IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 7, 2020, accepted March 26, 2020, date of publication April 7, 2020, date of current version April 22, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2986234

Vision-Based Instant Measurement System for Driver Fatigue Monitoring

YIN-CHENG TSAI[1], (Student Member, IEEE), PENG-WEN LAI[1], (Student Member, IEEE), PO-WEI HUANG[1], (Student Member, IEEE), TZU-MIN LIN[1], AND BING-FEI WU[2], (Fellow, IEEE)

[1]Institute of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan
[2]Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan

Corresponding author: Po-Wei Huang (powei@cssp.cn.nctu.edu.tw)

This work was supported in part by the Ministry of Science and Technology, Taiwan, under Grant MOST 108-2221-E-009-123-MY2.

ABSTRACT In this paper, a vision-based physiological signal measurement system is proposed to instantly measure driver fatigue. A remote photoplethysmography (rPPG) signal is a type of physiological signal measured by a camera without any contact device, and it also retains the characteristics of the PPG, which is useful to evaluate fatigue. To solve the inconvenience caused by the traditional contact-based physiological fatigue detection system and to improve the accuracy, the system measures both the motional and physiological information by using one image sensor. In a practical application, the environmental noise would affect the measured signal, and therefore, we performed a preprocessing step on the signal to extract a clear signal. The experiment was designed in collaboration with Taipei Medical University, and a questionnaire-based method was used to define fatigue. The questionnaire that could directly react to the feeling of the subject was treated as our ground truth. The evaluated correlation was 0.89 and the root mean square error was 0.65 for ten-fold cross-validation on the dataset. The trend of driver fatigue could be evaluated without a contact device by the proposed system. This advantage ensures the safety of the driver and reliability of the system.

INDEX TERMS Fatigue monitoring, remote photoplethysmography, biomedical monitoring, image sequence analysis.

I. INTRODUCTION

Road traffic accidents have been predicted to be the third leading cause of death and disability in 2020 [1]. For most of the population of the world, the burden of road traffic injury in terms of societal and economic costs is rising substantially [2]. Statistics show that driver fatigue is a contributing factor in numerous accidents and yearly approximately 20% of the total accidents are related to sleepiness [3]. To reduce these economic costs, finding a method to detect driver fatigue has become an important issue.

A few years ago, vision-based methods were proposed to detect fatigue by capturing specific features with one or more cameras and image processing. In these methods, a driver monitor system captures the face of the driver and extracts features such as blinking, yawning, and the head movement of the driver [4]–[8]. Further, a driver assistance system captures the road image, labels the lane line to determine the lane departure, and decides the fatigue state [7], [8]. However, these methods are sensitive to the shield on the images. In addition, these methods raise an alarm after the fatigue has occurred. Physiological-based methods utilizing physiological signals for determining the fatigue state have been proposed to solve the above problem. In addition, a physiological signal directly transmits the information of the subject and is rarely affected by different scenarios. Electrocardiograms (ECGs), electroencephalograms (EEGs), electromyograms (EMGs), respiration, galvanic skin response (GSR), heart rate (HR), heart rate variability (HRV), and pulse rate variability (PRV) are common indicators for determining the fatigue state [9]–[12]. However, the above signals require a subject to be equipped with complex devices while driving, which might make the subject feel inconvenient. More seriously, this requirement might affect the driving and lead to accidents.

The associate editor coordinating the review of this manuscript and approving it for publication was Alberto Botter.

Photoplethysmography (PPG) is also one of the widely used physiological signals that can convert the data to HR and PRV. PPG can be sensed by wearable devices and improves the convenience. Al-Libawy *et al.* used PPG signals to calculate PRV and classified the fatigue state by a support vector machine (SVM) or an artificial neural network (ANN); the total accuracy reached 88.3% and 91.3% for six subjects [13]. Huang also used PPG and PRV to measure the fatigue state and set a threshold for its determination by a functional link-based fuzzy neural network (FLFNN) [14]. Choi combined various information such as HR, HRV, GSR, temperature, acceleration, and rate of rotation to determine the status of subjects [9]. Despite the extreme convenience of a PPG sensor, it still requires the subjects to wear extra devices. In addition, it has some limitations such as forbidding the subject to press or shake, which might influence the driver.

As mentioned above, conventional vision-based methods have the weakness that they cannot measure fatigue immediately, and physiological-based methods use inconvenient sensors. Consequently, measuring a physiological signal by vision-based methods is an effective solution for the above problems. A remote photoplethysmography (rPPG) signal is a type of physiological signal measured by a camera without any contact device while retaining the characteristics of PPG. Qi and Wang applied rPPG to calculate PRV and analyzed the relationship between fatigue and PRV [15]. Tayibnapis combined facial features, rPPG signals, and the SVM method to classify the fatigue state [16].

The above methods did not provide a unified and precise definition of fatigue estimation. To define an explicit criterion for fatigue estimation, we collaborated with Taipei Medical University and designed a questionnaire-based experiment to measure the fatigue state of a subject. In our study, Swedish Occupational Fatigue Inventory (SOFI) [17] and fatigue assessment scale (FAS) [18] which can reflect different aspects of fatigue were treated as baselines to evaluate the fatigue state. Moreover, in medical research, a the visual analogue scale (VAS) [19] is a type of commonly used questionnaire, and therefore, VAS was also employed to observe various fatigue states during the experiment. To measure fatigue, we requested the subjects to fill out questionnaires such as SOFI, FAS, and VAS. After recording the questionnaires, we converted the questionnaire scores into indicators for assessing fatigue.

Considering all the complexities, we proposed a system to obtain the trend of the subject fatigue state with only one camera as the sensor. In the practical application, the proposed system captured the face of the subject to measure the physiological information and facial features. To acquire a better signal for precisely estimating the fatigue trend, the system consisted of a preprocessing stage to detect the presence of the face and obtain the image with the frontal face of the subject. Further, then the system converted collected information such as rPPG and facial expression into the HR, PRV, percentage of eyelid closure (PERCLOS) and yawning state. Finally, it calculated the fatigue score for evaluating the trend of the fatigue state by an artificial neural network (ANN).

TABLE 1. Experiment devices.

Item	Description
Webcam	Logitech C920
Resolution	VGA
Frame Rate	30 fps
Image Format	bitmap
PPG sensor	CMS50EW

II. EXPERIMENTS AND DATA

A. EXPERIMENT DEVICES AND QUESTIONNAIRE

As summarized in Table 1, a Logitech C920 camera was chosen as the image sensor of the system. The camera resolution and frame rate were set as VGA and 30 fps, respectively. Contact measurement signals, such as ECG and PPG, could be the criteria of the contactless measurement signal. However, ECG devices often complicate the equipment and measurement, which can easily affect the subjects in executing the experiment. Thus, in the experiment, CMS50EW, which is a Food and Drug Administration (FDA) certified device, was used to obtain the PPG signal as the ground truth.

Fatigue is a type of physiological phenomenon, and the exact standards for fatigue have not been established yet. Because we aimed to establish a clear rule to quantify fatigue, the questionnaire method [20] was used to estimate the fatigue state. Here, SOFI, FAS, and VAS were chosen as our questionnaires. SOFI categorizes fatigue into five classifications: lack of energy, physical exertion, physical discomfort, lack of motivation and sleepiness. Following this, it designs 25 cross-reference questions related to at least one classification to investigate the degree of performance of the subject in various fatigues. According to Leung *et al.* translation of SOFI into Chinese did not have a strong cultural effect, and so we used the Chinese version of SOFI [21]. FAS easily classifies fatigue into physical and psychological categories, and then designs five questions to survey. VAS is a simple and quick survey for grading sleepiness levels from 1 to 10, and it is widely used in medical research owing to its convenience and effectiveness. Sleepiness fatigue is a mixed result of numerous factors, and although some classifications of SOFI and FAS do not directly focus on it, the impact of these classifications on the subjects still cannot be ignored. Therefore, SOFI, FAS, and VAS were mixed to evaluate the fatigue state.

B. PARTICIPANTS

The subjects included 20–50 years old healthy people who were required to have a driver's license and cooperate with the limitations of the clinical trials. All of the subjects had given their informed consents before participating in the

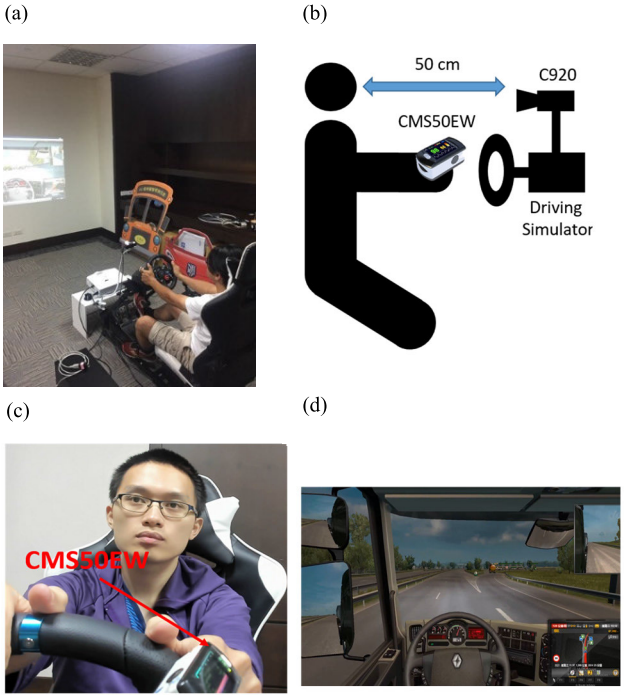

FIGURE 1. Driving simulation experiment scenario. (a) Experimental site. (b) Schematic of the experiment facility. (c) Actual shot. (d) Experiment scenario.

experiment. The subjects did not suffer from major diseases such as sleep apnea or medical history, use of psychotropic medications, tobacco, and alcohol, or drug addiction. In this experiment, there were 32 subjects, including 12 males and 20 females, and their age ranged from 22 to 46 years. The average age was 30.9±8.4 years. To prevent the data from interference by external factors, the subjects were demanded to follow the below three rules:

1. Consumption of drinks and alcoholic food was prohibited 24 h before the experiment.
2. Caffeinated beverages such as tea or coffee were forbidden for 4 h before the experiment.
3. They were required to complete their meal 1 h before the experiment.

C. EXPERIMENT PROCESS

Considering the risk of the driver fatigue, a driving simulation experiment is the appropriate method to simulate a real case and record the experiment data. In this study, the experimental site was set indoors with sufficient light to avoid the interference from the external environment and ambient light, as shown in Fig. 1 (a). During the experiment, the subjects were required to wear CMS50E to record the current HR and PPG signals. To capture the image for the rPPG calculation, webcam C920 was set in front of the subject, as shown in Fig. 1 (b) and (c). The used experiment software was Euro Truck Simulator 2, and the experiment scenario was set in a monotonous highway to reduce scenarios that may irritate a driver like a traffic light or pedestrians, as displayed in Fig. 1 (d).

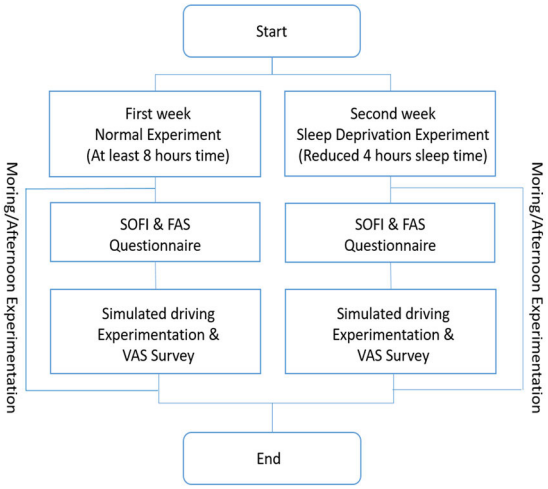

FIGURE 2. Driving simulation experiment process.

TABLE 2. Experiment and questionnaire for obtaining the data.

Status	Detail	Time	Used questionnaires
Normal	Questionnaire survey	5 min	SOFI, FAS
	Driving experiment	60 min	VAS
	Break	60 min	
	Questionnaire survey	5 min	SOFI, FAS
	Driving experiment	60 min	VAS
Sleep Deprivation	Questionnaire survey	5 min	SOFI, FAS
	Driving experiment	60 min	VAS
	Break	60 min	
	Questionnaire survey	5 min	SOFI, FAS
	Driving experiment	60 min	VAS

The driving simulation experiment process is shown in Fig. 2. We followed the standard procedures approved by Taipei Medical University to design the experiment treatments, and the experiment had received the Institutional Review Board (IRB) certification. The total experiment duration for each subject was two weeks. The subjects were demanded to sleep for more than eight hours before the experimentation to ensure adequate sleep in the first week of the normal state experimentation. In the second week of the sleep deprivation experimentation, the subjects were demanded to reduce their sleep time by four hours compared to the previous week so that the subjects remained tired during the experiment. The experiment details are listed in Table 2. On the day of the experimentation, each subject would fill out the SOFI and FAS questionnaires in 5 min before the driving experiment started and participate in the

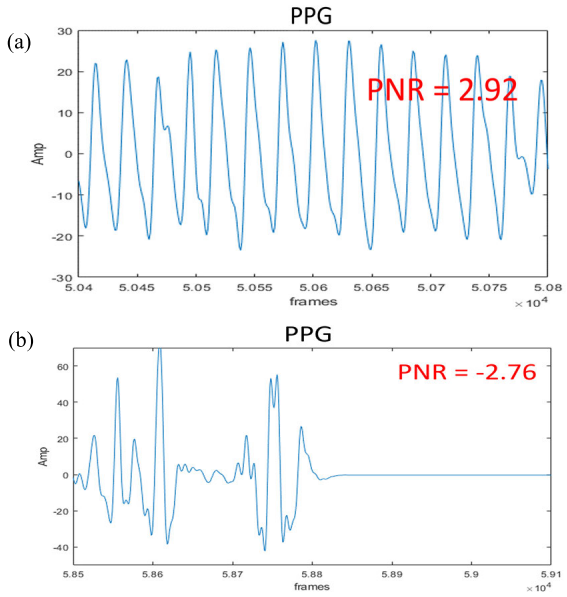

FIGURE 3. Waveform of CMS50E. (a) In the normal case. (b) In the abnormal case, the sensor was pressed by the user, and the waveform was destroyed.

driving experimentation for 60 min in the morning and afternoon. During the driving experiment execution, the subjects were requested to execute the VAS questionnaire survey every 10 min to confirm their fatigue state. After performing the experimentation in the morning, each subject would take a 60-minute break. To prevent the fatigue state from being affected by other factors, the subjects were reminded not to doze off during the break time.

D. DATASET DESCRIPTION

The purpose of the dataset was to record the images of the subjects for calculating the rPPG signal. The fatigue questionnaires were used for recording the fatigue state, and information such as the HR and PPG from the PPG sensor were employed for comparing the accuracy of the rPPG signal.

To simulate the actual driving scenario, we did not deliberately restrict the subject in this dataset. Because the subject needed to manipulate the steering wheel in the experimentation, finger shaking was inevitable, which might affect the accurate measurement of the PPG by the sensor. A wrong value of the PPG sensor would impede the accurate evaluation of the rPPG. Accordingly, the bad PPG signal should be filtered out before detecting fatigue. In addition, the necessary element of rPPG was the image with a face looking forward, and so the images with the face looking at the side mirror were also filtered out.

The pattern in Fig. 3 (a) displays remarkable peaks originating from heartbeats, and in Fig. 3 (b), the signal loses the feature of heartbeat and is deliberately polluted by noise. The energy of the maximum peak and noise ratio (PNR) in 512 frames at 30 fps is the PPG filter indicator. The threshold is chosen as zero which represents the energy of PPG signal is larger than the noise and PNR was selected as the property to monitor the PPG signal quality.

III. METHOD

The contactless fatigue detection system architecture is presented in Fig. 4. First, the system captures the image of the face of the driver by a camera or webcam, and then using the face detection algorithm and landmarks to mark the face position with the relevant features on the image. These coordinates decide the region of interest (ROI) to calculate the physiological signal. Second, the system converts the image data in the ROI into rPPG signals. Third, after the system collects more than 5 min of data, it captures the time and frequency domain characteristics of the rPPG signal and converts them into HR and PRV, serving as indicators of fatigue. In addition, facial expressions are also important indicators for determining fatigue, and so, behaviors such as blinking and yawning are also detected and used for fatigue decision after labeling the landmarks. Finally, the system uses the features, including HR, HRV, and behavior, as the input parameters to train the ANN model by a python code, and then the ANN model is applied and used for predicting the fatigue state.

A. ROI EXTRACTION

A human face contains significant information, which can be used for fatigue detection. For example, eyes and mouth can reflect behavior, and physiological information is also hidden in the skin. Hence, the system selected specific regions to measure, in the first step.

Multi-task Cascaded Convolutional Networks (MTCNN) [22] was chosen for finding the face position, and then Dlib [23] provided the facial landmarks. Facial landmarks can indicate the eyes and mouth position on the image directly. However, to obtain a clear physiological signal, it is necessary to select a small area with a strong signal to reduce the efficiency of the operation and avoid the influence of the noise. According to Kwon *et al.*, the signal quality from the cheeks and philtrum is extremely good, and so, these areas were selected as ROI [24]. The ROI extraction was performed using the result of the MTCNN and Dlib landmarks, as Fig. 5 (a) shows. The calculation process is expressed in (1) where *D(x,y)* is the Euclidean distance of x and y. The definitions of the notations are listed in Table 3, and the result is displayed in Fig. 5 (b).

$$\begin{cases} L = D(\frac{1}{2}E_{\mathbf{L}} + \frac{1}{2}E_R, \frac{1}{2}M_L + \frac{1}{2}M_R) \\ ROI_x = C_x - 0.65Face_{width} * L \\ ROI_y = C_y + 0.2Face_{height} * L \\ ROI_{width} = 1.3L \\ ROI_{height} = 0.6L \end{cases} \tag{1}$$

B. rPPG EXTRACTION

rPPG is a signal arising from the changes in the blood volume, and it reveals physiological information such as the HR

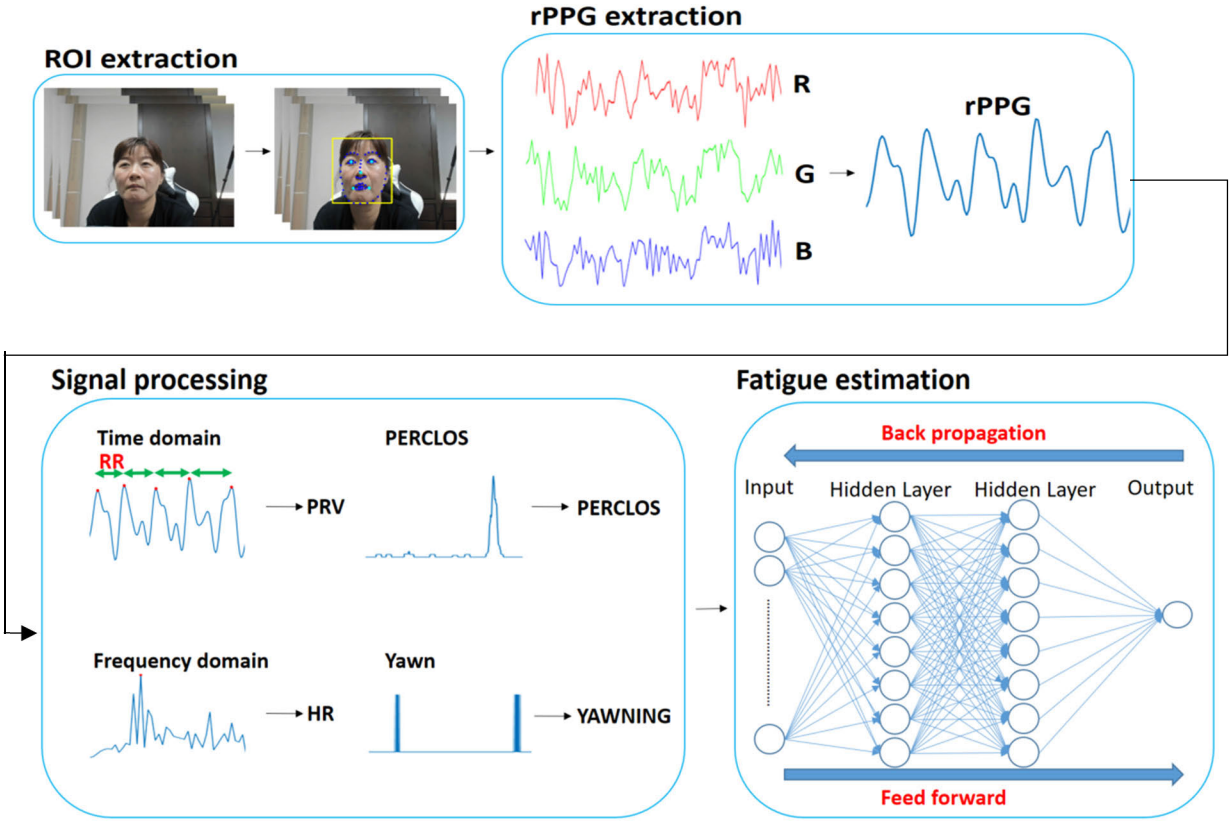

FIGURE 4. Overview of the proposed fatigue detection framework.

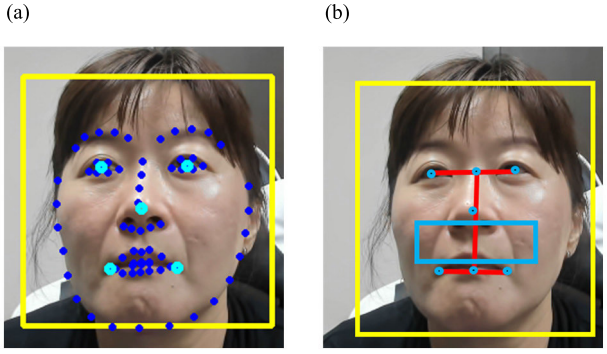

FIGURE 5. Result of the ROI extraction. (a) The yellow rectangle and cyan points are the MTCNN output, and blue points are the Dlib landmarks. (b) The blue rectangle is ROI for calculating the signal.

TABLE 3. Notation of ROI extraction.

Notation	Mean
E_{L}	Left eye's position
E_{R}	Right eye's position
M_{L}	Left mouth's position
M_{R}	Right mouth's position
C_x	The x coordinate of the center point of the face frame
C_y	The y coordinate of the center point of the face frame
$Face_{width}$	The width of the center point of the face frame
$Face_{height}$	The height of the center point of the face frame

and PRV. The system executed an optical flow and took the average for each pixel in ROI to obtain three-channel average value R, G, and B after extracting ROI [25]. Then, according to the method of chrominance-based rPPG proposed by De Haan *et al.* in 2013, the R, G, and B signals were converted to acquire the rPPG signal [26].

C. SIGNAL PROCESSING

Fatigue is a unique physiological state, which is different from the waking state. When people feel tired, there are numerous distinct reactions such as blinking and yawning. In addition to the visual behavior, physiological signals also contain some information related to fatigue. Liang *et al.* pointed out that HR and HRV were strong indicators for fatigue estimation [27]. Hence, the following content explains how to obtain these features.

According to the research of Moco, rPPG signals arise from the blood volume change in the blood vessels, and this change is produced by heartbeats [28]. Therefore, after a system performs a fast Fourier transform (FFT) on the rPPG signal, there will exist a strong component on the spectrum

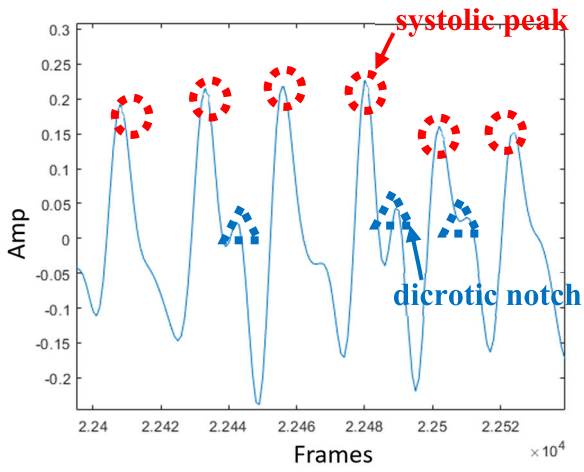

FIGURE 6. Systolic peak selection of the JMIPS. The red circles denote the systolic peaks of the rPPG signal, and the blue triangles denote the dicrotic notches of the rPPG signal.

representing the HR. However, an rPPG signal is highly susceptible to the environmental noise and motion signal. Our previous work proposed an algorithm, Motion Resistant Spectral Peak Tracking (MRSPT), to resist the motion signal and reduce the influence of environmental noise, which was helpful to predict the correct HR value [29].

1) PRV CALCULATION

The HRV is an indicator calculated by the interbeat interval (IBI) of the R peak in an ECG signal. The characteristic of the HRV can represent the relationship between the heartbeats and the fatigue state [27]. An ECG is necessary to calculate the HRV; however, in the existing methods, the ECG must be measured by contact devices. PRV is an indicator that uses the systolic peak of a pulse wave to replace the R peak in an ECG signal to obtain the IBI. It is calculated similarly as the HRV. N. Pinheiro *et al.* pointed out that for people who did not suffer from cardiovascular diseases, the result of the PRV was similar to that of the HRV [30]. Therefore, here, the PRV originating from the rPPG was selected to replace the HRV. Because the rPPG waveform can more easily be affected by the environment than the PPG, a robust peak selector was required. Here, the method called joint magnitude and interval features peak selection (JMIPS) was used to detect the systolic peaks of the rPPG and prevent it from being influenced by the dicrotic notches of the rPPG [31]. As shown in Fig. 6, JMIPS can detect the systolic peaks of the rPPG signal and avoid selecting the position of the dicrotic notches. This algorithm considers both the influence of the magnitude and time interval, and so regardless of the presence of a dicrotic notch, it can still decide if a reasonable systolic peak exists or not. After the system obtained the systolic peak positions, the HRV toolkit was used to calculate the PRV [32], and the standard deviations of all the normal to normal intervals (SDNN), power in low frequency (0.04–0.15 Hz) of the PRV (LF), power in high frequency (0.15–0.4 Hz) of the PRV (HF), total power in range 0–0.4Hz (TPW), and ratio of LF and HF were obtained.

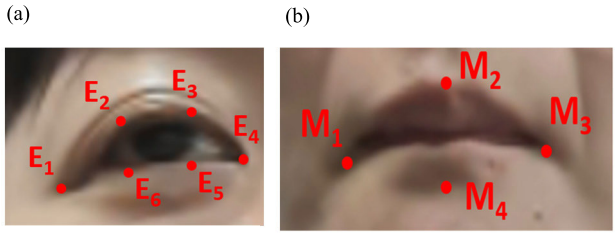

FIGURE 7. Landmarks detected by Dlib. (a) Six landmarks around the eye. (b) Four landmarks around the mouth.

2) BEHAVIOR DETECTION

Behavior can be observed mainly from the positions of the eyelids and the lips. With the eye aspect ratio (EAR) proposed by Soukupov'a, the closure degree of an eyelid can be decided [33]. The definition of EAR is

$$\delta = \frac{||E_2 - E_6|| + ||E_3 - E_5||}{2||E_1 - E_4||}, \tag{2}$$

where δ is EAR and E_1 to E_6 are the landmarks around the eyelid, as shown in Fig. 7 (a). After the system detects the landmarks from Dlib, blink detection can be performed by checking if EAR exceeds the threshold.

In addition to blink detection, Luo *et al.* proposed PERCLOS as an indicator of fatigue [4]. The definition of PERCLOS is

$$\varepsilon = \frac{\sum C_n}{\sum F_m}, \tag{3}$$

where ε is the notation of PERCLOS, C_n is the closed-eye picture in 1 min and F_m is the total number of the frames in 1 min. Then EAR can help to determine if the eyelids in the frame are close to the calculated PERCLOS.

Yawning is also a common reaction of fatigue, and it can be detected by the open degree of the mouth. The mouth aspect ratio (MAR) can be defined as the analogous method to EAR as

$$\zeta = \frac{||M_2 - M_4||}{||M_1 - M_3||}, \tag{4}$$

where ζ is MAR and M_1 to M_4 are the landmarks around the mouth, as presented in Fig. 7 (b). MAR can demonstrate the open degree of the mouth, and yawning can be detected by deciding the threshold of MAR.

D. FATIGUE ESTIMATION

Fatigue is a mixed physiological state. Prior detection is much more difficult than posterior detection. Posterior detection can be done after a specific phenomenon occurs. Prior detection should consider a large amount of information; however, which type of information correlates with fatigue has not been asserted. Besides, the relationship between the information and fatigue is also uncertain. Thus, the following content

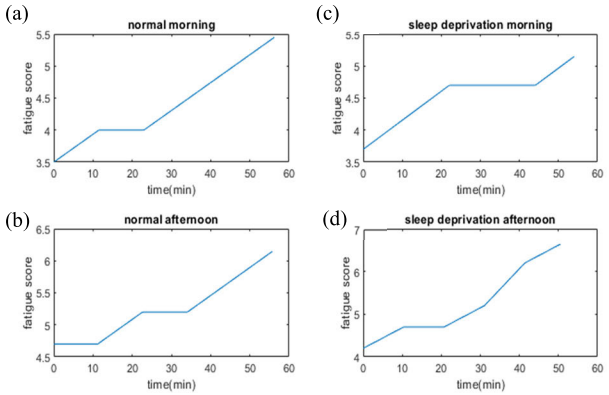

FIGURE 8. Mixture of the fatigue scores. (a) Morning fatigue score in the normal experiment. (b) Afternoon fatigue score in the normal experiment. (c) Morning fatigue score in the sleep deprivation experiment. (d) Afternoon fatigue score in the sleep deprivation experiment.

explains how to choose a reliable determination indicator and design a model that can learn the necessary information.

1) DEFINITION OF FATIGUE

The filled questionnaires provided by the subjects, such as SOFI, FAS, and VAS, were selected to define fatigue. As opposed to the methods that define the fatigue degree based on the experiment time slot or observe specific actions of the subjects, a questionnaire survey can better reflect the feelings of the subjects and reduce the influence of a wrong detection. Each type of questionnaire has its focus field of fatigue, and we mixed them into one score to evaluate the fatigue state. To avoid different questionnaires affecting the accuracy of the fatigue score, each score of the questionnaire was preprocessed by the min–max normalization defined as

$$y = \frac{10(x - x_{min})}{x_{max} - x_{min}}, \tag{5}$$

where x is the survey score of each questionnaire, x_{min} and x_{max} are the minimum and maximum of the questionnaire, respectively, and y is the normalized score. Then we summed the normalized scores of all the questionnaires for each subject and interpolated the fatigue score every second. The result is displayed in Fig. 8, and it is treated as an indicator to evaluate fatigue.

2) FEATURE SELECTION

A good feature could help model learning and reduce the calculation resource for processing the noise in the data. Eleven features were extracted from the signal to estimate the fatigue score.

The physiological signals contained a large amount of information that the conventional visualize method could not observe. However, such a method can detect the behavior and has been widely used to determine fatigue in the past. Therefore, the features detected by the conventional visualize method were also added in our feature sets. In addition, the fatigue state or physiological information varied from person to person, which highlights the importance of the individual parameters. The system combined the features from the physiological signal, conventional visualize method, and personal parameters, as listed in Table 4.

TABLE 4. Features notation and definitions.

Class	Notation	Detail	Observation Time
Physiological signals	H1	Mean of HR	1 min
	H2	Standard deviation of HR	
	V1	SDNN	5 min
	V2	Ratio of LF and HF	
	V3	LF	
	V4	HF	
	V5	TPW	
Conventional visualization method	B1	PERCLOS	10 sec
	B2	Yawning or not	
ersonal parameters	P1	Mean of HR	Whole experiment for each person
	P2	Standard deviation of HR	

TABLE 5. Parameters of SVR and ANN model.

SVR	Detail
Kernel function	Radial basis function
C	1
γ	1
ε	0.1
ANN	**Detail**
Hidden layer	5
Node	thrice the # of features
Activation function	Relu
Loss function	Mean square error

3) MODEL DESIGN AND TRAINING

Because the relationship between the used features and fatigue state was not extremely obscure, the complex learning methods may suffer the risk of overfitting. The supervised learning methods, support vector regression (SVR) and ANN, were used for training the regression models here. Both can determine the hidden relationship between input features and the target by either the optimized support vectors or the best weights between each node in the neural network. The setting of the methods is summarized in Table 5. The SVR selected the radial basis function as the kernel function, and set C to 1, gamma to 1 and epsilon to 0.1 by a grid search. The ANN used five hidden layers and each layer set thrice the number of nodes, and the activation function was Relu. The loss function chose the mean square error to minimize the variance and bias of the estimated result. During the training process, we separated 10% of the data as the validation data to determine when to terminate the training to avoid overfitting. Ten-fold cross-validation was used for verification. This validation

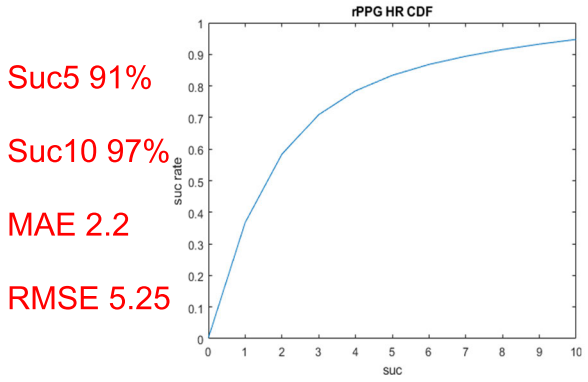

FIGURE 9. Estimated result of the rPPG HR.

could prevent training only specific data and avoided an inaccurate estimated result.

IV. RESULT AND DISCUSSION

A. HR ESTIMATION

1) EVALUATION METRICS

In several types of research, the physiological parameters originated from the PPG signal, and here, we used the rPPG signal to replace the PPG signal. Hence, comparison of the accuracies of the PPG and rPPG signals was first required. Here, we compared the accuracy of HR. The PPG signal and HR were measured using CMS50EW and then compared to the HR calculated by the rPPG signal. The percentage of the HR estimation difference within 5 or 10 bpm (Suc5 or 10), mean absolute error (MAE), and root mean square error (RMSE) were treated as the evaluation indices [34].

2) ESTIMATION RESULT

The result is shown in Fig. 9. Suc5 is 91% and Suc10 is 97%. The MAE and RMSE are also low in our dataset, which exhibits that using the rPPG signal and frequency domain peak selection can predict the HR appropriately.

B. PEAK SELECTION ESTIMATION

1) EVALUATION METRICS

The PRV is an interesting feature in the system, and it was also calculated by the PPG signal in previous research; therefore, we needed to compare the accuracy of the PRV. However, the PRV summarized numerous indicators with different calculation methods and each error peak position would result in different extent of effects in each indicator. Thus, the system chose the peak positions that were the origin of the PRV as the comparison object, instead of the PRV.

Considering the time difference in the blood flowing from heart to the face, the alignment between the rPPG and PPG peaks should be handled before evaluating the peak selector. Hence, first, the alignment between the rPPG and PPG peaks was calculated by cross-correlation using MATLAB. The performance of the peak selector was evaluated by the evaluation metrics quoted from JMIPS [31]. True positive

TABLE 6. Performance of the peak selector.

	Total beats	TP	FP	FN	Se (%)	PPV (%)	DER (%)
JMIPS	72525	55394	7567	9564	84.9	87.5	23.6

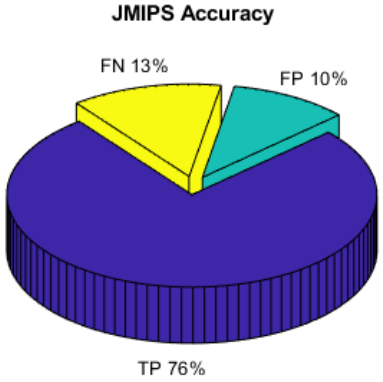

FIGURE 10. Pie chart of the peak selection accuracy.

(TP) implied the peaks of the rPPG aligned to the peaks of the PPG. False positive (FP) implied the peak selector chose the error position where the PPG peak did not exist. False negative (FN) implied the peak selector missed the peak of the PPG. Sensitivity (Se) is the hit rate of the PPG peaks; the positive predictive value (PPV) is the accuracy of the rPPG peaks, and the detection rate is the total error rate of the peak selector. The definitions of Se, PPV, and DER are respectively defined as

$$\text{Sensitivity } (Se) = \frac{TP}{TP + FN}, \tag{6}$$

$$\text{Positive Predictive value } (PPV) = \frac{TP}{TP + FP}, \tag{7}$$

$$\text{Detection Error Rate } (Se) = \frac{FP + FN}{TP + FN + FP}. \tag{8}$$

2) ESTIMATION RESULT

Owing to the elasticity and the resistance of the blood vessels, here, two indices of the peak alignment were accepted. The results are listed in Table 6 and presented in Fig. 10, where most of the peak positions of the rPPG are aligned with the peak positions of the PPG. However, some peaks mismatch the peaks of the PPG, owing to the FN increase and some peaks being situated at the dicrotic notch and response on FP. In Fig. 11, the blue lines represent the rPPG signal, the red circles are the peaks of the rPPG selected by JMIPS, and the blue triangles are the peaks of the PPG selected by JMIPS. Note that the positions of the peaks were adjusted by calculating the cross-correlation between the rPPG and PPG signal. In this experiment, the posture of the subjects remained intact, and in most of the cases corresponded to a stable case. The subjects kept looking straight ahead, and the faces remained stable; therefore, the rPPG signal is clean as shown in Fig. 11 (a).

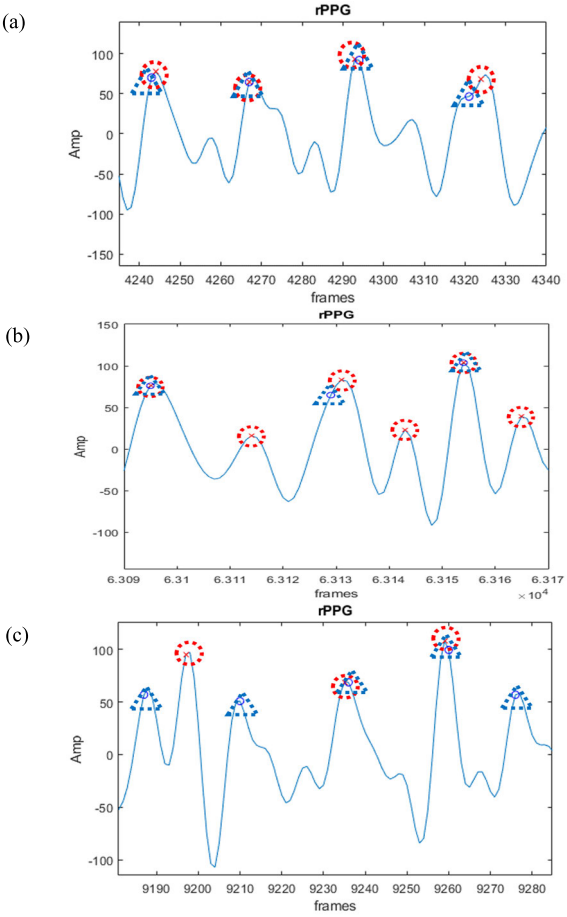

FIGURE 11. The selected peaks by JMIPS in the rPPG and PPG signals. (a) The case where rPPG and PPG peaks are matching. (b) The case where JMIPS detects a redundant dicrotic notch. (c) The case where rPPG and PPG peaks are mismatching or missing.

However, the time domain analysis can be easily affected. The subjects might have some subtle movement in their face that can cause an unstable signal like in Fig. 11 (b) and (c). This movement will enhance the noise, which either reinforces the dicrotic notch or conceals the specific waveform of the rPPG. The above conditions will cause JMIPS choose the error peak position, and so, need to be avoid. Hence, in the experiment, the subjects remained stable and tried to avoid violent motion. The TP of JMIPS reached 80%, which was sufficient to calculate the exact PRV in most cases.

C. FATIGUE PREDICTION

1) EVALUATION METRICS

To analyze the accuracy of the fatigue regression, correlation and RMSE were used for the prediction indices. Correlation is defined as

$$\rho(X, Y) = \frac{\sum_{m=1}^{N} (X_m - \mu_X)(Y_m - \mu_Y)}{(N-1)\sigma_X \sigma_Y}, \tag{9}$$

where X_m is the result of the regression, Y_m is the ground truth, N is the length of X and Y, μ_X and μ_Y are the mean of X and Y, respectively, and σ_X and σ_Y are the standard deviations of X and Y, respectively. The correlation value can illustrate the relevance between the regression result and ground truth. The value of correlation represents the existence of a close relationship between the result and ground truth when the absolute value of correlation is close to 1. Contrarily, the relationship becomes far as the absolute value of correlation approaches zero.

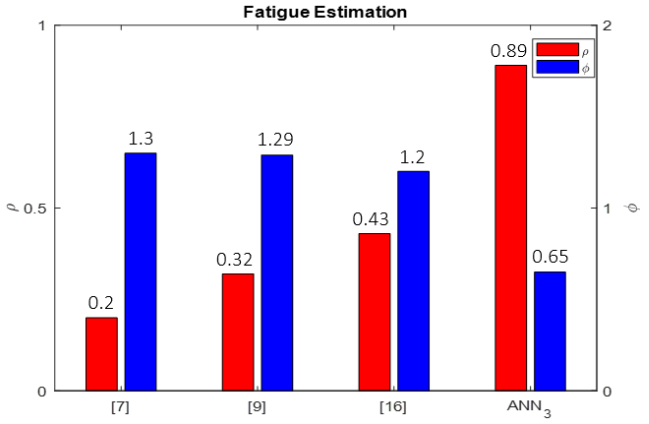

FIGURE 12. The histogram of correlation and RMSE for different fatigue estimation method.

The definition of RMSE is

$$\phi(\mathrm{X}, \mathrm{Y}) = \sqrt{\frac{\sum_{\mathrm{m}=1}^{\mathrm{N}} (\mathrm{X_m} - \mathrm{Y_m})^2}{\mathrm{N}}}, \tag{10}$$

where X_m is the result of the regression, Y_m is the ground truth, and N is the length of X and Y. Although correlation can reflect the relevance between the regression result and the ground truth, it cannot point out the prediction bias. The RMSE can resolve this issue, and as ϕ lowers, both the bias and variance of the result become lower, and the precision of regression is higher.

2) ESTIMATION RESULT

Fatigue is a cumulative phenomenon. A fatigue state gradually accumulates, and when the threshold is reached, it will cause yawning, blinking, and other behavior. Conventional visualization methods measure the behavior caused by fatigue. Physiological-based methods can find the change of state by observing physiological signals. However, the general rule to measure the fatigue state is difficult to find. Hence, personalization parameters can be expected to help in the estimation. To verify the above argument and find useful indicators for estimation, the following context presents the analysis of the performance of each method and the influence of each indicator.

Fig. 12 shows the performance of each method in fatigue estimation. The conventional visualization method was the one used in [7]. Only visualization-related features were selected to estimate the fatigue state. In [9], the PPG,

TABLE 7. Result of fatigue regression.

Name	Method	Correlation	RMSE	Selected Features
[7]	SVR_{rbf}	0.2	1.3	B1,B2
[9]	SVR_{rbf}	0.32	1.29	V1,V2,V3,V4,V5,H1,H2
[16]	SVR_{rbf}	0.43	1.2	V1,V2,V3,V4,V5,H1,H2,B1,B2
ANN_1	ANN	0.44	1.28	V1,V2,V3,V4,V5,H1,H2,B1,B2
ANN_2	ANN	0.8	0.84	P1,P2,H1,H2,B1,B2
ANN_3	ANN	**0.89**	**0.65**	V1,V2,V3,V4,V5,P1,P2,B1,B2
ANN_4	ANN	0.86	0.74	V1,V2,V3,V4,V5,P1,P2,H1,H2
ANN_5	ANN	0.85	0.75	V1,V2,V3,V4,V5,P1,P2,H1,H2,B1,B2

temperature, acceleration, and galvanic skin response were measured during the experiment, and the best combination was attempted to be chosen for fatigue estimation. However, the experiment in this study did not have the sensor to measure the temperature, acceleration, and galvanic skin response. Therefore, we selected the features that could be measured by the PPG sensor as the feature set. Conventional visualization and physiological features were mixed in [16].

Referring to Fig. 12, the conventional visualization method is not good at fatigue estimation. The visualization method can detect the phenomena caused by fatigue that might not suit for estimation. Although physiological information can be effectively used for fatigue estimation, the performance is still not very good. The key point is that fatigue is a personalized phenomenon. Personalization parameters can help to find the personal rule for fatigue estimation. Hence, in this study, we applied the personalization parameters, and the result is highly significant. The correlation coefficient of the method with the personalization parameters and physiological information is 0.89, which is the best result. The correlation coefficient of the physiological-based method reached 0.43, which has a slightly large difference from the best result. Lastly, the conventional visualization method has a correlation coefficient of only 0.2.

Table 7 lists the performance and used features for each method. In this study, we were interested in checking which parameters were related to fatigue estimation. Hence, ANN_1 to ANN_5 chose different combinations of the parameters to train the model. The result of ANN_1 shows that the physiological parameters are difficult to utilize without the personalization parameters. In comparison to ANN_5, ANN_2 yields the worse result because the PRV-related parameters are removed from the feature list. It also verifies that the PRV is a strong index to predict fatigue in previous research [27]. The result of ANN_4 is not much different from that of ANN_5, which exhibits that the conventional visualization parameters play an unimportant role in fatigue estimation. Based on the above discussion, we can draw the following conclusions:

1. Personalization parameters play an important role when using physiological parameter in fatigue estimation.
2. PRV can effectively reflect the changes in the state of the human by observing the change of IBI.
3. Conventional visualization parameters are not suitable for fatigue estimation.

V. CONCLUSION

The paper proposed a system that could measure physiological information and specific behavior, and then used these features to estimate the fatigue state. Because fatigue is a personalized phenomenon, physiological parameters also needed the assistance of personalization parameters, and then the physiological parameters could exhibit their value. The system combined behavior, physiological information, and personalization parameters to estimate the fatigue state and achieved 0.89 correlation coefficient, which was the best result. Besides the personalization parameters, we also found that the PRV was an effective indicator for fatigue estimation, during the experiment. Contrastingly, the conventional visualization parameters could only detect the behavior after the fatigue occurred, which was not suitable for fatigue estimation, and it only achieved a 0.2 correlation coefficient in our experiment.

The proposed system was convenient and useful to the driver. All of the information could be obtained by one camera and did not affect the driver. Most of all, the system could inform the driver of his/her fatigue status in advance to avoid accident occurrence. These results contributed greatly to traffic safety.

ACKNOWLEDGMENT

The authors would like to express sincere appreciation to Taipei Medical University Joint Institutional Review Board (TMU-JIRB) and Ministry of Science and Technology, Taiwan (MOST). The TMU-JIRB approved the above experiment. Certificate of TMU-JIRB No is N201805008.

REFERENCES

[1] C. J. Murray and A. D. Lopez, "Alternative projections of mortality and disability by cause 1990–2020: Global burden of disease study," *Lancet*, vol. 349, no. 9064, pp. 1498–1504, May 1997.

[2] S. Ameratunga, M. Hijar, and R. Norton, "Road-traffic injuries: Confronting disparities to address a global-health problem," *Lancet*, vol. 367, no. 9521, pp. 1533–1540, May 2006.

[3] A. W. MacLean, D. R. T. Davies, and K. Thiele, "The hazards and prevention of driving while sleepy," *Sleep Med. Rev.*, vol. 7, no. 6, pp. 507–521, Jan. 2003.
[4] X.-Q. Luo, R. Hu, and T.-E. Fan, "The driver fatigue monitoring system based on face recognition technology," in *Proc. 4th Int. Conf. Intell. Control Inf. Process. (ICICIP)*, Jun. 2013, pp. 384–388.
[5] A. Punitha, M. K. Geetha, and A. Sivaprakash, "Driver fatigue monitoring system based on eye state analysis," in *Proc. Int. Conf. Circuits, Power Comput. Technol. [ICCPCT]*, Mar. 2014, pp. 1405–1408.
[6] K. J. Raman, A. Azman, V. Arumugam, S. Z. Ibrahim, S. Yogarayan, M. F. Azli Abdullah, S. F. Abdul Razak, A. H. Muhamad Amin, and K. Sonaimuthu, "Fatigue monitoring based on yawning and head movement," in *Proc. 6th Int. Conf. Inf. Commun. Technol. (ICoICT)*, May 2018, pp. 343–347.
[7] M. Li and H.-L. Meng, "A method of driver fatigue detection based on multi-features," *Int. J. Signal Process., Image Process. Pattern Recognit.*, vol. 8, no. 10, pp. 107–114, Oct. 2015.
[8] R. Ahmed, K. Emrul Kayes Emon, and M. F. Hossain, "Robust driver fatigue recognition using image processing," in *Proc. Int. Conf. Informat., Electron. Vis. (ICIEV)*, May 2014, pp. 1–6.
[9] M. Choi, G. Koo, M. Seo, and S. W. Kim, "Wearable device-based system to monitor a Driver's stress, fatigue, and drowsiness," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 3, pp. 634–645, Mar. 2018.
[10] K. T. Chui, K. F. Tsang, H. R. Chi, B. W. K. Ling, and C. K. Wu, "An accurate ECG-based transportation safety drowsiness detection scheme," *IEEE Trans. Ind. Informat.*, vol. 12, no. 4, pp. 1438–1452, Aug. 2016.
[11] S. Kar, M. Bhagat, and A. Routray, "EEG signal analysis for the assessment and quantification of driver's fatigue," *Transp. Res. F, Traffic Psychol. Behaviour*, vol. 13, no. 5, pp. 297–306, Sep. 2010.
[12] A. Sahayadhas, K. Sundaraj, and M. Murugappan, "Detecting driver drowsiness based on sensors: A review," *Sensors*, vol. 12, no. 12, pp. 16937–16953, 2012.
[13] H. Al-Libawy, A. Al-Ataby, W. Al-Nuaimy, and M. A. Al-Taee, "HRV-based operator fatigue analysis and classification using wearable sensors," in *Proc. 13th Int. Multi-Conf. Syst., Signals Devices (SSD)*, Mar. 2016, pp. 268–273.
[14] Y.-P. Huang, N. N. Sari, and T.-T. Lee, "Early detection of driver drowsiness by WPT and FLFNN models," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2016, pp. 000463–000468.
[15] H. Qi, Z. J. Wang, and C. Miao, "Non-contact driver cardiac physiological monitoring using video data," in *Proc. IEEE China Summit Int. Conf. Signal Inf. Process. (ChinaSIP)*, Jul. 2015, pp. 418–422.
[16] I. R. Tayibnapis, D.-Y. Koo, M.-K. Choi, and S. Kwon, "A novel driver fatigue monitoring using optical imaging of face on safe driving system," in *Proc. Int. Conf. Control, Electron., Renew. Energy Commun. (ICCEREC)*, Sep. 2016, pp. 115–120.
[17] E. Åhsberg, F. Garnberale, and A. Kjellberg, "Perceived quality of fatigue during different occupational tasks development of a questionnaire," *Int. J. Ind. Ergonom.*, vol. 20, no. 2, pp. 121–135, Aug. 1997.
[18] J. De Vries, "Assessment of fatigue among working people: A comparison of six questionnaires," *Occupational Environ. Med.*, vol. 60, pp. 10–15, Jun. 2003.
[19] K. A. Lee, G. Hicks, and G. Nino-Murcia, "Validity and reliability of a scale to assess fatigue," *Psychiatry Res.*, vol. 36, no. 3, pp. 291–298, Mar. 1991.
[20] E. M. A. Smets, B. Garssen, B. Bonke, and J. C. J. M. De Haes, "The multidimensional fatigue inventory (MFI) psychometric qualities of an instrument to assess fatigue," *J. Psychosomatic Res.*, vol. 39, no. 3, pp. 315–325, Apr. 1995.
[21] A. W. S. Leung, C. C. H. Chan, and J. He, "Structural stability and reliability of the swedish occupational fatigue inventory among chinese VDT workers," *Appl. Ergonom.*, vol. 35, no. 3, pp. 233–241, May 2004.
[22] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1499–1503, Oct. 2016.
[23] D. E. King, "Dlib-ml: A machine learning toolkit," *J. Mach. Learn. Res.*, vol. 10, pp. 1755–1758, Jan. 2009.
[24] S. Kwon, J. Kim, D. Lee, and K. Park, "ROI analysis for remote photoplethysmography on facial video," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 4938–4941.
[25] W. Wang, S. Stuijk, and G. de Haan, "Exploiting spatial redundancy of image sensor for motion robust rPPG," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 415–425, Feb. 2015.
[26] G. de Haan and V. Jeanne, "Robust pulse rate from chrominance-based rPPG," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2878–2886, Oct. 2013.
[27] W. Liang, J. Yuan, D. Sun, and M. Lin, "Changes in physiological parameters induced by indoor simulated driving: Effect of lower body exercise at mid-term break," *Sensors*, vol. 9, no. 9, pp. 6913–6933, 2009.
[28] A. V. Moço, S. Stuijk, and G. de Haan, "New insights into the origin of remote PPG signals in visible light and infrared," *Sci. Rep.*, vol. 8, no. 1, p. 8501, Dec. 2018.
[29] B.-F. Wu, P.-W. Huang, C.-H. Lin, M.-L. Chung, T.-Y. Tsou, and Y.-L. Wu, "Motion resistant image-photoplethysmography based on spectral peak tracking algorithm," *IEEE Access*, vol. 6, pp. 21621–21634, 2018.
[30] N. Pinheiro, R. Couceiro, J. Henriques, J. Muehlsteff, I. Quintal, L. Goncalves, and P. Carvalho, "Can PPG be used for HRV analysis?" in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 2945–2949.
[31] B.-F. Wu, Y.-Y. Yang, B.-R. Tsai, P.-W. Huang, Y.-C. Tsai, and K.-H. Chen, "Remote HeartRate measurement based on signal feature detection in time domain," in *Proc. Int. Conf. Syst. Sci. Eng. (ICSSE)*, Jul. 2019, pp. 88–93.
[32] B. S. J. E. Mietus and M. D. A. L. Goldberger. (2009). *Heart Rate Variability Analysis with the HRV Toolkit*. [Online]. Available: https://physionet.org/tutorials/hrv-toolkit/
[33] T. Soukupová and J. Cech, "Real-time eye blink detection using facial landmarks," in *Proc. 21st Comput. Vis. Winter Workshop*, 2016, pp. 1–55.
[34] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7–11, Jan. 2011.

YIN-CHENG TSAI (Student Member, IEEE) received the B.S. degree in electrical and computer engineering from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 2018. His researches are focused on image-based physiological signal processing and driver fatigue detection. He is also interested in algorithm, digital signal processing, and artificial intelligence.

PENG-WEN LAI (Student Member, IEEE) received the B.S. degree in surveying engineering from the National Chung Cheng Institute of Technology (CCIT), Taoyuan, Taiwan, in 1991. He is currently pursuing the Ph.D. degree in electrical control engineering with NCTU. His research is focused on image-based physiological signal measurement and biomedical signal processing. He is also interested in artificial intelligence, data analysis, and image processing.

PO-WEI HUANG (Student Member, IEEE) received the B.S. degree in electrical and computer engineering from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 2017, where he is currently pursuing the Ph.D. degree in electrical control engineering. He received the IEEE SMC Outstanding B.Sc. Theses Grant, in 2017. His research is focused on image-based vital signs measurement.

TZU-MIN LIN received the M.D. degree from Fu Jen Catholic University, New Taipei City, Taiwan, in 2009. She is currently pursuing the Ph.D. degree in electrical control engineering with National Chiao Tung University (NCTU), Hsinchu, Taiwan. She received internship with the National Taiwan University Hospital, Taipei City, Taiwan, from August 2009 to July 2010, resident specialized in internal medicine with the Taipei Medical University Hospital, Taipei, from August 2010 to July 2013. She received a Chief Resident and a Fellowship in rheumatology, allergy, and immunology with the Taipei Medical University Hospital, Taipei, from August 2013 to January 2016. She has been working as an Attending Physician with the Department of Rheumatology, Immunology, and Allergy, Taipei Medical University Hospital, since February 2016.

BING-FEI WU (Fellow, IEEE) received the B.S. and M.S. degrees in control engineering from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1981 and 1983, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1992.

Since 1992, he has been with the Department of Electrical and Computer Engineering, where he was promoted to be a Professor, in 1998, and a Distinguished Professor, in 2010, respectively. He founded and has served as the Chair of the Taipei Chapter of IEEE Systems, Man, and Cybernetics Society (SMCS), in 2003, and was the Chair of the Technical Committee on Intelligent Transportation Systems of IEEE SMCS, in 2011. He has served as the Director of the Institute of Electrical and Control Engineering, NCTU, in 2011. He was the President of the Taiwan Association of System Science and Engineering and the Director of the Control Engineering Program, Ministry of Science and Technology, Taiwan, both in 2019. He is currently an Associate Editor of the IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS: SYSTEMS. His research interests include image recognition, vehicle driving safety and control, intelligent robotic systems, intelligent transportation systems, and multimedia signal analysis.

Dr. Wu is a Fellow of IET and CACS. He received many research honors, including the Outstanding Automatic Control Engineering Award from the Chinese Automatic Control Society, in 2007, the Outstanding Research Award of Pan Wen Yuan Foundation, in 2012, the Best Paper Award in The 12th International Conference on ITS Telecommunications, in 2012, the Best Technology Transfer Contribution Award from National Science Council, Taiwan, in 2012, the National Invention and Creation Award of Ministry of Economic Affairs, Taiwan, in 2012 and 2013, respectively, the Technology Invention Award of Y. Z. Hsu Scientific Award from Y. Z. Hsu Foundation, in 2014, the Outstanding Research Award, in 2015, all from Ministry of Science and Technology, Taiwan, and the FutureTech Breakthrough Award, in 2017 and 2019.