on using a single eyelid movement indicator to detect driver's drowsiness and therefore may suffer from low robustness and low accuracy. Moreover, how various eyelid movement features jointly reflect the change of drivers' drowsiness level, i.e., how eyelid movement features can be used to accurately and robustly predict drowsiness, has not yet been well investigated. Therefore, it is necessary to study the relations between the driver drowsiness degree and eyelid movement features.

The purpose of this paper is to study the approach of using multiple eyelid movement features to detect the drivers' drowsiness with a newly developed machine learning technique – Support Vector Machine (SVM) to detect the drowsiness of the drivers with multiple eyelid movement features. With more than one features, it is expected that both the precision and the robustness of predication could be significantly improved. In this paper, the data is collected from a simulated driving experiment conducted in VTI (Swedish National Road and Transport Research Institute), Sweden within the EU project SENSATION.

The paper is organized as follows: in the section below, first the SVM is reviewed for classification problems; secondly, the experiment is introduced; and then the application of SVM to predict driver drowsiness with the eyelid movement parameters from

* Corresponding author. Tel.: +86 10 8233 8091.
E-mail address: huyan@sa.buaa.edu.cn (S. Hu).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.09.030

physiological signals is presented. Finally the results are discussed and conclusions are derived, respectively.

2. Support Vector Machine

Support Vector Machine (SVM), a novel machine learning algorithm, has been recently proven to be a promising tool for both data classification and pattern recognition. SVM is a particular learning system that based on the margin-maximization principle (See Fig. 1). This learning strategy introduced by Vapnik and coworkers is a powerful method that in the few years since its introduction has already been successfully applied to a wide variety of applications ranging from particle identification, face identification and text categorization to intrusion detection, bioinformatics and database marketing (SVM application list, 2006). Here we only give a very brief introduction to SVM, more detailed descriptions can be found in Burges (1998), Campbell (2000), Cristiannini and Shawe-Taylor (2000).

Given a binary classification task with training datasets

$$(x_1, y_1), \ldots, (x_l, y_l), \quad \mathbf{x} \in R^n, \quad y \in \{+1, -1\}$$

the separating hyperplane will be $\mathbf{w}^T \cdot \mathbf{X} + b = 0$. Therefore, the decision function will be $f(\mathbf{X}) = \mathrm{sign}(\mathbf{w}^T \cdot \mathbf{X} + b = 0)$ and after a positive rescaling of the argument of the sign-function the task is to find $\mathbf{w}$, b with the maximal distance between $\mathbf{w}^T \cdot \mathbf{x} = \pm 1$ the training dataset is linear separable, we will obtain the following primal problem:

$$\begin{aligned} &\min_{\mathbf{w},\mathbf{b}} && \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} \\ &\text{subject to} && y_i(\mathbf{w}^T \cdot \mathbf{x_i} + b) \geqslant 1 \end{aligned} \tag{1}$$

This optimization problem can be transformed into its corresponding dual problem as follows:

$$L(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2}\mathbf{w}^T\mathbf{w} - \sum_{i=1}^{l} \alpha_i [y_i(\mathbf{w} \cdot \mathbf{x} + b) - 1] \tag{2}$$

where $\alpha_i \geqslant 0$ are the Lagrange multipliers. Differentiating with respect to b and $\mathbf{w}$ resubstituting back into the primal problem, we can obtain

$$\begin{aligned} &\max && W(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \\ &\text{subject to} && \sum_{i=1}^{l} y_i \alpha_i = 0, \quad \alpha_i \geqslant 0 \end{aligned} \tag{3}$$

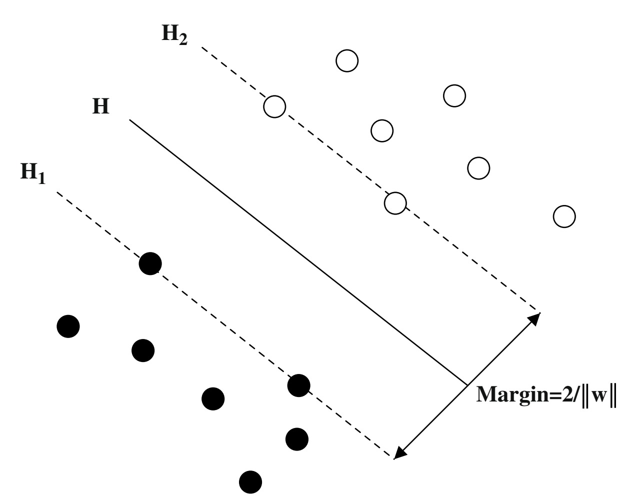


Fig. 1. A maximal margin hyperplane with its support vectors circled.

Table 1
Typical kernel functions

Kernel	$K(\mathbf{x}_i\mathbf{x}_j)$
Polynomial kernel	$K(\mathbf{x}_i\mathbf{x}_j) = (\mathbf{x}_i \cdot \mathbf{x}_j / a + b)^d$
Radial basis function kernel	$K(\mathbf{x}_i\mathbf{x}_j) = e^{-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2}$
Sigmoid kernel	$K(\mathbf{x}_i\mathbf{x}_j) = \tanh(\beta\mathbf{x}_i \cdot \mathbf{x}_j + b)$

The input data are mapped into a higher dimensional Hilbert space $\mathbf{x}_i \cdot \mathbf{x}_j \to \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)$, and then the kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}_j)$ is introduced here so the functional form of the mapping $\phi(\mathbf{x}_i)$ does not need to be known. With a suitable choice of kernel, the data can become separable in the feature space despite being non-separable in the original input space. Three typical kernel functions are listed in Table 1. Then the optimization problem of Eq. (3) becomes

$$\max W(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i \cdot \mathbf{x}_j) \tag{4}$$

Typically in real-world application, the training datasets may not be separable due to noise in the data. When the classes are non-separable, a soft margin Support Vector Machine which takes advantage of more robust measures is created. The robust measure introduces slack variables $\xi_i \geqslant 0,\ i = 1, \ldots, l$ into the constraints which then become:

$$y_i(w^T \Phi(x_i) + b) \geqslant 1 - \xi_i \quad \xi_i \geqslant 0,\ i = 1, \ldots, l. \tag{5}$$

Thus, $\sum_{i=1}^{l} \xi_i$ is an upper bound of the number of training errors and the objective function is changed to $\min_{\mathbf{w},\mathbf{b},\xi} \frac{1}{2}\mathbf{w}^{\mathrm{T}}\mathbf{w} + C\sum_{i=1}^{l} \xi_i$, where C is a parameter to be chosen by the user and corresponds to the assignment of penalty to errors.

From the above introduction and all the related literatures, we will see that SVM has several merits:

1. It has a strong theoretical foundation, motivated by statistical learning theory (SLT). According to SLT, the bound on the actual risk is the sum of the empirical risk and the confidence interval, i.e., $R(\alpha) \leqslant R_{\mathrm{emp}}(\alpha) + \Phi(\frac{h}{l})$, where l means the number of the datasets and h is the VC dimension (for Vapnik–Chervonenkis dimension), a measure of the capacity of a statistical classification algorithm. We can see that the confidence interval increases with the VC dimension h SLT presents the structural risk minimization principle to minimize the actual risk with respect to both terms, the empirical risk and the confidence interval, which defines a trade-off between the quality of the approximation of the given data and the complexity of the approximating function. The idea of SVM is to minimize the second term while empirical risk remains unchanged (e.g. let the training error be zero), which is different from Empirical Risk Minimization (ERM) in traditional neural network. SLT also shows that larger margin hyperplanes have smaller VC dimension (Vapnik, 1998).
2. The model (classifier) constructed only depends on the support vectors.
3. The training task involves optimization of a convex cost function; therefore we will obtain a globe rather than false local minima in the learning process. Moreover, due to the introduction of kernel function all computations are performed directly in input space, which successfully avoids the curse of dimensionality.
4. There are relatively few free parameters for the model to adjust, i.e., C and the kernel parameters. For example, if RBF kernel is selected, there will be only two parameters C, γ to be controlled by the user.

3. Experiment

The eyelid related parameters used here are provided from an EU Project SENSATION experiment. The experiment was conducted in Sweden by VTI using their third generation moving base driving simulator as shown in Fig. 2. There were 37 sleep-deprived subjects in this dataset. They were night shift workers, recruited through an advertisement in a local newspaper (no professional drivers included), who arrived to the driving simulator just after their night shifts and all of them drove in one lane with rumble strips on each side. Hitting the rumble strip, either left or right, was considered to be a well defined critical event that would have led to an accident if the driver had not been warned. All subjects drove at least 45 min.

Horizontal and vertical (left and right) EOG were recorded with three channels and Electromyogram (EMG) was measured with one channel, as shown in Fig. 3. The sampling frequency was 512 Hz for EOG. Simultaneously, the EEG was measured with three bipolar derivations positioned at Fz-A1, Cz-A2 and Oz-Pz, as shown in Fig. 4. Silver cup electrodes were used for all signals. An additional marker channel was used to provide a synchronisation sign.

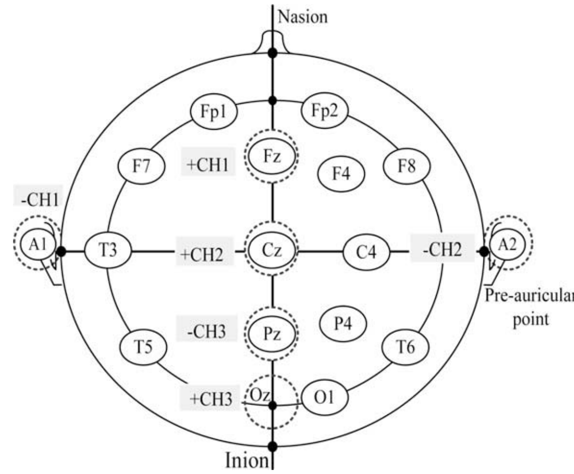


Fig. 4. EEG measurement positions.

Fig. 2. Moving base driving simulator at VTI.

The number of hits to the rumble strips varies from 0 to 54 among the 37 subjects during the course of driving. However, only data before the first hit for each subject are used in our training and validation which is consistent with the real situation that after their first accident the drivers are not able to or not supposed to drive again.

To evaluate the subjects' drowsiness, the Karolinska drowsiness score (KDS) and Karolinska Sleepiness Scale (KSS) are used as references here.

4. Karolinska Drowsiness Score (KDS) & Karolinska Sleepiness Scale (KSS)

Karolinska drowsiness score (KDS), proposed by Karolinska Institute (KI), is a method to score the drivers' drowsiness level which is done off-line.

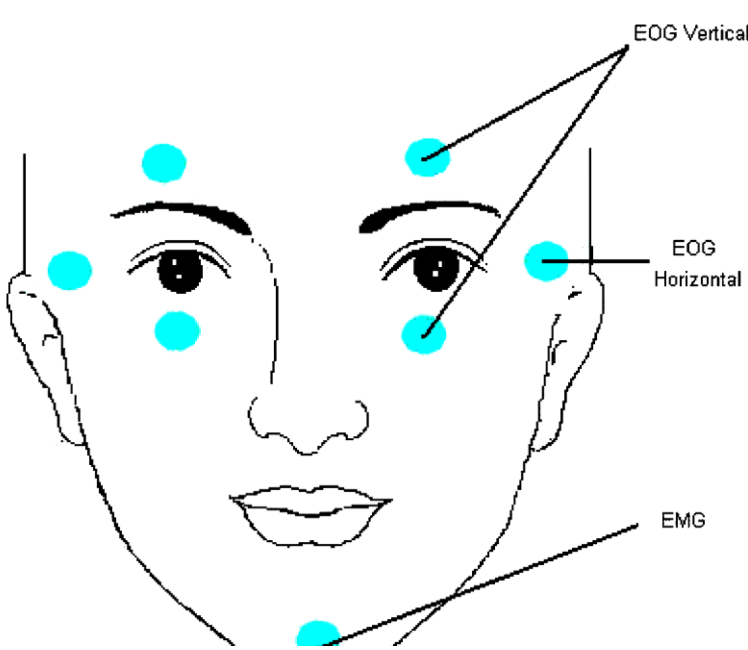


Fig. 3. EOG and EMG measurement.

This scoring method is developed for the quantification of sleepiness in "active" situations (e.g. situations when the subject should be awake). It is based mainly on EEG and EOG recordings. EMG may be a valuable indicator of arousals and artefacts but is not included among the criteria for drowsiness. The scoring method is based on Rechtschaffen & Kales (1968) scoring rules. The data is divided into 20 s epochs. Then the entire epoch is separated into 2 s bins and the outcome describes how many bins in the epoch contain slow eye movements, alpha activity and/or theta activity (i.e., sign of sleepiness). If no sleepiness signs are found, the epoch is assigned the value 0 (= 0% of sleepiness). If one 2-s bin includes signs of sleepiness, the epoch is assigned the value 10 (meaning that the epoch contains 10% of sleepiness). If the epoch contains two bins with sleepiness, it is assigned 20 (20% of sleepiness). The outcome includes 11 steps and starts with 0, followed by 10, 20 and so on. The maximum level is 100%, which means that all bins include sleepiness signs. In some cases, KDS >50% have been assigned the value 'sleep onset'.

Subjective ratings of sleepiness are also obtained using a modified version of Karolinska Sleepiness Scale (KSS) (Åkerstedt & Gillberg, 1990). The scale consists of a nine-point scale with verbal anchors for each step, see Table 2. Furthermore, the subjects were allowed to use intermediate steps of 0.5 as well.

The participants rated their level of sleepiness before driving at three occasions (see Table 3) and then during each driving session, the subjects give a verbal rating of sleepiness through the intercom every 5 min. One subject rated the level of sleepiness as 10 and actually fell asleep. After the driving they rated their sleepiness, too.

Table 2
Karolinska sleepiness scale (KSS)

Rate	Verbal descriptions
1	Extremely alert
2	Very alert
3	Alert
4	Fairly alert
5	Neither alert nor sleepy
6	Some signs of sleepiness
7	Sleepy, but no effort to keep alert
8	Sleepy, some effort to keep alert
9	Very sleepy, great effort to keep alert, fighting sleep

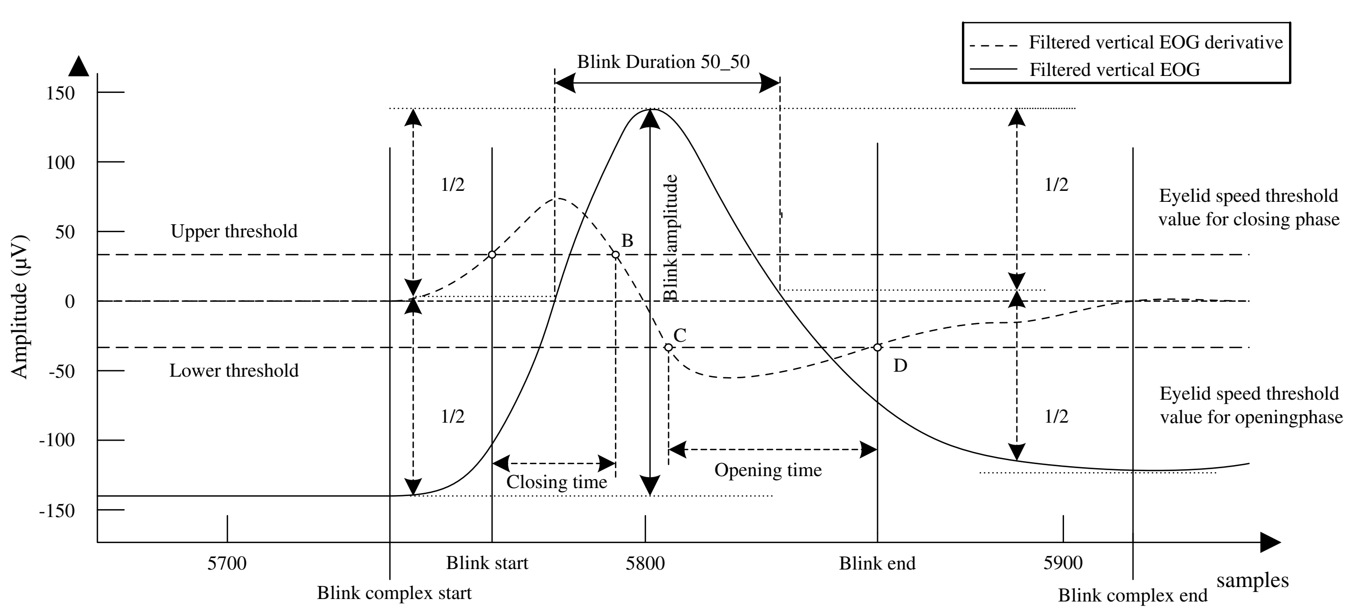
Table 3
KSS rating data for the selected 38 subjects before and after driving

Event	KSS rating mean (min–max), SD
Arriving at VTI	5.3 (4.0–7.5), 1.1
Before biocalibration	5.9 (4.0–8.0), 1.0
After pupillometry	7.2 (4.0–9.0), 1.1
After driving	7.1 (4.0–9.0) 1.2

5. Features extraction from EOG

From the EOG experimental data, eyelid movement features needed for SVM training and validation are extracted. The first stage is to identify eye blinks from EOG data. In order not to lose any blinks, the two EOG vertical channels are combined into one by averaging the two signals. This combined signal is then analysed and approximately 95% of the blinks are correctly identified, which is considered adequate for the analysis in the next section. The algorithm consists of the following steps:

Step 1: Define the threshold parameters- the minimum duration, threshold of eyelid closing speed, threshold of amplitude, etc.
Step 2: Filter high frequency of EOG signal.
Step 3: Compute the derivative (difference) of the filtered signal.
Step 4: Apply a threshold to select blinks. In this step, the blinks and the zone, in which the derivative of the signal (i.e., speed) is abnormally larger than corresponding thresholds, are detected.
Step 5: Test the possibility of long blink. With a new detected blink, search in the neighborhood for the next blink and then verify whether the two blinks form one long blink.
Step 6: Verify some constraints, including the conditions of maximum amplitude of closing phase and the opening phase, maximum duration, to validate the detected blinks as true blinks.



(A:start pOsition of blink B: the moment wheneyelidfinishes closing C: the moment when eyelid begins opening D: end position of the blink)

Fig. 5. The blink feature definition.

Table 4
The eyelid movement parameters

Variables	Description	Comment
X_1	Blink duration [s]	Calculated from the start position of blink to the end value of blink
X_2	Blink duration 50_50 [s]	Calculated from the half rise amplitude to half fall point
X_3	Amplitude [μV]	Calculated from the start of blink to the peak of this blink
X_4	Lid closure speed [μV/sample]	The average velocity in closing phase
X_5	PCV [μV/sample]	The peak closing velocity
X_6	Lid opening speed [μV/sample]	The average velocity in opening phase
X_7	POV [μV/sample]	The peak opening velocity
X_8	Delay of eyelid reopening [s]	The reopening time starts at the moment the eyelid begins its main movement in the opening direction. As a stopping point we use the moment of highest velocity during the opening.
X_9	Duration at 80% [s]	Calculated from the 80% rise amplitude to 20% fall amplitude
X_{10}	Closing time [s]	
X_{11}	Opening time [s]	

Secondly, with the above algorithm, the position parameters of the validated blink are acquired as shown in Fig. 5 and all extracted features are listed in Table 4.

6. Training/validation datasets selection and processing

Like many other machine learning algorithms, the quality of the data used in the training process is also critical to the successful application of the SVM. Therefore, in this part, we will study the selection of training/validation datasets and the processing of the datasets.

In order to smooth the data, we start with the average of all features on a 20-s which also corresponds to the KDS epoch. Then we consider classifying the sleepiness of the drivers mainly based on KDS and KSS into three incremental levels: alert (marked as label 1), sleepy (marked as label 2) and very sleepy (marked as label 3). When a subject is ‘sleepy’, the system should sound alarms to make the driver awake. When a subject is ‘very sleepy’, it is very dangerous if the driver carries on his driving since he puts himself at high risk of being involved in a serious car accident. In some cases the engine must be stopped.

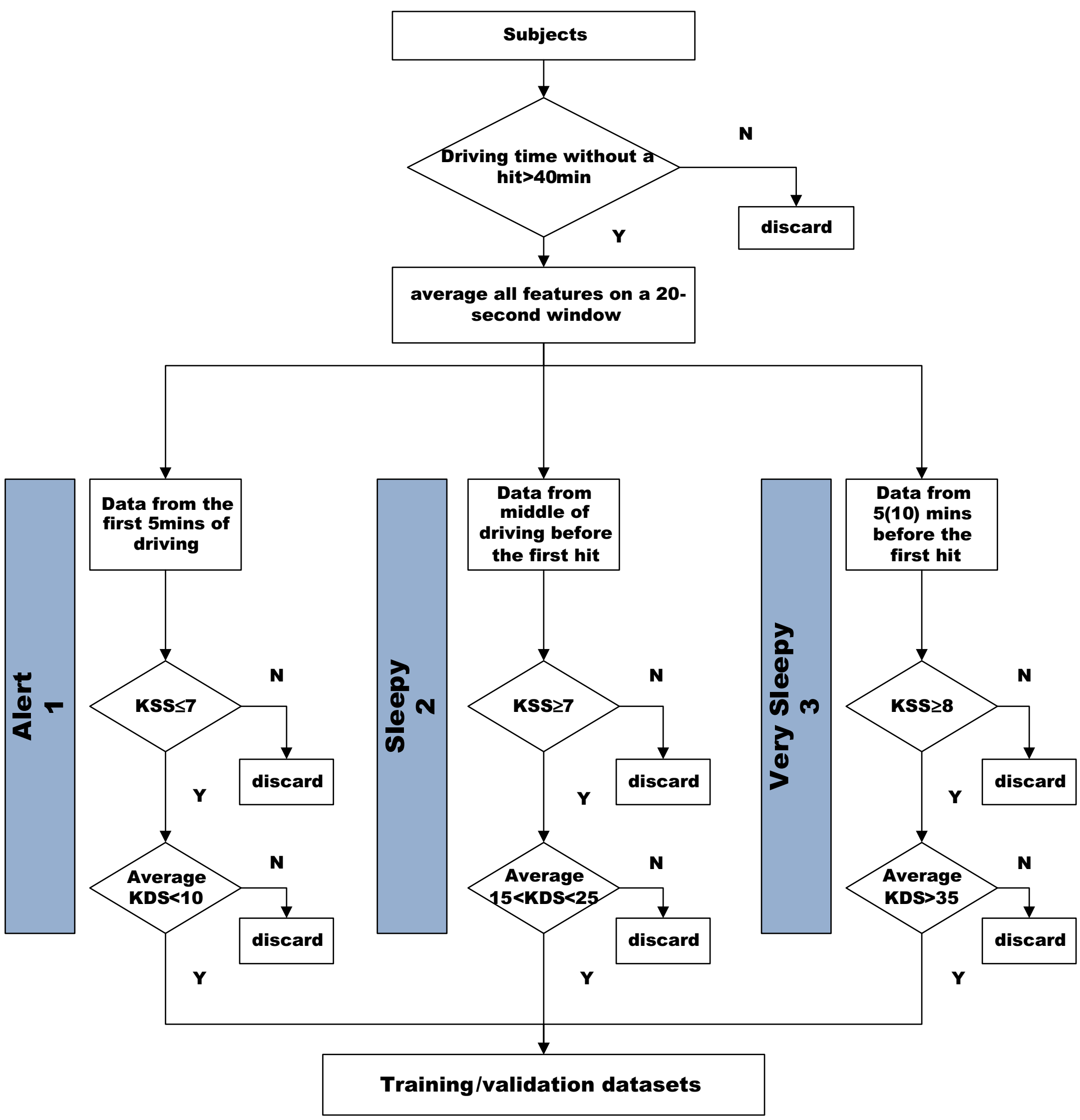


Fig. 6. The flow chart for selecting the training/validation dataset

Table 5
Training and validation datasets

Sex	Training			Validation	
	Female	Male	Female	Male	Male
Hit time (min)	52	50	44	47	82
Average KDS	1 (alert)	6.67 (45 samples)		6.33 (30 samples)	
	2 (sleepy)	18.06 (45 samples)		25 (30 samples)	
	3 (very sleepy)	43.11 (45 samples)		48.67 (30 samples)	

The training/validation samples selected with drowsiness levels (1, 2 and 3) are based on the rules contained in the flow chart as shown in Fig. 6.

Here we choose five subjects for training and validation who drove for at least 40 min without any hits. Data of those who hit the strips immediately are not included because we need to investigate the three different levels of sleepiness for an adequate amount of time. Based on the dataset selection rules, KDS and KSS are checked out to make sure every hit of the subjects we selected are due to drowsiness and also at the beginning of driving the subjects are alert. Detailed information about the selected five subjects is in Table 5.

Since there are five subjects are involved in the training/validation task, it is unavoidable that individual difference will bias the validation results. Therefore, proper data processing method of data should be employed to improve the prediction performance. In this paper, we attempt to scale all the features to $[0,1]$ with following standardization procedure:

$$x_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{6}$$

where x_i denotes the ith eyelid movement feature extracted from EOG.

In this study, on account of the fact that the experiment was conducted in early mornings with all subjects being night shift workers, it is safe to assume that the hits of the rumble strips are only caused by sleepiness. Therefore, with all the eyelid movement-related features obtained from three different drowsiness levels, besides the random error the changes of the feature values are due to drowsiness, i.e., if the features for different drowsiness levels differ, we can say these features are differentially expressed for the three levels. In order to investigate how these eyelid movement-related features extracted from EOG are associated with the sleepy conditions, a paired t-test is performed.

The paired t-test can be used to determine whether or not the means of two features ($\overline{X_1}$ and $\overline{X_2}$) for two different drowsiness levels are equal in a significant way. It computes the differences between values of the two features and tests whether the average equals to zero under the assumptions that the paired differences are independent and identically normally distributed. For the paired t-test, the computed test statistic, $t = \frac{\overline{D}-d_0}{S_D}$, $(\overline{X_1} - \overline{X_2} = d_0)$, has a Student's t-distribution with $v = n-1$ degrees of freedom. For groups 1 and 2 and data points $j = 1, 2, 3, \ldots, n$, $D = X_{1j} - X_{2j}$, $\overline{D} = \frac{1}{n}\sum_{j=1}^{n} D_j$ and '$S_D = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(D_j - \overline{D})^2}$.

Given the t-value and the degree of freedom, a p-value can be found using a table of values from Student's t-distribution. If the p-value is less than the chosen statistical significance (usually 0.05 or 0.01), the null hypothesis that there is no significance difference between two groups is rejected in favor of an alternative hypothesis, which typically states that the groups are different and the features are statistically differential.

Firstly, we should make sure that the differences of two groups are from normal distribution. Here the One-Sample Kolmogorov–Smirnov Test is employed for this purpose. Moreover, we perform the paired t-test for feature selection only on the training dataset so as not to bias the validation results. All statistical calculations are conducted by SPSS. The one sample K-S test results are shown in Table 6.

As we can see from table 6, all the differences come from normal distribution at the 0.01 level. The paired t-test results are listed in Table 7.

As it can be seen that *p-values* (the t-value significance levels) of X_2–X_7, X_9 and X_{10} are below 0.05 (threshold of t-value for significance level 0.05 is 2.01), which indicates there are highly statistically significant increase/decrease in the seven eyelid movement-related parameters extracted from EOG. As to X_1 (i.e., duration of the blink), the first *p-value* is 0.13, which is a chance that 'duration' may not be a statistically distinguished feature in our training dataset when we do Alert–Sleepy classification. However, it plays a role in the Sleepy–Very Sleepy classification with p-value of 0.0027, so it cannot be ruled out. Same conclusions will be drawn for X_8 and X_{11}. Therefore, all these features will be used in our SVM model. It should be noticed that the bigger the $|t|$ value is, the more differential the feature is, and also more useful the feature is in the classification task.

7. SVM experiment results

The SVM model (classifier) we constructed in this paper will be trained with the 11 eyelid related features extracted from EOG with three different drowsiness levels (1, 2 and 3), and then it will be used to detect the driver's condition with the validation dataset. Multi-class implementation of Libsvm (Chang & Lin, 2001), which

Table 6
Significance levels of K-S test

	Significance level										
Differences of features	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
Alert–sleepy	0.55	0.39	0.99	0.76	0.98	0.94	0.70	0.06	0.45	0.14	0.80
Alert–very sleepy	0.20	0.18	0.67	0.61	0.67	0.83	0.69	0.40	0.17	0.76	0.79
Sleepy–very sleepy	0.81	0.49	0.66	1.00	0.86	0.42	0.50	0.13	0.25	0.95	0.31

Table 7
Paired t-test results of eyelid movement-related features

Features		X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
Alert–sleepy	$\|t\|$	1.53	2.71	7.43	6.50	6.22	4.81	7.15	0.24	2.27	2.90	2.22
Alert–very sleepy	$\|t\|$	1.92	7.71	15.48	21.04	18.24	8.87	15.08	3.33	6.65	5.53	2.36
Sleepy–very sleepy	$\|t\|$	3.18	5.26	6.88	8.33	8.19	3.22	5.17	3.75	5.38	3.77	0.16

$v = 44$, $t_{(0.05)} = 2.01$, $t_{(0.01)} = 2.69$.

Table 8
Validation results

Validation results				
	1 Alert	2 Sleepy	3 Very sleepy	Correct %
1 Alert	25	3	2	83.33
2 Sleepy	0	26	4	86.67
3 Very sleepy	0	0	30	100

uses the one-against-one approach, is utilized for this purpose. In this experiment, RBF kernel functions are selected. In order to determine the parameters C and γ to yield the highest predictive accuracy, a five-fold cross-validation is done with Libsvm. After conducting the grid search on the training data, the optimal (C, γ) is $(4, 1)$ with a cross-validated accuracy of 80.74%.

Table 8 shows the prediction results of the three sleepiness levels. As we can see, all the 'very sleepy' conditions have been successfully detected which shows that a SVM model conducted with eyelid movement-related features could be a reliable tool for drowsiness detection. Up to 86.67% of the trials accurately detect when the subject is 'sleepy'. 16.67% of the trials which are supposed to be 'alert' are wrongly detected as 'sleepy'/'very sleepy'. However, we cannot easily claim that they are false alarms since the drivers are suffering from sleep deprivations. In practical applications, if the detections continuously are 'sleepy' or 'very sleepy', an alarm should be triggered.

We also give an all-driving-session-prediction of subject 05 compared to KDS from the beginning to his first hit every 20 s as illustrated in Fig. 7 and 8 where the lower parts of the displays indicate the prediction results of our SVM model.

In these figures, we can see that since the beginning (from start to 400 s) the subject has been quite sleepy with an average KDS above 20; therefore, 'sleepy'/'very sleepy' warnings have been given. As the driving session proceeds, the low average KDS indicates the driver regains his awareness slightly and the SVM outcomes are 'alert'/'sleepy'. The driver then goes through another two 'sleepy→less sleepy (or alert)' processes which can also be observed from the corresponding-period prediction results. The driver is always struggling to keep clear-headed when he is driving, as a result, these 'sleepy→alert→sleepy' fluctuations are anticipated both in the experiment and the prediction results.

Fig. 8 shows that there are considerable 'very sleepy' warnings even 20 m before the hit. Even though the drivers are sleepy or very sleepy, they might not immediately hit the strip. However, whether the driver hits the strips or not, warning measures must be done with sufficient time to warn the driver (we all get a positive warning when KDS >40); therefore, we cannot assume these warnings to be false alarms.

8. Future work

Driver drowsiness is believed as one of the main contributing causes of serious industrial and automobile accidents, which draws more and more attention of research communities from road safety to industrial environment protection. Among all the technologies, using eyelid movement features to detect the state of driver drowsiness is a reliable and promising approach. However, there are some issues remain to be discussed in the future work:

1. In this experiment, only sleep-deprived subjects are included and no data from alert driving condition. Moreover, simulator driving is not exactly the same as real driving even with a high fidelity simulator as the one used in this experiment. Field trials often end up with noisy data which might be difficult to analyze and require some careful considerations. It should be noted that our approach developed on the simulator data need to be validated on the real driving.

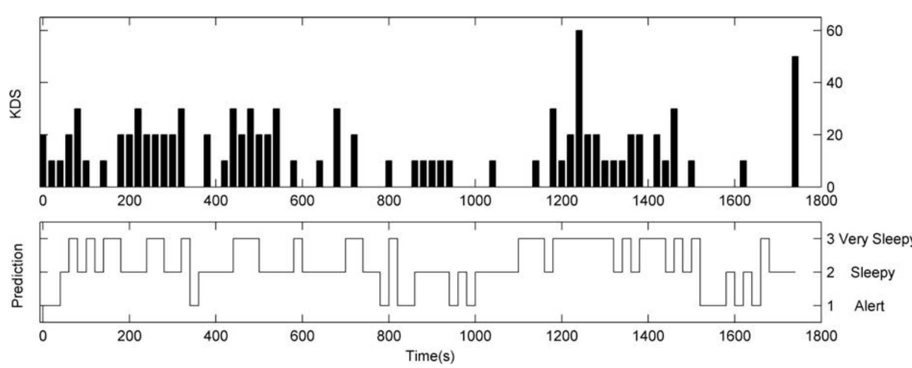


Fig. 7. Prediction results for the first 1780 s of the driving (subject 05).

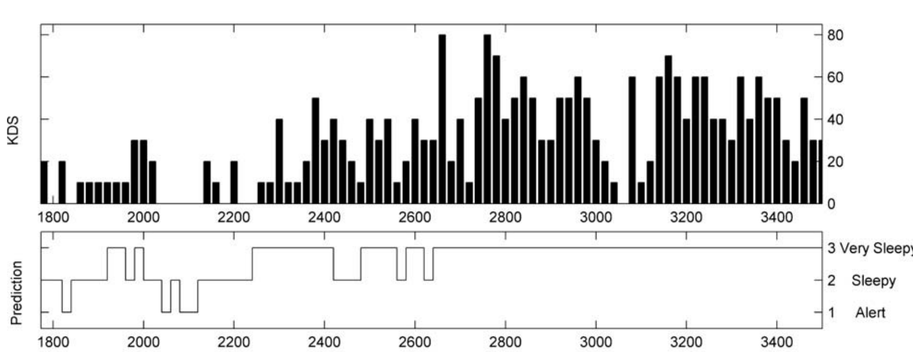


Fig. 8. Prediction results for 1720 s before the hit (subject 05).

2. Since driver drowsiness may lead to a major accident, a device with high drowsiness detection accuracy would be very useful as a safety aid for drivers to prevent potential hazards. On the other hand, in order to ensure the comfort of driving, low rate of false alarms should also be considered as an important factor in developing a method and/or device. In our training strategy, it is important that we evaluate of the subjects' drowsiness levels wisely, which could be critical to our two above goals. In this paper, KDS, KSS and hits of the strip have all been used as sleepy references. In real situations, it still needs to be carefully discussed.

Acknowledgements

This research is jointly sponsored by EU project SENSATION and Ministry of Science and Technology of China. The authors would like to make a grateful acknowledgement for VTI, LAAS-CNRS and KI.

References

Åkerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience, 52*, 29–37.

Boverie, S., Leqellec, J. M., & Hirl, A. (1998). Intelligent systems for video monitoring of vehicle cockpit. In *The 1998 international congress and exposition ITS: Advanced controls and vehicle navigation systems* pp. 1–5.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery, 2*(2), 121–167.

Campbell, C. (2000). Algorithmic approaches to training support vector machines: A survey. In *Proceedings of ESANN2000* (pp. 27–36). Belgium: D-Facto Publications.

Chang, C. C., & Lin, C. J. (2001). LIBSVM: A library for support vector machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Cristiannini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge: Cambridge University Press.

Driver fatigue (2006). Driver fatigue: http://www.driverfatigue.co.uk.

Grace, R. (2001). Drowsy driver monitor and warning system. In *International driving symposium on human factors in driver assessment, training and vehicle design, Aspen, CO*.

Hamada, T., Ito, T., Adachi, K., Nakano, T., & Yamamoto, S. (2004). Detecting method for drivers' drowsiness applicable to individual features. *IEICE Transactions on Information and Systems, E87-D*(1), 88–96.

Hayami, T., Matsunaga, K., Shidoji, K., & Matsuki, Y. (2002). Detecting drowsiness while driving by measuring eye movement – A pilot study. In *Proceedings of the fifth international conference on intelligent transportation systems* (pp. 156–161).

Ji, Q., & Yang, X. J. (2002). Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging, 8*, 357–377.

Liu, X., Xu, F. L., & Fujimura, K. (2002). Real-time eye detection and tracking for driver observation under various light conditions. *IEEE Intelligent Vehicle Symposium, 2*, 344–351.

Rechtschaffen, A., & Kales, A. (1968). *A manual of standardized terminology, techniques and scoring for sleep stages of human subjects*. Washington, DC: US Government Printing Office.

Seeing Machines (2004). Facelab transport http://www.seeingmachines.com/transport.htm.

SVM application list (2006): http://www.clopinet.com/isabelle/Projects/SVM/applist.html.

Ueno, H., Kaneda, M., & Tsukino, M. (1994). Development of drowsiness detection system. In *Proceedings of 1994 vehicle navigation and information systems conference, Yokohama, Japan* (pp. 15–20).

Vapnik, V. N. (1998). *Statistical learning theory*. New York: Wiley.

Wierwille, W. W., Lewin, M. G., & Fairbanks, R. J. (1996). Final reports: Research on vehicle-based driver status/performance monitoring, part III (Rep. No. DOT HS 808 640). Virginia: Srpingfield.