

A Motion Robust Remote-PPG Approach to Driver's Health State Monitoring

Bing-Fei Wu¹, Yun-Wei Chu¹(✉), Po-Wei Huang²,
Meng-Liang Chung¹, and Tzu-Min Lin³

¹ Institute of Electrical Control Engineering,
National Chiao Tung University, Hsinchu, Taiwan
yunwei@cssp.cn.nctu.edu.tw

² Department of Electrical and Computer Engineering,
National Chiao Tung University, Hsinchu, Taiwan

³ Division of Rheumatology, Immunology and Allergy,
Department of Internal Medicine,
Taipei Medical University Hospital, Taipei, Taiwan

Abstract. With the surging significance of personal health care, driver's physiological state is no longer negligible nowadays. Among all the indicators of health state in human, heart rate (HR) is one of the most cardinal indicators. The commonly used HR measurement is contact-type, might result in driver's distraction and discomfort in the vehicle applications. To cope with this problem, remote photoplethysmography (rPPG) is utilized to monitor HR at a distance via a web camera. Nevertheless, the rPPG is not without its flaw. The main concern of the rPPG technique is the potential not-robustness result from the arbitrary motion. Consequently, the contribution of this paper is to conquer the motion noise when the car is driving and the driver's health state is well monitored to enhance the public safety. The proposed algorithm is investigated in not only the indoor environment but as well the outdoor driving, which contains much more unpredictable motion. With k-nearest neighbor (kNN) classifier on chrominance-based features, the mean square error can be reduced from 30.6 to 2.79 bpm, approaching the medical instrument level. The proposed method can be applied to human improving driving safety for Advanced Driver Assistance Systems.

1 Introduction

Advanced Driver Assistance Systems (ADAS) are utilized to assist drivers in their driving process and have gathered increasing attention in the public recently. With external and internal imaging process algorithms, applications of computer vision in ADAS provide a more convenient and effective way to improve the intelligence of vehicles.

The imaging processing to detect the driving behavior and/or healthy condition of a driver has generated relatively little discussion in the considerable amount of research work for vision-based ADAS around the vehicle. Health care of drivers, as a typical issue of driver's behavior, can be of enormous value to public traffic safety. Consequently, this

work is to address the monitoring the driver's health state by a camera to enhance the public safety.

Along with the surging significance of personal health care, driver's physiological state is no longer negligible nowadays. Among all the indicators of health state in human, heart rate (HR) is one of the most significant indicators. In contrast with conventional techniques such as Electrocardiography (ECG), Photoplethysmography (PPG) is one of the technique that is capable of not only avoiding the skin irritation but as well providing a non-invasive and user-friendly recording. PPG, first described by Alrick Hertzman in 1937 [1], provides an electro-optic technique that detects the transmitted light on human's skin. The measurement sensed the periodic variation of light, which caused by the absorption difference when blood flows through the vessel under the skin (Zijlstra [2], Cheang [3], Allen [4]). Owing to PPG's simple, low-cost and non-invasive properties, contact sensors are commonly used to detect the variation of blood volume (e.g. the finger clipping device). Nevertheless, PPG is not without its flaw. As the skin contact is still required in practice, it may not suit for all scenario in real life (e.g. burn or contusion patient). Furthermore, applying PPG in vehicles to detect drivers' HR might cause drivers' distraction and discomfort.

Recently, several work have reported that even with a distance, blood volume variations during the cardiac cycle can also be detected from human's face with a web camera. This non-contact, image-based method for monitoring driver's HR is an exceptional approach that provides driver better flexibility and comfort experience during driving. In this paper, we apply the computer vision method on vehicles in order to estimate the driver's health and elevate the driving safety without interfering driver.

Lately, a non-contact and remote heart rate measurement has published by Poh et al. [5] who analyzed cardiac pulse from the three channel color video based on Independent Component Analysis (ICA), which is known as a Blind Source Separation (BSS) technique, leading to remote PPG (rPPG). The basic thought of BSS method has been applied to recover a clean signal (or called source) from a set of observations that are the linear combinations of the implicit source. This concept was expanded by Lewandowska et al. [6], who utilized Principal Component Analysis (PCA), a BSS method as well, claimed to reduce computational complexity in compared with ICA method. From that time, ample relative BSS based work has been reported to evaluate the retrieved HR signal more robustly from the video (Benjamin D Holton [7], Daniel Wedekind [8]). Mannapperuma [9] illustrated a comparison of different ICA methods (Fast ICA, JADE ICA, RADICAL ICA) and recorded the limitations of ICA-based HR detection in rPPG. Nonetheless, the accuracy of rPPG technique was highly affected by the motion of the subject up to now. To conquer the arbitrary motion challenge, Wenjin Wang [10] implemented face detection and tracking technique, Farneback dense optical flow algorithm [11] and PCA method to overcome the influence of global motion (due to face shifting and the rotating) and local motion (due to blinking and talking). Hamed Monkarezi [12] replicated Poh's ICA method and implemented the machine learning technique aimed to improve the robustness of subject motion in three different scenarios (resting, naturalistic HCI, in door cycling). Though these BSS-based methods retrieved independent and clean sources from the compound observations, there was no immediate way to determine which separated signal is the HR signal. All previously published BSS-based literatures selected

the most periodic signal from a collection of the independent signal as HR signal, resulting in distortion when arbitrary motion occurred.

de Haan and Jeanne [13] presented a chrominance-based rPPG (CrPPG) that focused on improving motion robustness. They eliminated the component selection issue, which mattered in the BSS-based method, by constructing a linear combination of the normalized color signals orthogonal to the assumed distortions, regardless the color of the illumination. After that time, abound relative works (de Hann [14], van Gastel M [15]) have continued improving motion robustness based on CrPPG in indoor cycling and stepping device scenario. Ren-You Huang [16] extended the concept of signal recovery from CrPPG and integrated with the ICA method, trying to separate a cleaner HR signal in indoor treadmill scenario from observation of face's x-position, y-position and CrPPG. Yung-Chien Hsu [17] also based on CrPPG method and utilized support vector regression (SVR), a machine learning technique, for predicting HR more accurately in indoor naturalistic HCI scenario.

Most of the previous work aimed to enhance the motion robustness of rPPG or CrPPG in the indoor and controllable scenario; nonetheless, the restriction of the scenario environment will limit the application of this promising technique. Consequently, the aim of this paper is to provide an enhanced motion-robust rPPG technique in different scenarios, especially the outdoor driving vehicles scenario for monitoring driver's heart rate in real time. In this paper, we reduce the noise result from artificial motion by conducting Empirical Mode Decomposition (EMD) based on CrPPG method and predict the HR by k-Nearest Neighbors (kNN), a machine learning technique for classification problem. The results confirm that the proposed algorithm can reduce the error between the predicted HR and actual HR, which is detected by Scosche Rhythm+ [18] during the experiment as ground truth. The accuracy of the proposed algorithm reduce the error from 30.6 to 2.79 bpm in the outdoor and uncontrollable sunlight environment during driving. The proposed application of rPPG on vehicles can be applied in intelligent vehicles system and monitor the driver's heart rate continuously, ensuring driver's health and safety without drivers' awareness.

The remainder of this paper is organized as follows. Sections 2 and 3 introduce our HR method and the experimental scenario respectively. The experimental results and comparisons are illustrated in Sect. 4. Ultimately, conclusions are summarized in Sect. 5.

2 Our Approach

In this section, the details of our algorithms are depicted. In addition, the estimation problems and corresponding solutions are introduced as well. The diagram of the proposed algorithm is shown in Fig. 1 First, a video sequence containing the participants' face is regarded as the input images. Then, a time domain raw signal, CrPPG, was obtained from the linear combination of the input's RGB channel. For reducing noise of raw signal, bandpass filter and Empirical Mode Decomposition (EMD) were utilized. Ultimately, k-nearest neighbor (kNN) classification, the machine learning based method, estimates HR from the frequency domain features. Each algorithm is

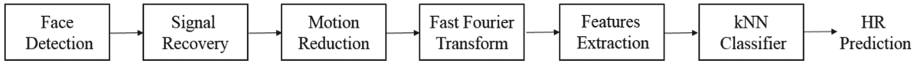


Fig. 1. Flowchart of proposed machine learning method for HR extraction from video recording.

discussed in detail in the following sections. All the experimental videos were recorded by a web camera (Logitech C920R).

2.1 Face Detection and HR Signal Recovery

Heart rate detection based on rPPG requires monitoring the skin region, especially the face. Dlib [19], the open toolkit commonly used in both industry and academia for a wide range of domains, was used for frontal face detection. At first, the whole frame is detected to figure out the entire face region as our region of interest (ROI). To lessen the computational load, the face region of the latter frame is based on the 20% expansion of the previous face ROI (fROI). The following HR estimation regard proportion of the fROI, which denoted as pROI, as input images. Then we compute the mean value of all pixels in pROI on R, G, B channels, which denoted as $\mu(R)_i, \mu(G)_i, \mu(B)_i$, where $i = 1, 2, 3, \dots$ representing the different frames. A realization of de Haan and Jeanne’s CrPPG method linearly combines the three channels as follows:

$$X_i = 3\mu(R)_i - 2\mu(G)_i \text{ and} \tag{1}$$

$$Y_i = 1.5\mu(R)_i + \mu(G)_i - 1.5\mu(B)_i \tag{2}$$

HR signal, denoted as S , computed by Eqs. (1) and (2) and illustrated as follows:

$$S = X_f - \alpha Y_f \tag{3}$$

with

$$\alpha = \frac{\sigma(X_f)}{\sigma(Y_f)} \tag{4}$$

where $\sigma(X_f)$ and $\sigma(Y_f)$ are the standard deviations of X_f and Y_f , respectively, and X_f and Y_f are the length-64 collections of X_i and Y_i . Nevertheless, the algorithm suffers from several limitations of face detection. First, the angle of the face rotation might exceed the limitation of the face detection algorithm. To cope with this kind of missing face problem, we use skin detection based on the previous fROI to ensure the HR signal’s continuity. Secondly, as a result of arbitrary motion, the pROI is not certainly the exact same portion of face. Consequently, the noise reduction algorithms are implemented.

2.2 Filtering and Noise Reduction

The original CrPPG signal is frequently polluted by unpredictable channel noise, especially in arbitrary motion scenario. Consequently, the noise reduction process is a must-have. Two steps to eliminate the noise are summarized as follow. First, since human HR is in the range between 0.7 Hz and 3Hz, an FIR Band Pass Filter (BPF) along with Hamming Window is utilized to extract the HR signal. Secondly, as the noise of which frequency band is the same as HR cannot be eliminated by the BPF, Empirical Mode Decomposition (EMD) is utilized to split the polluted signal into HR and noise.

For the design of BPF, the cutoff frequencies are 0.7 Hz and 3Hz with 128 order. The time domain and frequency domain results are illustrated in Fig. 2(a). If there were little noise, the peak frequency would represent the HR. Nonetheless, although the BPF singles out the signal of which frequency energy lay on the range from 0.7 to 3 Hz, there could still be some short duration but strong noise in the same frequency band, as shown in Fig. 2(b). This sort of noise appears frequently in the motion scenario and would confuse the choice of peaks frequency, leading to low estimation accuracy.

So as to reduce the unpredictable noise caused by arbitrary motion, EMD serves as a nice filter to extract the main component of the signal. EMD is used to decompose the signal to several unique intrinsic mode functions (IMFs) and one residue function. The iteration is shown in Eq. (5).

$$\tilde{g}(t) = \sum_{i=1}^n IMF_i(t) + R_n(t) \tag{5}$$

Where $IMF_i(t)$ is the i^{th} IMF at time step t , $R_n(t)$ is the residue function at time step t , and n is the number of EMD iterative. The EMD ensures that every IMF is

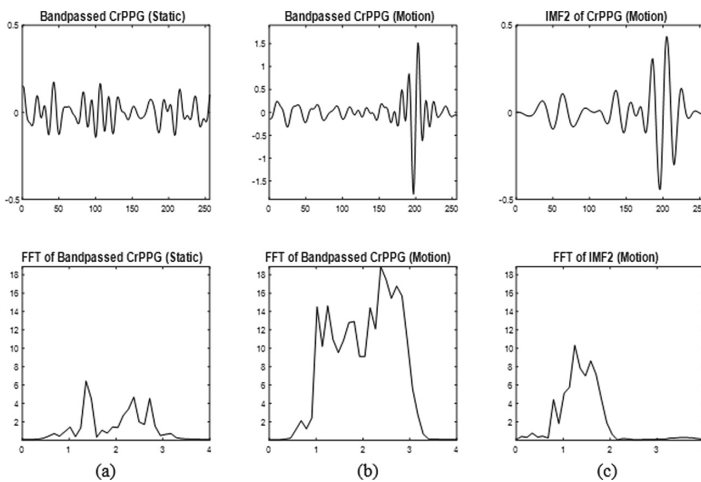


Fig. 2. The comparison of band passed signal and IMF in static and motion scenarios.

symmetrical with zero mean and the same amounts of peaks and zero-crossing points. With the EMD process, the main component of signal in the Fig. 2(b) is extracted as illustrated in Fig. 2(c), making the peak frequency more lucid. As a result, compared with the original band passed signal, the extracted IMF signal serves as more reliable input variables of the following kNN process.

2.3 k-nearest Neighbor Classification and Prediction

Most of the previous work estimate their accuracy by comparing the max peak among all frequency band (MPA, Maximum Peak among All) as an indicator. Nevertheless, the channel noise in motion would distort the frequency response, resulting in several pseudo peak frequencies. The kNN classifier [20] which figures out the closest distance between the testing data and the categories of the training data is implemented in our method. In contrast with MPA, kNN based frequencies selection is more resistant to the distortion of signal. Due to kNN's high accuracy and strong ability to classify unknown and non-Gaussian distribution data, it is of high suitability to our study. Also, its simple and instant property enhances the feasibility of abound real-time applications such as ADAS.

The top five peaks of the FFT spectrum are regarded as the features for the kNN classifier ($k = 1$). Since kNN is a supervised learning model, the real HR as well as frequency features are required for training. With the assistance of EMD and bandpass filter, the better features are extracted in the frequency band and enhance our classifier. For each volunteer, 40000–45000 training data and 12000–13000 testing data are recorded. With the predicted HR, a Kalman filter is utilized to achieve the smooth and optimal estimation of real HR and avoid the unreasonable huge variation between short time duration. Ultimately, the accuracy of our method is investigated by mean square error (MSE), maximum error (ME) and root mean squared error (RMSE).

3 Experiment Setup

Our experiment contains three scenarios. The first and the second scenarios were evaluated in an indoor environment. First, the participants were allowed to sit in front of the computer casually. In the second study, participants were asked to sway their bodies that caused strong artificial motion. The third study was measured in the driving, at which encountered unpredictable artificial motion.

3.1 First scenario—Sitting Casually

Seven volunteers (five males and two females, age between 22 and 25) participated in our first and second studies. All participants are seated in front of the same computer running Windows 10 in an indoor environment. Video recording was carried out using a web camera (Logitech C920R) mounted on the screen in front of the participants at 80 cm (see Fig. 3). The only illumination source was the ambient ceiling light. All videos were recorded at 30fps in 24-bit RGB color at 1280×720 resolution and saved



Fig. 3. The experimental environment of the first and the second studies. The camera was mounted on the computer in front of the participants. Scosche Rhythm+ is banded on participants' wrist to get the heart rate data simultaneously when being video recorded.

in BitMaP (BMP) format. During the experiment, video sequences and the real heart rate were recorded simultaneously. Participants wore Scosche Rhythm+, which inseparably banded on their wrist and transmitted with Bluetooth in 30 Hz, to get the heart rate data when being video recorded. The participants were allowed to either talking on the cell-phone or eating in front of the computer during the measurement. The experiment duration was about 10 min for each participant.

3.2 Second scenario—Strong Artificial Motion

This study was conducted using the same instruments used in the first study. All the participants were asked to swing their body so as to conduct a strong motion scenario. This study was conducted using the same instruments used in the first study. The movement even caused the face's location to be variable. The location of the face was recorded as X-axis (horizontal position) and Y- axis (vertical position). We measured the range and the standard deviation of the face's position as the indicator of the motion. This experiment duration was approximately 10 min for each participant.

3.3 Third scenario—Car Driving

In this study, two males from above studies participated in the car driving with different road section. Same camera was mounted on instrument panel below the front of the driver at the distance of 80 cm. The frames captured by the camera were the driver's face with a slight elevation. Participants as well wore Scosche Rhythm+ to get the heart rate data when being recorded (see Fig. 4). The driving involved flat and slightly uphill asphalt road section in Hsinchu. Driver's face swung while turning the vehicle or sometimes driving through a bumpy road. The illumination source was sunlight that shines in at the windshield. Two participants drove 4.4 and 5.2 km respectively. We divided the driving process to several intervals and classified in two states, one is the



Fig. 4. The experimental environment of the third study. The camera was mounted on instrument panel below the front of the driver. Scosche Rhythm+ is also banded on participants’ wrist to get the heart rate data simultaneously when being video recorded.

Table 1. Frames of each interval and the traffic states for the first participant are recorded as below.

	# of frames	State
Interval I	5290	Driving
Interval II	4580	Stops at the traffic light
Interval III	4220	Driving
Interval IV	3810	Stops at the traffic light
Interval V	2500	Driving

Table 2. Frames of each interval and the traffic states for the second participant are recorded as below.

	# of frames	State
Interval VI	2700	Driving
Interval VII	800	Stops at the traffic light
Interval VIII	7800	Driving
Interval IX	1700	Stops at the traffic light
Interval X	2100	Driving
Interval XI	800	Stops at the traffic light
Interval XII	9159	Driving

interval while driving and the other is the interval while waiting at the traffic light. Two participants drove in different road section and divided to separate intervals recorded in Tables 1 and 2, respectively.

4 Experimental Results

For each participants, the kNN model consisted of 200000–225000 training data. The testing data used the user-dependent training model to estimate the HR. According to the experimental setup, there are three scenario leaving for discussion. The mean square error (MSE), maximum error (ME) and the root mean square error (RMSE) between the predicted HR and the ground truth are regarded as performance criteria. In the meantime, the comparison with MPA, which is commonly used in other rPPG works to compare the efficiency, is illustrated.

4.1 Result for Sitting Casually

For the first scenario, participants were asked to act normally like eating or talking on the cellphone in front of the camera, which is mounted on top of the computer. The video sequence of each participant consists of 8800–12300 frames as testing data. The standard deviation of the x and y coordinates, representing face's movement of all participants were 11.1 and 2.85 pixels respectively. Though the face movement of each participant was quite small, the MSE of the MPA method still can up to 13.05. The noise reduction method can eliminate the wrong feature for training the kNN model. MSE of the model can be reduced to 2–4 bpm, compared to 3–13 bpm of MPA (see Table 3).

Table 3. This table shows the motion level by recorded the face movement of each participant in the first scenario. Therefore, the comparison of our proposed method and MPA was measured in MSE, ME and RMSE.

ID	#of frames	Range of x	Std of x	Range of y	Std of y	Method	MSE	ME	RMSE
1	9201	94	6.4	96	1.7	MPA	5.00	55.21	0.02
						kNN	3.23	14.5	0.008
2	12310	203	16.3	123	2.4	MPA	13.05	42.01	0.02
						kNN	4.32	12.41	0.048
3	8801	196	13.2	158	3.4	MPA	10.30	49.62	0.047
						kNN	3.30	11.23	0.041
4	9101	179	10.2	193	4.2	MPA	3.61	15.62	0.037
						kNN	2.00	6.89	0.024
5	9880	211	10.6	154	3.8	MPA	9.81	50.34	0.03
						kNN	3.41	11.3	0.007
6	10300	262	11.3	169	2.9	MPA	11.61	44.50	0.04
						kNN	2.29	12.68	0.013
7	9760	164	9.8	106	1.6	MPA	10.16	38.51	0.069
						kNN	2.91	10.67	0.037

4.2 Result for Strong Artificial Motion

For the second scenario, participants were asked to swing their body so as to conduct a strong motion scenario with the same instruments in first scenario. The goal of this experimental setup was to evaluate the robustness of our HR method with strong artificial motion. The video sequence of each participant regards 9100–10500 frames as testing data. The standard deviation of the x and y coordinates representing the face's movement of all participants were 46.1 and 14.37 pixels respectively, which were much stronger than that in the first scenario. The strong motion of the face would lead FFT spectrum of the time sequences into chaos. MPA regards the strongest but the wrong peak in the distorted spectrum as HR and the MSE could arise up to 26 bpm. With the user-dependent model, kNN method predicts the HR that's more suitable for corresponding user. The MSE of our proposed method can reduced to 2–8 bpm in comparison with 15–26 bpm of the MPA method (see Table 4).

Table 4. This table shows the stronger motion level compared with the first scenario. Therefore, the comparison of our proposed method and the MPA method was measured in MSE, ME and RMSE.

ID	#of frames	Range of x	Std of x	Range of y	Std of y	Method	MSE	ME	RMSE
1	10565	229	21.3	73	6.7	MPA	20.78	49.21	0.247
						kNN	8.95	23.81	0.102
2	10400	361	49.6	196	15.3	MPA	15.19	54.01	0.179
						kNN	2.75	13.10	0.03
3	9135	413	51.2	217	13.4	MPA	26.24	65.81	0.298
						kNN	2.59	11.56	0.036
4	9700	395	56.1	226	16.2	MPA	15.23	59.21	0.214
						kNN	7.79	27.12	0.116
5	10100	334	53.9	234	19.7	MPA	24.31	67.21	0.243
						kNN	2.44	10.43	0.041
6	9665	449	54.6	236	17.0	MPA	15.62	51.30	0.164
						kNN	6.62	19.6	0.093
7	9890	294	36.4	116	12.3	MPA	24.3	49.80	0.261
						kNN	2.23	10.62	0.039

4.3 Result for Car Driving

In the third scenario, two participants drove the same vehicle in different road section. One drove 4.4 km and the other drove 5.2 km, and the recorded video contains 20501 and 25326 frames respectively. Therefore, the driving was divided into several intervals and two states, one state is the intervals recorded in driving scenario and the other is the interval recorded while waiting at the traffic light. The MSE, ME and RMSE for each interval are regarded as the performance criteria. The same devices of the previous scenarios are used, but the position of the camera was below the front of the participant

that mounted on instrument panel. The goal of this experimental setup was to evaluate the feasibility of the proposed algorithm in an outdoor driving scenario.

The standard deviation of the x-direction movement of the first and the second participant were 20.54 and 13.7 pixels, and 6.59 and 10.6 pixels in y-direction respectively. We used the same training model as the previous scenarios, and the recorded data were tested by participants’ own model for predicting HR. With the proposed algorithm, the result shows the advance of robustness in uncontrollable motion during car driving. The MSE can reduce to 2–6 and 3–7 bpm for the first and the second participant respectively. The corresponding intervals and the MPA method’s MSE are recorded in Tables 5 and 6. Figures 5 and 6 depict the ground truth which

Table 5. We divided five intervals for the first participant’s car driving. The error measurements for each interval were reported in this table.

Interval	#of frames	Range of x	Std of x	Range of y	Std of y	Method	MSE	ME	RMSE
I	5290	143	22.6	38	6.47	MPA	30.6	68.0	0.46
						kNN	2.79	7.27	0.04
II	4580	161	16.7	36	4.81	MPA	22.9	51.2	0.39
						kNN	4.1	11.5	0.07
III	4220	130	20.4	38	5.97	MPA	28.1	64.2	0.50
						kNN	3.57	12.6	0.06
IV	3810	145	21.2	86	9.32	MPA	20.5	60.2	0.41
						kNN	5.62	14.0	0.10
V	2500	155	21.8	46	6.41	MPA	24.1	55.4	0.54
						kNN	6.61	13.4	0.14

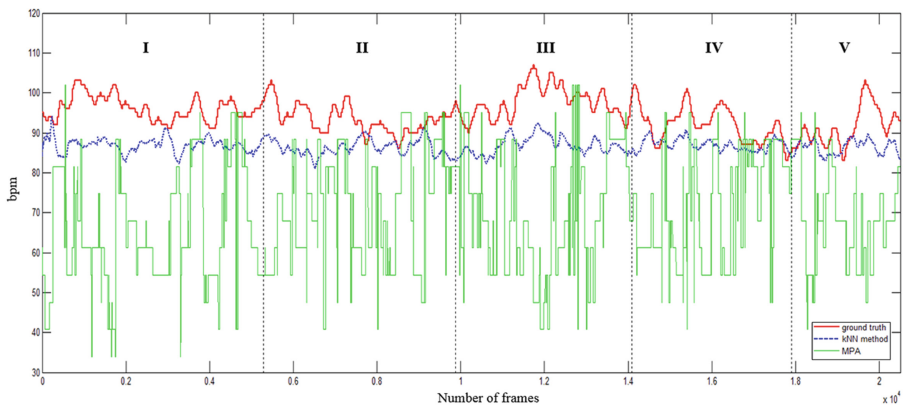


Fig. 5. The red bold line represents the first driver’s real HR and the blue dotted line is the HR predicted by our method. The compared method, MPA, is illustrated in green solid line and suffered abruptly drop and climb during the process (Color figure online).

Table 6. We divided seven intervals for the first participant’s driving. The error measurements for each interval were reported in this table.

Interval	#of frames	Range of x	Std of x	Range of y	Std of y	Method	MSE	ME	RMSE
VI	2700	100	12.9	54	9.19	MPA	11.3	40.4	0.27
						kNN	6.48	17.3	0.13
VII	800	82	10.6	51	9.67	MPA	12.6	46.1	0.31
						kNN	5.39	13.5	0.11
VII	7800	68	9.68	55	10.03	MPA	11.2	50.2	0.16
						kNN	3.58	13.2	0.05
IX	1700	128	21.2	66	11.3	MPA	10.0	29.4	0.32
						kNN	4.15	8.70	0.11
X	2100	74	11.0	64	10.9	MPA	14.3	33.4	0.36
						kNN	3.87	11.2	0.10
XI	800	34	8.21	53	12.4	MPA	17.9	38.4	0.70
						kNN	7.53	16.6	0.32
XII	9195	128	22.9	64	10.9	MPA	14.0	44.4	0.17
						kNN	4.35	17.0	0.05

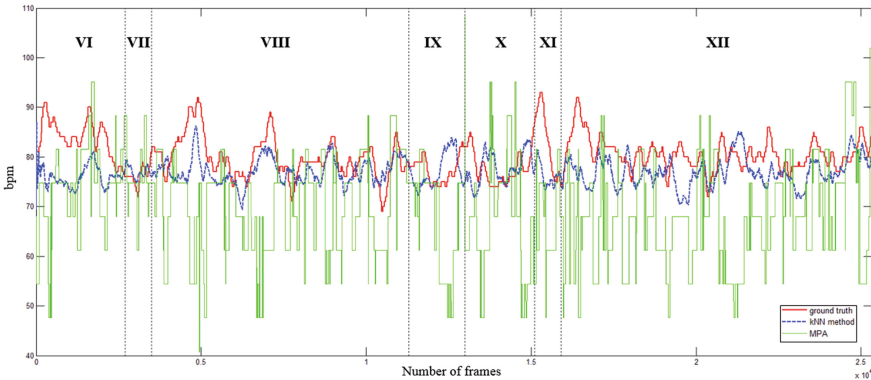


Fig. 6. The red bold line represents the first driver’s real HR and the blue dotted line is the HR predicted by our method. Our method can predict the trend of HR with enhanced accuracy in compared with MPA method, which is illustrate in green solid line (Color figure online).

denoted as the red bold line, the HR predicted by our method and MPA method which denoted as the blue dotted line and green solid line respectively for each participants. According to the figures, the HR predicted by the MPA method suffered some suddenly drop and climb during the car driving. While the HR evaluated by our method was not exactly equal to the ground truth, it closely follows the trend of real HR with higher accuracy. The final result and monitoring user interface was illustrated in Fig. 7.

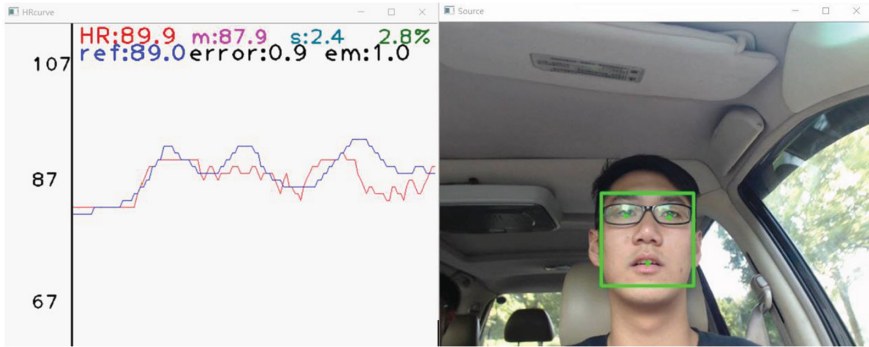


Fig. 7. This figure shows the user interface of the proposed HR monitoring system.

5 Conclusions and Future Work

In this paper, rPPG and our approach are investigated in three different scenarios: an indoor room with the participants sitting static and with strong motion respectively, and the outdoor car driving. To transfer the HR monitoring system from the indoor scenario to an outdoor vehicle requires the stronger anti-interference ability. Consequently, the goal of this paper is to improve motion robustness and increased the feasibility of rPPG method in vehicle application, in which the motion is stronger and unpredictable. To cope with the uncontrollable noise, the user-dependent kNN classifier can customize the HR for each user and reduce abruptly drop and climb phenomena as seen in the MPA method. The new approach of the machine learning based rPPG yields the more accurate and reliable result, reducing the RMSE from 0.208 to 0.036.

Although the accuracy of the proposed approach is increased, we still need to evaluate HR more precisely in other environment conditions. That is, the light changes as well affect the time domain signal during the car driving. So our next investigation will also determine the illuminance source variation in the outdoor environment. In addition, kNN is user-dependent classifier; therefore, the training process for new user is still required. The total solution of the classification model for every user should be considered in our future progress.

In conclusion, with the proposed internal imaging process, the algorithm is beneficial for ADAS. In contrast with contact-type HR sensors, the proposed rPPG can estimate the HR without the concern of distraction and discomfort. The application of computer vision in monitoring driver's heart rate can ensure driver's health and enhance the safety in public.

Acknowledgement. This work was supported by Ministry of Science and Technology under Grand no. MOST105-2622-E-009-013-CC2.

References

1. Hertzman, A.B.: Photoelectric plethysmography of the fingers and toes in man. *Exp. Biol. Med.* **37**, 290–292 (1937)
2. Zijlstra, W., Buursma, A.: Spectrophotometry of hemoglobin: absorption spectra of bovine oxyhemoglobin, deoxyhemoglobin, carboxyhemoglobin, and methemoglobin. *Comp. Biochem. Physiol. B* **118**, 743–749 (1997)
3. Cheang, P.Y.S., Smith, P.R.: An overview of non-contact photoplethysmography. *Electron. Syst. Control Div. Res.* 57–59 (2003)
4. Allen, J: Photoplethysmography and its application in clinical physiological measurement. *Physiol. Meas.*, **28**, R1 (2007)
5. Poh, M.-Z., McDuff, D.J., Picard, R.W.: Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Opt. Express* **18**, 10762–10774 (2010)
6. Lewandowska, M., Ruminski, J., Kocejko, T., Nowak, J.: Measuring pulse rate with a webcam—a non-contact method for evaluating cardiac activity. In: (IEEE) Proceedings of Federated Conference Computer Science and Information Systems, pp. 405–410 (2011)
7. Holton, B.D., Mannapperuma, K., Lesniewski, P.J., Thomas, J.C.: Signal recovery in imaging photoplethysmography. *Physiol. Meas.* **34**, 1499–1511 (2013)
8. Wedekind, D., Trumpp, A., Andreotti, F., Gaetjen, F., Rasche, S., Matschke, K., et al.: Assessment of source separation techniques to extract vital parameters from videos. In: 23rd European Signal Processing Conference, pp. 434–438 (2015)
9. Mannapperuma, K., Holton, B.D., Lesniewski, P.J., Thomas, J.C.: Performance limits of ICA-based heart rate identification techniques in imaging photoplethysmography. *Physiol. Meas.* **36**(1), 67–83 (2015)
10. Wang, W.: Exploiting spatial redundancy of image sensor for motion robust rPPG. *IEEE Trans. Biomed. Eng.* **62**(2), 415–425 (2015)
11. Farneböck, G.: Two-frame motion estimation based on polynomial expansion. In: Bigun, J., Gustavsson, T. (eds.) SCIA 2003. LNCS, vol. 2749, pp. 363–370. Springer, Heidelberg (2003). doi:10.1007/3-540-45103-X_50
12. Monkaresi, H., Calvo, R.A., Yan, H.: A machine learning approach to improve contactless heart rate monitoring using a webcam. *IEEE J. Biomed. Health Inf.* **13**, 1153–1160 (2014)
13. de Haan, G., Jeanne, V.: Robust pulse rate from chrominance-based rPPG *IEEE Trans. Biomed. Eng.* **60**, 2878–2886 (2013)
14. de Haan, G., van Leest, A.: Improved motion robustness of remote-PPG by using the blood volume pulse signature. *Physiological measurement* **35**(9), 1913–1926 (2014)
15. van Gastel, M.: Motion robust remote-PPG in infrared. *IEEE Trans. Biomed. Eng.* **62**(5), 1425–1433 (2015)
16. Huang, R.Y., Dung, L.R.: A motion-robust contactless photoplethysmography using chrominance and adaptive filtering, 1–4 (2015)
17. Hsu, Y.: Learning-based heart rate detection from remote photoplethysmography features. In: Proceedings of IEEE International Conference Acoustics Speech Signal Processing, pp. 4433–4437 (2014)
18. SCOSCHE, Rhythm+ (2015). <http://www.scosche.com/rhythm-plus-heart-rate-monitor-armband>
19. Dlib library, imaging processing toolbox (2016), <http://dlib.net/imaging.html>
20. Aha, D., Kibler, D., Albert, M.: Instance-based learning algorithms. *Mach. Learn.* **6**(1), 37–66 (1991)