Non-contact Heart Rate Monitoring: A Comparative Study of Computer Vision and Radar Approaches

Gengqian Yang[1][0000-0002-4759-4091], Benjamin Metcalfe[0000-0003-4279-8930], Robert Watson[0000-0002-6596-7568], and Adrian Evans[0000-0001-8586-8295]

Department of Electronic and Electrical Engineering, University of Bath, UK
{gy299,a.n.evans}@bath.ac.uk

Abstract. The Heart Rate (HR) is a vital sign that is used to assess the physical and mental state of an individual. There is a growing interest in incorporating HR measurement into Driver Monitoring Systems (DMS), providing physiological measurements to help address long-existing road safety issues by minimising human error. In real-world driving scenarios, the HR must be measured using non-contact approaches that avoid distracting or restricting the driver. The most common approaches to non-contact HR measurement use either computer vision (CV) or mm-wave radar, both showing acceptable performances in controlled studies. However, the relative merits of different sensor modalities for real-world scenarios remain unclear, and the potential benefits of a combined approach are unquantified. To address these questions, this paper first proposes and implements non-contact HR measurement architectures for both CV and mm-wave radar systems and characterises their HR estimation performance, using electrocardiography (ECG) to provide ground truth measurements. The effects of distance to sensors and of illumination variations on HR estimation are also studied, showing the relative errors for both modalities to be less than 0.5% for the distances found in practical DMS. These results also highlight the distinctive characteristics of each modality and the benefits of a multi-modality approach for DMS.

Keywords: Non-contact Heart Rate Monitoring, Remote Photoplethysmography, mm-wave Radar, Driver Monitoring Systems.

1 Introduction

Improving road safety remains a major challenge in the automotive sector, with more than one million deaths globally on the road each year and recent studies showing that human factors contributing in more than 90% of road accidents [1]. Two approaches to improving road safety are vehicle automation and driver monitoring using Driver Monitoring Systems (DMS). As the widespread adoption of fully autonomous vehicles is still many years away, driver behavior will continue to have an important safety role

[1] Gengqian Yang is supported by a scholarship from the EPSRC Centre for Doctoral Training in Advanced Automotive Propulsion System (AAPS), under the project EP/S023364/1. The underlying data for this paper is available at: https://github.com/GengqianYang/.

and pressures from both regulatory bodies and industries continue to drive the development of DMS.

As one of the vital signs, the Heart Rate (HR) is critical to the diagnosis of physiological and psychological states and is ubiquitous in medicine. Furthermore, recent studies demonstrating the correlation between HR and fatigue have reinforced the importance of vital signs in DMS [2]. In clinical settings, HR is continuously measured by either electrocardiography (ECG) or photoplethysmography (PPG), both of which are contact methods. However, due to the restrictions and distractions introduced by body-attached sensors, non-contact HR monitoring has gained popularity in applications such as DMS, with many studies using camera-based Computer Vision (CV) or radar-based methods [3].

CV approaches to HR measurement operate in a similar manner to PPG, in which changes in blood flow and blood oxygenation alter the optical absorption and reflection properties of the tissue. From Beer's law, the percentage of light reflected from the skin is inversely proportional to changes in blood volume [4]. This is due to different absorption rates of oxyhemoglobin and deoxyhemoglobin, with the percentage of total hemoglobin varying throughout the cardiac cycle [5]. Alternatively, radar systems monitor chest wall movement with modern mm-wave Doppler radars having the ability to detect mm or even sub-mm motion, enabling the measurement of the small movements of the chest wall caused by cardiac function. Although the effectiveness of both CV and radar non-contact HR estimation has been validated in controlled environments, how the specific characteristics of each modality are influenced by environmental factors such as lighting and movement remains relatively unexplored.

To address this knowledge gap, this paper investigates the characteristics and performance of CV and radar-based non-contact HR monitoring systems. The findings of this work could directly inform the development of DMS and be extended to related fields such as healthcare, aerospace, and navigation. Specifically, the major contributions of this work are as given below:

1. A non-contact HR measurement system for each modality is proposed and implemented using commercially available automotive-grade devices. Performance is validated using a reference ECG device, which is widely used in clinical settings and provides an accurate and reliable reference. The proposed hardware and algorithms for each modality demonstrate good accuracy with relatively low complexity, achieving relative errors of 0.3% and 0.15% for CV and radar-based approaches, respectively.
2. Through a comparative study, the accuracy and robustness of each modality are evaluated in different environments with varying distance and illumination, emphasising the need for a multi-modality approach when designing real-world DMS.

The remainder of this paper is organised as follows. A literature review in Section 2 summarises recent work in the field. Section 3 explains the reference non-contact architectures proposed for this study and Section 4 details the experimental setting and results discussion, followed by discussion and conclusions in Section 5.

2 Non-contact Heart Rate Monitoring

2.1 CV-based HR Monitoring

Since the concept of remote (non-contact) PPG (rPPG) was first demonstrated [6], it has been applied across a range of applications, most commonly using the green channel of an RGB video as this contains the strongest PPG signal due to the corresponding peak in the absorption spectrum of hemoglobin [6]. A conventional CV-based rPPG pipeline typically consists of the following main steps: Region of Interest (ROI) selection, signal extraction, signal processing, and HR estimation.

In the ROI selection phase, face detection algorithms such as HAAR [7], HOG [8], or deep neural networks are first used to achieve face detection, and then the ROI is selected for raw signal extraction. Several areas of the face, including the forehead and cheeks have been used as the ROI [8-11]. Using 68 facial landmarks was shown to enable the selection of ROI at the locations of large skin areas where the Signal-to-Noise Ratio (SNR) is higher [12]. However, any inaccuracies found in facial landmarks may also introduce noise. Alternatively, simply finding a rectangular bounding box of the face, or part of it, may introduce interference from hair, nose, mouth, and background but enhance the stability of ROI [7, 13].

The signal extraction phase can directly select the green channel [6] or develop a new color space from a linear combination of RGB channels, helping to address the limitations of RGB color space towards motion [14]. A study of the effectiveness of linear combinations of GB, GR, and GBGR found that GBGR achieved the best performance [11]. Instead of exploiting the color space, direct averaging can be applied to each channel and the exploitation task left to the subsequent signal processing stage [7, 13]. A recent work used the near-infrared light spectrum, where the raw signal was shown to significantly suppress the negative effect of ambient light variation at the cost of a low SNR [10]. In the signal processing phase, denoising and signal decomposition are the two most commonly used techniques One of the most significant denoising methods is the smoothness prior approach [15] which, due to its effectiveness and simplicity, remains in current use [16-17].

Advances in Artificial Neural Networks (ANN) have led to the development of neural-network-based denoising approaches in recent years. Examples include employing an inverse attention mechanism to increase the SNR when the signal of interest is weak [18] or utilizing Action Units to tackle the noise introduced by facial expressions [19]. However, the use of ANN raises concerns about the storage and computational power required for the real-time processing scenarios found in DMS.

The most frequently used signal decomposition technique is Blind Source Separation (BSS), including Independent Component Analysis (ICA) [7, 11, 13, 16] and Principal Component Analysis (PCA), used in [10] to extract the signal of interest. BSS assumes that all components are statistically independent and follow a non-Gaussian distribution. It also introduces two uncertainties: the order of and the magnitude of the recovered components. The order can be solved by power spectrum analysis or empirical methods, for example it has been argued that the signal of interest is usually the second component [11]. Finally, HR estimation is performed by extracting the frequency of the

highest power from the power spectrum or applying peak detection algorithms to the pre-processed signal. It is worth noting that the accuracy of the power spectrum analysis method is limited by the frequency resolution, determined by the time window length.

Like many other disciplines, deep learning provides a new avenue for research in this field. Qiu et al. proposed the CosTHR architecture, using color space transformation layers to learn the optimal color space and an attention convolutional neural network (CNN) to achieve estimation [20]. Other researchers combined Multi-scale Retinex (MSR) theory with the RGB space and fed the outcome into a CNN for HR estimation [9]. The MSR theory is inspired by the observation that the human vision system is robust to different illumination conditions, thus being used to counter the illumination variation issue of rPPG. Given the inaccuracy of handcrafted features, end-to-end networks were also explored to extract HR directly from the spatiotemporal maps constructed from videos [21-22].

2.2 Radar-based HR Monitoring

Since the recent widespread adoption of mm-wave radars in Advanced driver-assistance systems (ADAS), the feasibility of using radar in interior sensing has been studied, especially for HR monitoring. Automotive mm-wave radars operate in the 24 – 60 GHz range, thus providing the required Doppler resolution to detect the micro-motions caused by heartbeats. However, there are several sources of unwanted artefacts including movements from respiratory function (and harmonics of the same), motion artefacts, background reflection, multi-path interference and noise. To overcome these, typical radar-based HR measurement algorithms employ various filtering techniques and power spectrum analysis. One recent study first employed a 24 GHz Continuous-Wave (CW) Doppler Radar to extract the raw in-phase and quadrature (I/Q) signals and then used spectrum analysis to provide an approximate HR [23]. Finally, a set of more accurate estimations from several narrowband bandpass filters were compared with the previous rough estimation to determine the valid output.

The performance of radar-based sensing is highly dependent on power spectrum analysis – the highest peak may be the HR component but could also be the second respiration harmonic or other interference. To reduce noise, a signal elimination method can be used, for example subtracting the low- and high-frequency noise using two cascaded bandpass filters and feeding the output into a peak detection algorithm [24], though the performance is constrained by the accuracy of the lower and upper bounds of filters which are based on observations. The performances of peak spectrum super resolution techniques, such as the Multiple Signal Classification (MUSIC), and particle filter methods have been compared on both a driving simulator and on-vehicle testing, showing the complex and noisy nature of the driving scenario [25].

Recent radar-based HR monitoring has used a Frequency-modulated Continuous-wave (FMCW) radar to compensate for body motion by introducing the range-azimuth map [26]. This algorithm can estimate and remove large body motion artefacts and background reflection at the cost of more complex signal processing. However, the artefacts caused by micro-body motion such as random body movements and other irregular motions remain a problem because of the restricted range resolution.

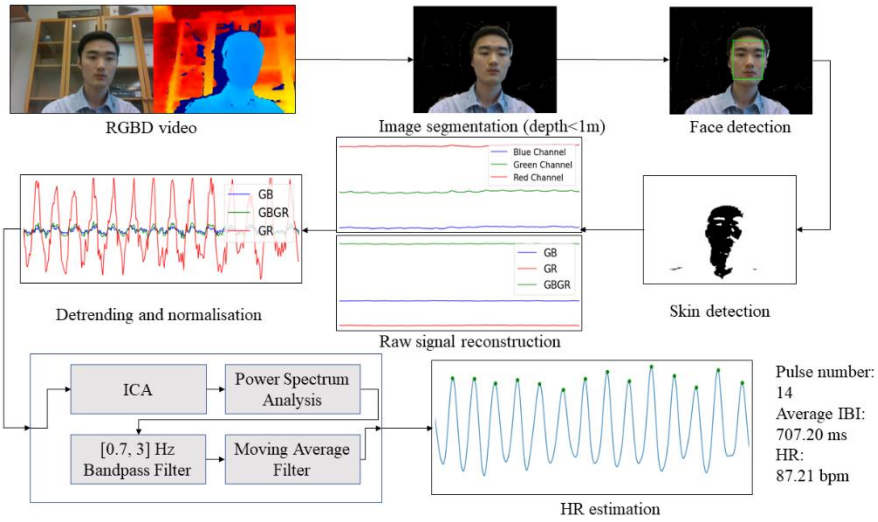

Fig. 1. Processing pipeline of the proposed CV-based HR monitoring architecture.

3 CV and Radar-Based DMS Testbench Architectures

To summarize the review in Section 2, CV-based approaches are mainly influenced by illumination variations and large body motions while radar-based approaches are robust to illumination change but prone to other interference. This observation provides the motivation for further investigation of the failure modes of each sensing modality and the extent to which each may degrade the system performance. To achieve this aim, the CV and radar-based non-contact HR monitoring architectures described below are used to simulate a DMS, enabling the characteristics and robustness of each modality to be investigated. A detailed description of the architectures and the underlying motivations are provided below.

3.1 Proposed CV-based HR Monitoring Architecture

The proposed CV-based HR monitoring algorithm is based on an RGBD camera [27], which is widely adopted by vehicle manufacturers for interior sensing. The processing pipeline shown in Fig. 1 aims to achieve a trade-off between performance and efficiency and consists of 5 main stages: (1) image segmentation; (2) ROI identification and tracking; (3) raw signal reconstruction; (4) signal processing; and (5) HR estimation from the interbeat intervals (IBI).

The system's input is an RGBD video sequence containing RGB channels and aligned depth information for each pixel and the time index for each frame. To remove the background and exclude passengers each frame is segmented using a simple depth range with a default threshold of 1m, based on the typical distance between the dashboard and the driver, see image in top center of Fig.1. As an alternative to the depth

threshold, image segmentation techniques such as the histogram of depth could be used to provide improved robustness with the cost of greatly increased processing.

Next, a combination of face detection and skin detection algorithms is applied to achieve ROI identification and tracking. The single-shot-detector (SSD) algorithm is significantly faster than most of the widely used deep-learning-based face detectors while outperforming most feature-based methods such as HAAR and HOG, and therefore to locate and track the facial region an SSD is implemented using the DeepFace library [28]. The initial ROI is defined as 80% of the height and width of the given bounding box of the face and, to maintain a stable bounding box, the face coordinates are updated only when the non-overlapping facial area from the previous and the current frames exceeds 3%. However, as the extracted ROI can include hair, nose, mouth, and other interferences in addition to the skin region, a skin detection algorithm based on a YCbCr color space transformation [29] is employed to refine the ROI and enhance the overall SNR. Following [29], The threshold values for the skin detection are:

$$85 \leq Cb \leq 135 \wedge 135 \leq Cr \leq 180, \tag{1}$$

giving a binary mask shown in center right of Fig. 1.

The raw signals are obtained by direct averaging of the values in each RGB channel within the ROI:

$$I\,(C)_{C\in\{R,G,B\}} = \frac{1}{n}\sum_{i=pixel\ value\ in\ ROI}^{n} i(C)_{C\in\{R,G,B\}}, \tag{2}$$

where $I\,(C)$ and $i(C)$ are raw signals and ROI pixel values in each channel, respectively. It has been observed that the ratio of these three signals can mitigate the fluctuations in the raw signals caused by light variation and movement, based on the assumption that all channels are equally influenced [11]. Hence, the raw signals used here are reconstructed by $GB = I(G)/I(B)$, $GR = I(G)/I(R)$ and $GBGR = GB + GR$.

In the signal processing stage, a smoothness prior detrending method is employed, which mimics a time-varying FIR high-pass filter [15]. The value of the regularization parameter λ is 50, and the corresponding bandpass frequency is 0.75 Hz when the sampling rate is 30 Hz. After a K-normalization, the three signals are fed into a fast ICA algorithm to recover the HR signal. A power spectrum analysis is performed to select the component within the 0.7 – 3 Hz frequency band which contains the highest peak. Finally, a 6th-order 0.7 – 3 Hz Butterworth bandpass filter and a five-point moving average filter are applied to suppress noise. It is worth noting that the 0.7 – 3 Hz frequency band corresponds to the normal HR range of 40 – 180 beats per minute (bpm).

The HR is then extracted using a prominence-distance-based peak detection algorithm to determine the peak position of each heartbeat, allowing comparison with the ground truth.

3.2 Proposed Radar-based HR Monitoring Architecture

The radar-based algorithm is developed using the BGT60LTR11AIP from Infineon [30], which is a 60GHz low-power automotive CW Doppler radar. The challenge for a radar-based system is that the amplitude of chest wall motion caused by heartbeats is

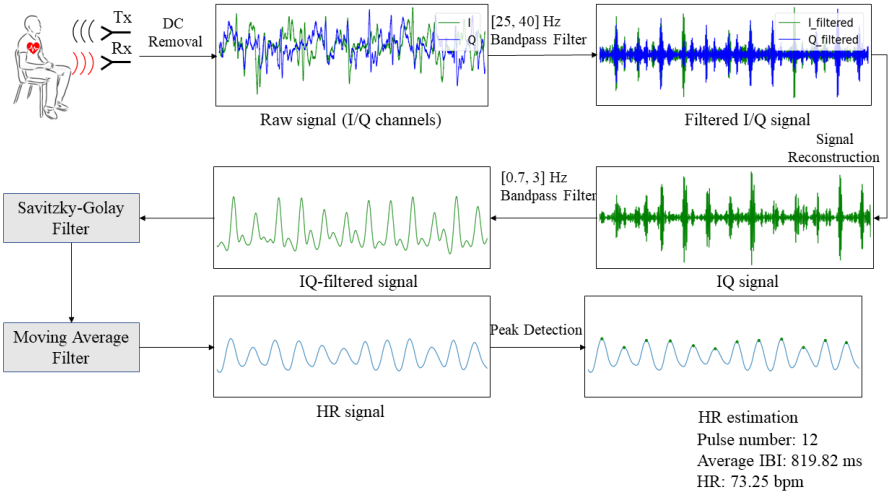

Fig. 2. System flow diagram of the proposed radar-based HR monitoring architecture.

small when compared to the overall motion of the driver. Moreover, the motion detected is the superposition of cardiac activity and respiration and therefore a set of filtering techniques must be employed to recover the signal of interest before any peaks can be recognized. The proposed processing architecture is shown in Fig. 2.

For a CW Doppler radar, the transmitter emits a frequency-modulated signal and the received signal is a motion-modulated signal. By mixing the transmitted and received signals, two baseband signals I and Q are generated, which are the raw signals used. After DC removal, the signals are fed into a 4th-order 25 – 40 Hz Butterworth bandpass filter to extract the periodic chest wall motion. The 25 – 40 Hz frequency band has been shown to contain the Doppler frequency range of the heartbeat motion in both synchronized radar signals and ECG ground truth. After complex signal demodulation,

$$IQ(t) = I(t) + jQ(t), \tag{3}$$

the magnitude of $IQ(t)$ is given by the absolute value of the complex number. Although the periodic signal can be visually recognized at this point, interference from other motion sources hinders the use of an automatic peak detection to extract the HR. Hence, a subsequent 2nd-order Butterworth bandpass filter is applied to further suppress interference. The IQ-filtered signal in Fig. 2 shows an interesting phenomenon in which the major peak is always followed by a smaller peak, which is most likely the dicrotic notch, reinforcing the previous observation that smaller regular motions can also be detected. However, these small spikes can sometimes interfere with peak detection, resulting in higher estimations, and to address this problem two further filters are applied. A 10th-order Savitzky-Golay smoothing filter with window size 1/3 of the sampling rate is used to merge the main peak and the dicrotic notch and a moving average filter with a window size 1/4 of the sampling rate is applied to reduce random noise. Finally, the HR estimation is achieved by a prominence-distance-based peak detection algorithm.

4 Experiment Design and Results Analysis

4.1 Performance Validation

A comparative study was designed under different testing environments to expose the performance and characteristics of each modality. Experiments were conducted in a laboratory setting using the CV and radar-based architectures described in Section 3 and two subjects with different characteristics (age, weight, and appearance). The RGBD video was captured using an Intel RealSense D435 depth camera [27] at 30 fps and resolution cropped to 640×480. The camera was mounted on a tripod 30 cm from the subject's face and, following a short initial transient period, 30 s of video was recorded. To obtain the ground truth HR, time-synchronized ECG data was collected using a MIKROE ECG 2 Click [31]. To collect the radar data, a BGT60LTR11AIP [30] sensor was mounted on a 3D-printed plastic mount located 30 cm away from the subject's chest and 30 s of data was recorded using a sampling frequency of 2000 Hz, with synchronized ECG again collected.

For each modality, the interbeat intervals (IBI) were found as the time differences between successive peaks and then averaged over the total number of IBIs n to estimate the HR using:

$$\mathrm{HR} = 60 \,/\, \left(\frac{1}{n}\sum_{i=1}^{n} IBI_n\right). \tag{4}$$

After investigating the variation of IBI distributions, and hence HR, for each modality, an experimental investigation is used to quantify the sensitivity of the different modalities to distance, illumination and motion.

4.2 Variation of IBI Distribution with Sensor Modality

Due to the different underlying principles of the ECG and non-contact HR monitoring methods, there are variations in the IBI distributions obtained from each sensor type. For example, the ECG measures electrical activity so the timing of the peak with the highest amplitude (the R-peak) does not exactly coincide with the resulting cardiac muscle contraction, and the peaks in radar signals are related to mechanical changes occurring sometime after the muscle contraction. For the CV-based rPPG method, the detected peaks lag the ECG due to the time taken for the blood volume to change. To illustrate the impact of sensor modality on IBI distribution and hence the corresponding HR estimation, Figs. 3 and 4 compare the IBI distributions from the non-contact methods with the ECG ground truth.

The Bland-Altman plots in Fig. 3 show the relationship between the paired non-contact HR estimations and ECG ground truth, with their 95% confidence intervals. It can be seen that no consistent bias exists for either non-contact method. The results from Figs. 3 and 4 show that both the IBIs and the corresponding HR estimations extracted from the CV-based rPPG method are distributed over a wider range than those extracted from the radar method. These results clearly illustrate that the peak of the blood volume change depends on many other factors, including the blood pressure, respiration, etc. However, the overall accuracy of the two methods is comparable, such

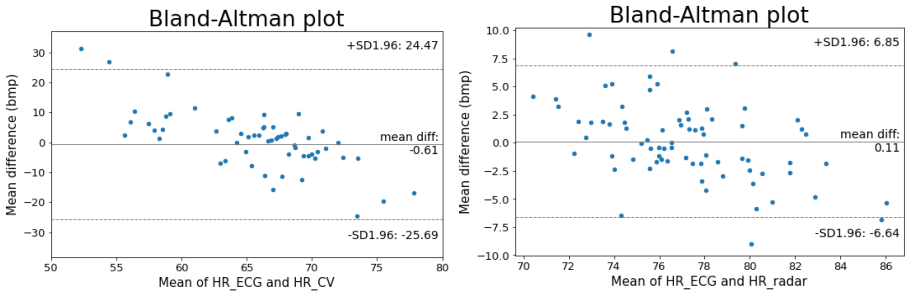

Fig. 3. Bland-Altman plots of the HR distributions of each method.

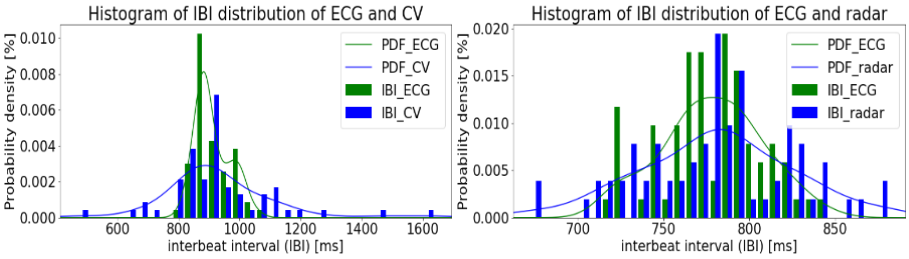

Fig. 4. Histograms of the IBI distributions of each method.

that when the IBI from the rPPG method is averaged over the 30 s period it has a minimal impact on the final estimation. Hence, the CV-based rPPG method is equally suitable for measurements over a long period while the radar method retains its accuracy over short-time measurements.

4.3 Impact of Distance on CV and Radar-Based HR Detection

Non-contact sensors are inevitably influenced by the physical distance between the sensor and the subject, often in the form of a lower SNR. Table 1 presents HR estimates for both modalities at distances of 30, 50 and 80 cm, showing that the radar-based method exhibits higher accuracy compared to the CV-based method at short and medium distances. However, the performance of radar diminishes with distance and is almost comparable to the CV-based approach at 80 cm, indicating the radar might be more sensitive to distance than the camera, possibly due to the power loss.

Although the relative error of both methods increases exponentially with distance, implying the existence of thresholds where the algorithms may fail completely, the distance between sensors and the driver in DMS is normally < 80 cm and so they operate in the distance range where the proposed architectures for both modalities are effective. The relative errors for this scenario are all below 0.7%.

4.4 Impact of Illumination on CV-based HR Detection

Illumination is widely recognized as one of the major barriers to CV-based HR monitoring. This contrasts with radar-based monitoring, which is invariant to illumination

Table 1. Performance of non-contact Heart Rate (HR) estimation modalities with distance.

Sensor type	Distance (cm)	Subject number	Ground truth (bpm)	HR estimation (bpm)	Absolute Error (AE)	Relative Error (RE)
CV-based	30	1	68.62	68.93	0.31	0.45%
		2	63.37	63.28	0.09	0.14%
				Mean	**0.20**	**0.30%**
	50	1	73.29	72.73	0.56	0.76%
		2	63.10	62.71	0.39	0.62%
				Mean	**0.95**	**0.69%**
	80	1	68.65	68.54	0.11	0.16%
		2	69.93	73.59	3.66	5.23%
				Mean	**1.89**	**2.70%**
Radar-based	30	1	78.96	78.95	0.01	0.01%
		2	75.41	75.20	0.21	0.28%
				Mean	**0.11**	**0.15%**
	50	1	72.26	72.06	0.20	0.28%
		2	72.05	71.49	0.56	0.78%
				Mean	**0.38**	**0.53%**
	80	1	72.54	69.84	2.70	3.72%
		2	88.32	86.90	1.42	1.61%
		Mean			**2.06**	**2.67%**

Table 2. Performance of CV-based Heart Rate (HR) detection under varying lighting conditions.

Illumination	Subject number	Ground Truth	HR estimation	Absolute Error (AE)	Relative Error (RE)
Low	1	65.91	65.34	0.57	0.86%
	2	64.46	66.00	1.54	2.39%
			Mean	**1.06**	**1.62%**
Flickering light	1	67.09	66.78	0.31	0.46%
	2	66.53	69.54	3.01	4.52%
			Mean	**1.66**	**2.49%**

changes. To investigate the impact of illumination on the CV-based method, the performance of the CV-based architecture was studied under two adverse lighting conditions: low ambient light, where only limited natural sunlight was available, and flickering light conditions where a flash lamp was alternatively on and off for 2 s at a time.

Table 2 summarizes the results, where the distance between sensors and subjects was 50 cm. The results show that, using its raw signal reconstruction technique, the proposed CV-based algorithm is relatively invariant to the adverse illumination conditions and still demonstrates good results. Overall, the low and flickering light conditions only increased the relative error by 0.93% and 1.8%, respectively.

4.5 Impact of Motion on CV and Radar-Based HR Detection

Motion is another significant factor that may degrade the performance of both the CV and radar methods, though in slightly different ways. Typical driving behavior involves

Table 3. Performance of each Heart Rate (HR) detection method under different motion types.

Sensor type	Subject number	Ground truth	HR estimation	Absolute Error (AE)	Relative Error (RE)
Camera (motion)	1	71.98	71.60	0.38	0.53%
	2	65.33	62.19	3.14	4.81%
			Mean	**1.76**	**2.67%**
Radar (motion)	1	77.07	80.24	3.17	4.11%
	2	80.43	71.07	9.36	11.64%
			Mean	**6.27**	**7.88%**

motion of the head and body and to fully understand the impact of motion experiments were conducted with different simulated motion types.

Previous studies have found the CV-based method to be affected by head motion and invariant to body motion, while the radar-based method demonstrates the opposite. Therefore, experiments were conducted using translational movements of the head for the CV-based approach and back-and-forth body vibrations for the radar-based approach. Subjects were asked to repeat the motion every 5 seconds and the distance to the sensors was kept at 50 cm.

The results presented in Table 3 show that the CV-based method is more robust to head motion than the radar-based method is to body movement, which is in line with expectations given the former measures color variation while the latter detects motion directly.

5 Discussion and Conclusions

In this paper, two non-contact HR monitoring architectures are proposed using CV and radar approaches and implemented using low-cost commercially available products suitable for automotive use. The CV-based system first employs depth information to achieve human background segmentation, followed by a combination of face detection and skin detection algorithms to locate and track the ROI. A raw signal reconstruction technique based on ratios of color channels is employed to reduce the fluctuations caused by illumination variations. A detrending method using smoothness prior is applied to enhance the SNR. Finally, the HR signal is recovered by ICA, bandpass filter, power spectrum analysis, and a moving average filter. Despite employing several techniques to mitigate the problems caused by illumination changes, this system still has a limited capability to handle extreme cases found in real-world deployment, such as drastic changes in brightness, contrast, and shadows. Addressing this robustness is an area of further work.

The radar-based system uses a bandpass filter to extract the chest wall motion from the motion-modulated RF signals, followed by a signal reconstruction process to merge the I and Q channels. Three further filters (bandpass, smoothing and moving average) are then combined to extract the HR signal. Again, real-world factors such as vehicle vibration and drivers' body motions would need to be considered for practical use.

Through the comparison with the ground truth obtained from a reference ECG device, the HR results for both methods were found to demonstrate a high accuracy. A

comparative study was performed to investigate the characteristics and failure modes of each method, showing the sensor modality has a significant impact on its performance in different environments and therefore should be carefully selected according to the application scenario.

In the context of DMS, the two modalities each demonstrates different strengths and weaknesses to the illumination changes, vibrations and motion found in real-world driving scenarios. Based on the hypotheses that both modalities will not fail simultaneously and that a temporary loss of data in extremely adverse conditions is acceptable, the development of a multi-modality sensor fusion system has significant potential for future non-contact HR monitoring research and may offer substantial benefits for continuous robust operation as part of a DMS. Further research into the development of a reference-free signal quality index and motion compensation techniques for both sensors are promising areas for future studies.

References

1. Human Error in Road Accidents, https://www.visualexpert.com/Resources/roadaccidents.html, last accessed 2023/04/18.
2. Jo, S., Kim, J., Kim, D.: Heart rate change while drowsy driving. Journal of Korean Medical Science 34(8), 56 (2019).
3. Halin, A., Verly, J. G., Van Droogenbroeck, M.: Survey and synthesis of state of the art in driver monitoring. Sensors 21(16), 5558 (2021).
4. Rovas, G., Bikia, V., Stergiopulos, N.: Quantification of the phenomena affecting reflective arterial photoplethysmography. Bioengineering 10(4), 460 (2023).
5. Van Kampen, E. J., Zijlstra, W. G.: Determination of hemoglobin and its derivatives. Advances in Clinical Chemistry 8, 141-187 (1966).
6. Verkruysse, W., Svaasand, L. O., Nelson, J. S.: Remote plethysmographic imaging using ambient light. Optics Express 16(26), 21434-21445 (2008).
7. Poh, M. Z., McDuff, D. J., Picard, R. W.: Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Optics Express 18(10), 10762-10774 (2010).
8. Muratov, Y., Nikiforov, M., Melnik, O., Loskutov, A.: Heart rate measurements with a web camera based on the facial image moving in the frame. In: 2021 10th Mediterranean Conference on Embedded Computing (MECO), pp. 1-5. IEEE, Budva, Montenegro (2021).
9. Jaiswal, K. B., Meenpal, T.: rPPG-FuseNet: Non-contact heart rate estimation from facial video via RGB/MSR signal fusion. Biomedical Signal Processing and Control, 78, 104002, (2022).
10. Magdalena Nowara, E., Marks, T. K., Mansour, H., Veeraraghavan, A.: SparsePPG: Towards driver monitoring using camera-based vital signs estimation in near infrared. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1353-135309. IEEE, Salt Lake City, USA (2018).
11. Haugg, F., Elgendi, M., Menon, C.: GRGB rPPG: An efficient low-complexity remote photoplethysmography-based algorithm for heart rate estimation. Bioengineering 10(2), 243 (2023).
12. Asthana, A., Zafeiriou, S., Cheng, S., Pantic, M.: Robust discriminative response map fitting with constrained local models. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3444-3451. IEEE, Portland, USA (2013).

13. Poh, M. Z., McDuff, D. J., Picard. R. W.: Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Transactions on Biomedical Engineering 58(1), 7-11 (2011).
14. De Haan, G., Jeanne, V.: Robust pulse rate from chrominance-based rPPG. IEEE Transactions on Biomedical Engineering 60(10), 2878-2886 (2013).
15. Tarvainen, M. P., Ranta-Aho, P. O., Karjalainen, P. A.: An advanced detrending method with application to HRV analysis. IEEE Transactions on Biomedical Engineering 49(2), 172-175 (2002).
16. Qiao, D., Zulkernine, F., Masroor, R., Rasool, R., Jaffar, N.: Measuring heart rate and heart rate variability with smartphone camera. In: 2021 22nd IEEE International Conference on Mobile Data Management (MDM), pp. 248-249. IEEE, Toronto, Canada (2021).
17. Krishnamoorthy, A., Bairy, G. M., Siddeshappa, N., Mayrose, H., Sampathila, N., Chadaga, K.: Channel intensity and edge-based estimation of heart rate via smartphone recordings. Computers 12(2), 43 (2023).
18. Nowara, E. M., McDuff, D., Veeraraghavan, A.: The benefit of distraction: Denoising camera-based physiological measurements using inverse attention. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 4935-4944. IEEE, Montreal, Canada (2021).
19. Lokendra, B., Puneet, G.: AND-rPPG: A novel denoising-rPPG network for improving remote heart rate estimation. Computers in Biology and Medicine 141, 105146 (2022).
20. Qiu, Z., Liu, J., Sun, H., Lin, L., Chen, Y.: CoSTHR: A heart rate estimating network with adaptive color space transformation. IEEE Transactions on Instrumentation and Measurement 71, 1-10 (2022).
21. Yu, Z., Li, X., Zhao, G.: Remote photoplethysmography signal measurement from facial videos using spatial-temporal networks. arXiv: 1905.02419 (2019).
22. Yu, Z., Peng, W., Li, X., Hong, X., Zhao, G.: Remote heart rate measurement from highly compressed facial videos: An end-to-end deep learning solution with video enhancement. In: 2019 Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 151-160. IEEE, Seoul, Korea (2019).
23. Petrović, V. L., Janković, M. M., Lupšić, A. V., Mihajlović, V. R., Popović-Božović, J. S.: High-accuracy real-time monitoring of heart rate variability using 24 GHz continuous-wave Doppler radar. IEEE Access 7, 74721-74733 (2019).
24. Oh, S. H., Lee, S., Kim, S. M., Jeong, J. H.: Development of a heart rate detection algorithm using a non-contact Doppler radar via signal elimination. Biomedical Signal Processing and Control 64, 102314 (2021).
25. Schwarz, C., Zainab, H., Dasgupta, S., Kahl, J.: Heartbeat measurement with millimeter wave radar in the driving environment. In: 2021 IEEE Radar Conference (RadarConf21), pp. 1-6. IEEE, Atlanta, USA (2021).
26. Wang, F., Zeng, X., Wu, C., Wang, B., Liu, K. J. R.: Driver vital signs monitoring using millimeter wave radio. IEEE Internet of Things Journal 9(13), 11283-11298 (2022).
27. Intel RealSense Depth Camera D435, https://www.intelrealsense.com/depth-camera-d435/, last accessed 2023/04/18.
28. Deepface, https://github.com/serengil/deepface, last accessed 2023/04/18.
29. Dahmani, D., Cheref, M., Larabi, S.: Zero-sum game theory model for segmenting skin regions. Image and Vision Computing 99, 103925 (2020).
30. BGT60LTR11AIP, https://www.infineon.com/cms/en/product/sensor/radar-sensors/radar-sensors-for-iot/60ghz-radar/bgt60ltr11aip/, last accessed 2023/04/18.
31. ECG 2 click - board with ADS1194 AD converter from Texas Instruments, https://www.mikroe.com/ecg-2-click, last accessed 2023/04/18.