Real-Time Vision-Based Driver Drowsiness/Fatigue Detection System

[1]K. P. Yao, [1]W. H. Lin, [1]C. Y. Fang, [2]J. M. Wang, [3]S. L. Chang, and [1]S. W. Chen,
[1]Computer Science and Information Engineering, National Taiwan Normal University, Taiwan
[2]Computer Science and Information Engineering, National Taiwan University, Taiwan
[3]Electronic Engineering, St. John's University, Taiwan

Abstract

In this paper, a vision system for monitoring driver's vigilance is presented. The level of vigilance is determined by integrating a number of facial parameters. In order to estimate these parameters, the facial features of eyes, mouth and head are first located in the input video sequence. The located facial features are then tracked over the subsequent images. Facial parameters are estimated during facial feature tracking. The estimated parametric values are collected and analyzed every fixed time interval to provide a real-time vigilance level of the driver. A series of experiments on real sequences were demonstrated to reveal the feasibility of the proposed system.

Index Terms – In-vehicle vision system, driver's vigilance monitoring, facial parameters, fuzzy reasoning

1. Introduction

Most of the drivers agree to accept the prohibition of driving while intoxicated. Nevertheless, they think it is endurable to drive in a low level of vigilance, such as drowsiness, fatigue and distraction. However, drowsy, weary and distracted driving is equally as dangerous as drunk driving. All significantly degrade driving performance is due to the declines of visual activity, perceptual sensitivity, situational awareness, as well as decision-making capability. Moreover, many drivers have the experience of feeling sleepy, weary, or distracted at the wheel [12].

In this study, we are devoted to the detection of driver fatigue/drowsiness. Currently, available techniques can be divided into three categories: (1) physiological, (2) driving behavior, and (3) facial expression approaches. In the physiological approach, EEG (recording brain activity) [1] and EOG (recording eye movement) [2] have been shown to possess a high level of correlation with human alertness. However, their sensitivity and specificity are still low compared with visual data [2]. Moreover, the instruments for collecting physiological data somewhat interfere with driving operations and may make drivers uncomfortable.

Vigilance analysis based on driving behaviors count on a variety of driving data observed from the vehicle and the road [6]. Different types of driving data are acquired using different equipment. Assorted devices have to be deployed in the vehicle. However, the installment of exotic devices can be subject to the limitations of vehicle type, driver experience, and road condition [5]. Furthermore, inferring the vigilance level of the driver using driving data is typically more sophisticated and uncertain than that using physiological data.

The third class of technique [5][9] is motivated by the human visual system that can easily identify the vigilance/fatigue level of a person based on his/her facial expressions. Facial expressions convey rich information about inward sensations, including both psychological (e.g., cheer, anger, delight, disgust, fear, and surprise) and physiological (e.g., vitality, fatigue, alertness, drowsiness, drunkenness, and illness) reactions [7]. Facial expressions are primarily captured using visual sensors, which are low-cost and easy to install, operate and maintenance. These advantages have made vision-based vigilance monitoring systems increasingly attractive.

In this paper, a single video camera without supplementary light is utilized. We detect facial features primarily and count on their chromatic characteristics. Several strategies for overcoming illumination changes/variations are introduced. We address the configuration and workflow of the proposed system in Section 2. Implementing details of the system are discussed in Section 3. Section 4 demonstrates the feasibility and robustness of the proposed system. Concluding remarks and future work are finally given in Section 5.

2. Drowsiness Detection System

2.1. System Configuration

Three major components constitute the proposed system: a video camera, an in-vehicle computer, and a warning device. Of these, the camera, which provides the input data to the system, plays a critical role in determining the workflow and techniques used in the system. In this study, we mount a video camera on the top of the dashboard right behind the steering wheel (see Figure 1(a)). The camera is tilted at an angle of about 30 degrees to point at the driver's face. Figure 1(b) shows an example image of a driver taken by the video camera.

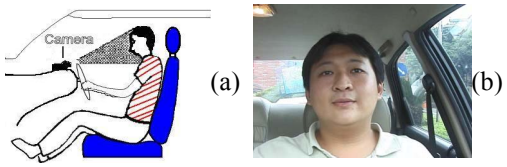


Figure 1. (a) Installation of the video camera. (b) Example image.

2.2. System Workflow

Figure 2 shows a block diagram of the overall operation of the system. There are five blocks in the diagram, preprocessing, facial feature extraction, face tracking, parameter estimation, and reasoning. Each block represents a main stage of the system operation and can be further divided into a number of steps.

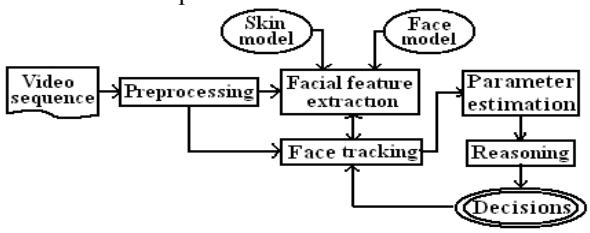


Figure 2. Block diagram for the drowsiness detection system.

The preprocessing stage consists of lighting compensation

978-1-4244-2519-8/10/$26.00 ©2010 IEEE

and color transform. The lighting compensation process, to be detailed in Section 3.1, normalizes the brightness of an image in order to reduce the effects of environmental lighting variations. In the final color transform step, we calculate a number of chrominance values for each image pixel for locating skin regions in the next stage.

In the facial feature extraction stage, the face, eyes, and mouth are detected. There are four steps involved in this stage: skin detection, face localization, eyes and mouth detection, and feature confirmation. We first examine the chrominance values of each image pixel to see whether they satisfy the criteria of a predefined skin model defined in [10]. The largest connected skin region is preserved so as to ignore noisy patches. Moreover, the largest skin area usually corresponds to the face region.

The located face is typically imperfect. Detecting eyes and mouth within the located face is unreliable. So, we search for facial features throughout the entire image (see Section 3.2). The imperfect face will later be used to confirm the detected facial features. Searching for facial features throughout the entire image is time consuming. However, the facial feature detection process is not necessarily applied to every video frame. We will apply Kalman filter [8] to track facial features over video images once they are located in a frame. The facial feature detection process is initiated only when it receives a "no feature" signal from the system.

The facial feature detection module often returns a number of eye and mouth candidates. The feature confirmation step is to select the actual features among the candidates based on both a predefined face model and the face region located earlier. The face model specifies the relationships between facial features in terms of scale-invariant constraints. The face model will be detailed in Section 3.2.

Two groups of the facial parameters will be estimated in the next stage. The first group, including head orientation (tilt, pan and rotation angles) and degree of gaze, are estimated every time the driver's face is located. The second group, including percentage of eye closure over time, eye blinking frequency, eye closure duration, and mouth openness duration, are estimated every n times (n = 300 in our experiments) of face localization. The definition of the facial parameters will be discussed in Section 3.3. A fuzzy integral, addressed in Section 3.4, later combines the parametric values of facial features to obtain a value, which indicates a level of drowsiness.

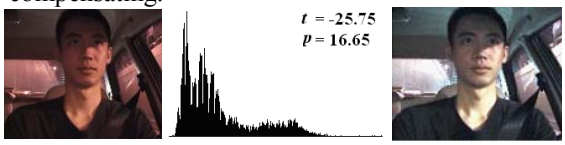
3. Implementations

In this section, the implementation details of lighting compensation, color transform, face model, facial feature extraction, and fuzzy reasoning are addressed.

3.1. Lighting Compensation

Consider a color image $C(R, G, B)$, where R, G, and B are the three color components of the image. First, we estimate the brightness level of the image by computing the intensity of the image as $I = (R+G+B)/3$. Let $h(\cdot)$ be the histogram of I. The distribution of $h(\cdot)$ gives a global description of the brightness of image C. Instead of calculating the skewness of $h(\cdot)$, we compute its distribution tendency. Both measure the asymmetricity of $h(\cdot)$ but with respect to its mean and mid-range, respectively. Let m denote the mid-range of $h(\cdot)$, where $m=L/2$ and L is the number of gray levels. The distribution tendency of $h(\cdot)$ is computed as $t = M_3/(M_2\sqrt{M_2})$, where M_2 and M_3 are the second and third moments of $h(\cdot)$ about m. The value of t can be positive or negative. A positive t indicates a relatively bright image, while a negative t indicates a relatively dark image.

Having computed the histogram distribution tendency of an image, if $t \le -10$, we calculate a fraction $p = \alpha e^{-\beta(t+138)}$, where α = 0.713 and β = 0.013. Then, for each color component of the image we find the pixels with the top p fraction of values and average them to obtain the averages (a_R, a_G, a_B). For each image pixel, we scale its (r, g, b) color values as $r' = 255\frac{r}{a_R}$, $g' = 255\frac{g}{a_G}$, and $b' = 255\frac{b}{a_B}$, which will raise the brightness of the image. In a similar way, if $t \ge 50$, we calculate p, where α =0.323 and β = 0.011. But, now we find the pixels in the bottom p fraction and average them to (a_R, a_G, a_B). The image pixels are scaled by $r' = 255(1-\frac{255-r}{255-a_R})$, $g' = 255(1-\frac{255-g}{255-a_G})$, and $b' = 255(1-\frac{255-b}{255-a_B})$. This results in a decrease of the image brightness. Finally, no operation is applied to the image if $-10 < t < 50$. Figure 3 shows an example after compensating.



(a) (b) (c)

Figure 3. Example illustrating lighting compensation; (a) input image, (b) histogram, distribution tendency, and fraction, and (c) resultant image.

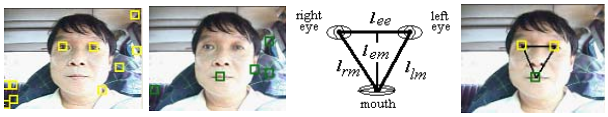
3.2. Facial Feature Detection

In facial feature extraction, eyes and mouth are detected in an image using the Hsu *et al*. technique [4]. First, the image is transformed from the RGB to the YC_rC_b color spaces [3]. Some maps calculated from the new transformed image emphasize the chrominance and luminance characteristics of eyes and mouths. Finally, a number of eye candidates are detected. The examples in Figure 4(a) and (b) show the detected eye and mouth candidates.

In Figure 4(c) of a face model, there are five constraints (two qualitative and three quantitative constraints) describing the face model: (1) the mouth is lower than the two eyes; (2) the horizontal position of the mouth is between those of the two eyes; (3) the tilt of the line connecting the two eyes from the horizontal is smaller than 10 degrees; (4) the angle between the line joining the two eyes and the line connecting the mouth and any eye is between 45 and 75 degrees; (5) $l_{ee} < l_{rm}$, $l_{ee} < l_{lm}$ and $l_{lm} \approx l_{rm}$. The lengths of l_{ee}, l_{rm}, l_{lm}, and l_{em} are determined from a few faces detected at the early stage of system processing.

Let S_e and S_m be the sets of eye and mouth candidates, respectively. For each pair of eye candidates in S_e and a mouth candidate in S_m, we examine their relationships to see whether

they satisfy the constraints of the face model. If so, the features form a face candidate represented by the triangle formed by sequentially connecting the three features. We collect all face candidates in set S_f. For each face candidate f_i, we also compute a level of confidence c_i. Let α indicate the degree of shape similarity between the triangle of f_i and an equilateral triangle, defined as $\alpha = \frac{1}{\pi}\max_{1\le i\le 3}\{|a_i - \frac{\pi}{3}|\}$ where a_i's are the internal angles of the triangle of f_i. Let β denote the degree of vicinity between f_i and the face region previously detected, defined by $\beta = \frac{1}{l_r} dis(\mathbf{c}_f, \mathbf{c}_s)$ where c_f and c_s are the centers of gravity of f_i and the face region, l_r specifies the row length of images and $dis(\cdot)$ is a distance function. The level of confidence of f_i is then calculated according to $c_i = e^{-(\alpha+\beta)/2}$. Finally, we determine the potential face by $f^* = \arg\max_{f_i \in S_f}\{c_i\}$. Note that only if the potential face is also observed in the following image, it is regarded as the actual face. Figure 4(d) shows a located face.



(a) (b) (c) (d)

Figure 4. Face localization: (a) eye candidates, (b) mouth candidates, (c) face model, (d) actual facial features.

3.3 Facial Parameters

Facial parameters are estimated around facial features. The neighborhoods of eyes and mouths, within which parameters are evaluated, are delimited by rectangular windows of sizes l_{ee} / 1.5 by l_{em} / 3.5 and l_{ee} by l_{em} / 2.5, respectively.

A. Eye Parameters

It is observed that the vertical edge magnitudes of an open eye are stronger than those of a closed eye. Let E_{le} and E_{re} be the averages of vertical edge magnitudes of the left and right eyes of a face, respectively. We define the degree of eye closure, d_e, for the face as $d_e = \min\{E_{le}, E_{re}\}/255$. Here, a small d_e indicates a high degree of eye closure.

In the example shown in Figure 5, the distribution of d_e values calculated from a sequence of 300 face images is depicted in Figure 5(a). There are a number of significant valleys along the distribution curve, referred to as the d_e-curve for simplicity, which correspond to the face images containing eyes with high degrees of closure. In order to highlight significant valleys, we transform the d_e-curve into the v_e-curve by $v_e = (d_e - m)^2$, where $m = \frac{1}{n}\sum_{i=1}^{n} d_{ei}$ and n (=300) is the number of images in the sequence. The above transformation actually performs a contrast enhancement on the d_e-curve. Figure 5 shows the example of transformation result.

Based on the b_e-plot, we can easily calculate the parameters of (a) percentage of eye closure over time, $PERCLOS = \frac{1}{n}\left(\sum_{i=1}^{n} b_{ei}\right)$ where b_{ei} is the binary value of the b_e-plot at image i, (b) blinking frequency, $BF = n_p / T_n$, where n_p is the number of pulses on the b-plot and T_n is the period of the image sequence, and (c) eye closure duration, $D = \max_{1\le i\le n_p}\{D_i\}$ where D_i is the time interval of pulse i.

B. Mouth Parameters

The open mouth has stronger edge magnitudes than the closed mouth. Let E_m be the average edge magnitude of a mouth. We define the degree of mouth openness as $d_m = E_m / 255$. A large value of d_m indicates a high degree of mouth openness. As the transformation in eye parameters, we transform d_m v_m-curve and b_m-plot. Based on the b_m-plot, the mouth openness duration is calculated by $D_m = \max_{1\le i\le n_p}\{D_i\}$, where n_p is the number of pulses on the b_m-plot and D_i is the time interval of pulse i.

C. Head Parameters

Head parameters include the tilt (α), pan (β) and rotation (γ) of the head. Referring to the face model in Figure 4(c), the line segment l_{em} will be foreshortened if the driver tilts his head. Let l'_{em} be the foreshortened line segment. The tilt angle is given by $\alpha = \cos^{-1}(l'_{em} / l_{em})$. Likewise, let l'_{ee} be the foreshortened line segment of l_{ee} when the driver's head pans. The pan angle is given by $\beta = \cos^{-1}(l'_{ee} / l_{ee})$. Finally, the rotation angle of the head is computed by $\gamma = \cos^{-1}\left(|x_l - x_r| / l_{ee}\right)$, where x_l and x_r are the horizontal coordinates of the left and right eyes, respectively.

D. Gaze Parameter

A person gazing at something will have his/her eyes open and hold his/her head stationary. Accordingly, we define the gaze parameter. Since the degree of eye openness is inversely related to the degree of eye closure (d_e) previously defined, the smaller d_e the larger the degree of eye openness and in turn the degree of gaze. A stationary head has both its orientation and position fixed. Head orientation (α, β, and γ) has been defined earlier. We define its time variation as $\delta = \max\{\Delta\alpha, \Delta\beta, \Delta\gamma\}$. Therefore, the smaller δ is the higher the degree of gaze. Likewise, the smaller is head movement m, which is readily evaluated in the course of face tracking, the higher the degree of gaze. Let $d_g = [m/l_r + \delta/90 + d_e]/3$, where l_r is the row length of images. Then, a small d_e indicates a high degree of gaze.

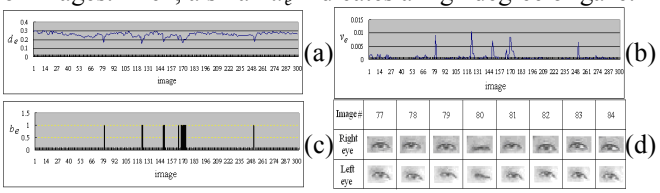


Figure 5. The (a) d_e -curve, (b) v_e-curve, and (c) b_e-plot of a sequence of 300 face images, (d) vicinity of image 80.

3.4. Fuzzy Reasoning

Table 1 lists the transfer functions for different parameters. The fuzzy integral provides an elegant nonlinear numerical approach to integrating multiple sources of information to arrive at a value that indicates the degree of support for a particular hypothesis or decision. We use this technique to integrate the values of facial parameters to deduce the drowsiness level of a face.

A. Fuzzy Integral

There are several variants of the fuzzy integral [8] based on either the Lebesque or Riemann integral. In this study, the Sugeno fuzzy integral based on the Lebesque integral is

considered. Let $f: S \to [0,1]$ be a function defined on a finite set S and $g: P(S) \to [0,1]$ be a set function defined over the power set of S. Function $g(\cdot)$, referred to as the Sugeno fuzzy measure function, satisfies

$$g(A \cap B) = g(A) + g(B) + \lambda g(A) g(B), \quad \lambda \geq 1. \tag{1}$$

The fuzzy integral of $f(\cdot)$ with respect to $g(\cdot)$ is defined as

$$e = \int_S f(s) \cdot g = \sup_{\alpha \in [0,1]} \{\alpha \wedge g(A_\alpha)\}, \tag{2}$$

where $\wedge$ represents fuzzy intersection and $A_\alpha = \{s \in S \mid f(s) \geq \alpha\}$.

Suppose we have a collection of information sources, $S = \{s_i, i = 1, \cdots, n\}$, and a set of hypotheses, $H = \{h_i, i = 1, \cdots, m\}$. A decision D is to be made from H based on S using the fuzzy integral approach. To this end, for each hypothesis $h \in H$ we calculate its fuzzy integral value e_h according to Equation (2). The decision is then determined as $D = h^* = \arg\max_{h \in H} e_h$. To calculate e_h, $f(\cdot)$ and $g(\cdot)$ are defined. Function $f(\cdot)$ receives an information source $s \in S$ and returns the value that indicates the level of support of s to hypothesis h. Function $g(\cdot)$ taking as input a subset A of S gives the value that specifies the degree of worth of subset A relative to the other sources in S.

To calculate $g(A)$, we have to first determine the degrees of worth of individual information sources, $W = \{w(s_i), i = 1, \cdots, n\}$, where $w(s_i) = g(\{s_i\})$. $g(A)$ can then be recursively calculated using Equation (1),

$$g(A) = \left(\prod_{s_i \in A} (1 + \lambda w(s_i)) - 1\right) / \lambda. \tag{3}$$

Since $g(S) = 1$ and Equation (3). Parameter λ can be determined by solving $\lambda + 1 = \prod_{s_i \in S} (1 + \lambda w(s_i))$.

Let $S' = \{s'_1, s'_2, \cdots, s'_n\}$ be the sorted version of S such that $f(s'_1) \geq f(s'_2) \geq \cdots \geq f(s'_n)$. Equation (2) can be rewritten as

$$e = \int_S f(s) \cdot g = \sup_{\alpha \in [0,1]} \{\alpha \wedge g(A_\alpha)\} = \bigvee_{1 \leq i \leq n} [f(s'_i) \wedge g(S'_i)], \tag{4}$$

where $\vee$ represents fuzzy union.

B. Reasoning

Let the eight facial parameters P = {*PERCLOS*, *BF*, *D*, α, β, γ, D_m, d_G} involved in our drowsiness analysis have degrees of worth, $W = \{w_1, w_2, \cdots, w_8\}$, which are determined a priori on the basis of three criteria: importance, accuracy and reliability. In this study, $W = \{0.93, 0.8, 0.85, 0.5, 0.3, 0.3, 0.5, 0.9\}$ is used in the next section. Based on W, the degree of worth $g(A)$ of any subset $A \subseteq P$ is readily computed by Equation (3).

Let $V = \{v_1, v_2, \cdots, v_8\}$ be the parametric values estimated from a face. This set of values is transferred into $S = \{s_1, s_2, \cdots, s_8\}$ using the transfer functions listed in Table 1. Set S here serves as the set of information sources. Next, we determine the hypothesis set H. Let $m = \left\lfloor 10 \times \min_{s_i \in S} s_i \right\rfloor / 10$ and $M = \left\lceil 10 \times \max_{s_i \in S} s_i \right\rceil / 10$, in which $\leq \cdot \rfloor$ and $\lceil \cdot \geq$ denote the floor and ceiling operators, respectively. The hypothesis set is then defined as $H = \{m, m+0.1, m+0.2, \cdots, M\}$, from which a hypothesis is to be selected as the drowsiness level of the face.

For each hypothesis $h \in H$, we calculate the levels of support of information sources $s_i \in S$ to the hypothesis h by $f_h(s_i) = 1 - |s_i - h|$. Clearly, the larger the difference between s_i and h the lower the level of support of s_i to h. We then sort the information sources in S according to their calculated levels of support. Let $S' = \{s'_1, s'_2, \cdots, s'_8\}$ be the sorted version of S such that $f(s'_1) \geq f(s'_2) \geq \cdots \geq f(s'_8)$. Substituting $f(s'_i)$ and $g(S'_i)$ into Equation (4), we obtain the fuzzy integral value e_h of h. The above process repeats for every hypothesis in H. Finally, the drowsiness level of the face is determined as $h^* = \arg\max_{h \in H} e_h$.

Table 1.Ttransfer functions for facial parameters

Percentage of eye closure over time ([0,1])	Eye blink frequency (times/minute)
$D(x) = \begin{cases} 0 & x \leq 0.12 \\ 3.57x - 0.43 & 0.12 < x < 0.4 \\ 1 & x \geq 0.4 \end{cases}$	$D(x) = \begin{cases} 0 & x \leq 20 \\ 0.05x - 1 & 20 < x < 40 \\ 1 & x \geq 40 \end{cases}$
Eye closure duration (seconds)	**Head pan and tilt angles ([0, 90])**
$D(x) = \begin{cases} 0 & x \leq 0.15 \\ 0.35x - 0.05 & 0.15 < x < 3 \\ 1 & x \geq 3 \end{cases}$	$D(x) = \begin{cases} 0 & x \leq 20 \\ 0.04x - 0.8 & 20 < x < 45 \\ 1 & x \geq 45 \end{cases}$
Head rotation angle ([0, 90])	**Mouth openness duration (seconds)**
$D(x) = \begin{cases} 0 & x \leq 10 \\ 0.05x - 0.5 & 10 < x < 30 \\ 1 & x \geq 30 \end{cases}$	$D(x) = \begin{cases} 0 & x \leq 3 \\ 0.33x - 1 & 3 < x < 6 \\ 1 & x \geq 6 \end{cases}$
Gaze degree ([0, 1])	
$D(x) = \begin{cases} 0 & x \leq 0.5 \\ 0.05x - 0.5 & 0.5 < x < 1.0 \\ 1 & x \geq 1.0 \end{cases}$	

4. Experimental Results

The proposed driver's drowsiness detection system has been developed using the Borland C++ Programming Language run on an Intel Solo T1300 1.66 GHz PC running under Windows *XP* Professional. The input video sequence to the system is at a rate of 30 frames/second. The size of video images is 320 x 240 pixels.

A number of experimental video sequences of various drivers were taken under different illumination conditions, including shiny light, sunny day, cloudy day, twilight, underground passage, and tunnel. Table 2 summarizes the performances of the system on these video sequences. The system performance of our concern includes facial feature detection rate r_d, face tracking rate r_t, and system speed s, indicated in the top row of Table 2. Let n be the number of frames. The facial feature detection rate r_d is defined as $r_d = n_1/n_d$, where n_d is the number of frames to which the facial feature extraction module is applied and n_1 is the number of frames in which facial features are correctly detected. Likewise, the face tracking rate r_t is defined as $r_t = n_2/n_t$, where $n_t = n - n_d$ is the number of frames to which the face tracking module is applied and n_2 is the number of frames in which the face constituting facial features is correctly located by the tracking module. Let T be the total processing time of the video sequence. The system speed s is defined as $s = t / n$.

As mentioned, the processing speed of our system is primarily governed by the facial feature extraction and face tracking modules. The former takes about 1/8 seconds to detect facial features in an image, while the latter takes about 1/25 seconds to locate facial features in an image. The system had a rather poor performance on video 8. The video was acquired while driving in a tunnel and possessed a very low intensity, which degrades the chromatic characteristics of the

video. Although the lighting compensation process incorporated in our system can recover the brightness of the video, its color is significantly distorted because vigorous brightness compensation is performed.

Table 2. Performances of the system on experimental video sequences.

Video	Video length n (#frames)	Facial feature detection rate r_d (%)	Face tracking rate r_t (%)	System speed s (f/sec)
1	19650	96.4	95.3	18.8
2	5790	94.8	89.9	17.7
3	11400	91.3	96.2	20.2
4	2070	88.7	92.6	17.3
5	3960	89.4	96.5	20.4
6	5850	90.6	97.4	21.7
7	6780	81.5	86.1	15.3
8	4470	47.9	58.3	6.2

Drowsiness level is evaluated for every video frame based on a number of facial parameter measurements. Since different parameters have different meanings and ranges of measured values, the measured values are transferred into the numbers within the interval [0, 1], which express degrees of drowsiness. The fuzzy integral process then combines the numbers to arrive at a value that indicates a level of drowsiness. Figure 6 displays the distributions of drowsiness levels evaluated from videos 1–7. Individual prominent peaks along the distribution curves will be ignored because a drowsy face would exhibit a number of continuous intervals of high drowsiness levels.

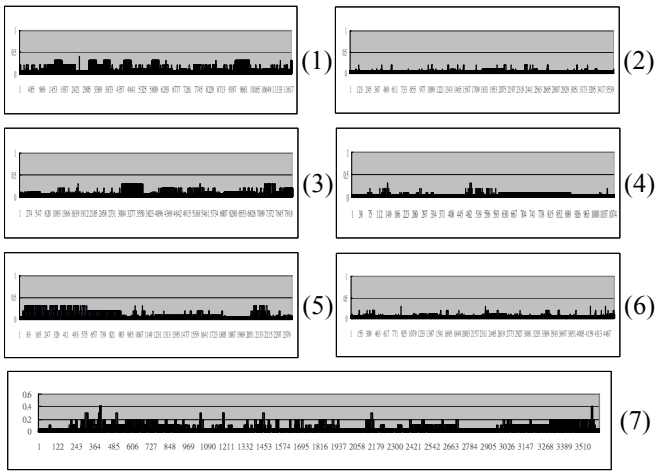


Figure 6. Distributions of drowsiness level evaluated from videos 1-7.

5. Concluding Remarks and Future Work

In this paper, an in-vehicle nonintrusive vision-based driver drowsiness detection system was presented. The system estimates the drowsiness level of a driver based on his/her facial expressions, for which we measure a number of facial feature values from the driver. The values are then integrated to deduce the drowsiness level of the driver. A series of experiments were conducted. Both the efficiency and effectiveness of the proposed system and its constituent stages were examined. Preliminary experimental results reveal the feasibility and robustness of our system under various environmental and illumination conditions.

Two major difficulties were encountered during development. The first difficulty results from the wide range of lighting sources and abrupt changes in lighting conditions. To compensate for this difficulty, a lighting compensation process was introduced. The second difficulty originates from vehicle vibration and leads to video instability. To reduce the influence, a strategy of repeat confirmations of detected facial features in successive images was employed. We will consider video stability techniques in our future work.

References

[1] J. L. Cantero, M. Atienza, and R. M. Salas, "Human Alpha Oscillations in Wakefulness, Drowsiness Period, and REM Sleep: Different Electroencephalographic Phenomena within the Alpha Band," *Neurophysiol Clin*, 32, pp. 54–71, 2002.

[2] I. G. Damousis and D. Tzovaras, Fuzzy Fusion of Eyelid Activity Indicators for Hypovigilance-Related Accident Prediction," *IEEE Trans. on Intelligent Transportation Systems*, vol. 9, no. 3, pp. 491-500, 2008.

[3] J. J. de Dios and N. Garcia, "Face Detection Based on A New Color Space YC_gC_r," *Proc. of Int'l Conf. on Image Processing*, vol. 3, pp. 909 -912, 2003.

[12] P. H. Gander, N. S. Marshall, R. B. Harris, and P. Reid, "Sleep, Sleepiness and Motor Vehicle Accidents: A National Survey," *Journal of Public Health*, 29(1), pp. 16-21, 2005.

[4] R.L. Hsu, A. M. Mottaleb and A. K. Jain, "Face Detection in Color Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696-706, 2002.

[5] Q. Ji, Z. Zhu, and P. Lan, "Real-time Nonintrusive Monitoring and Prediction of Driver Fatigue," *IEEE Trans. Vehicular Technology*, vol. 53, no. 4, pp. 1052-1068, 2004.

[6] Y. Liang, M. L. Reyes, and J. D. Lee, "Real-Time Detection of Driver Cognitive Distraction Using Support Vector Machines," *IEEE Trans. on Intelligent Transportation Systems*, vol. 8, no. 2, pp. 340-350, 2007.

[7] Y. Mitsukura, H. Takimoto, M. Fukumi and N. Akamatsu, "Face Detection and Emotional Extraction System Using Double Structure Neural Network," *Proc. of the Int'l Joint Conf. on Neural Networks*, vol. 2, pp. 1253 -1257, 2003.

[8] Z. Wang and G. J. Klir, *Fuzzy Measure Theory*, Plenum Press, New York, 1992.

[9] Y. Wu, H. Liu and H. Zha "A New Method of Detecting Human Eyelids Based on Deformable Templates," *IEEE Int'l Conf. on Systems, Man and Cybernetics*, vol. 1, pp. 604-609, 2004.

[10] J. M. Wang, H. W. Lin, C. Y. Fang, and S. W. Chen, "Detecting Driver's Eyes During Driving," *Proc. of 18th IPPR Conf. on Computer Vision, Graphics and Image Processing*, Taiwan, A4-5, pp. 941-9477, Aug 2005.