Remote real-time heart rate monitoring with recursive motion artifact removal using PPG signals from a smartphone camera

Asmaa Hosni[1,2] · Mohamed Atef[2,3]

Abstract

Remote photoplethysmography (rPPG) recorded by low-cost smartphone cameras is a promising method for noncontact monitoring of heart rate (HR). The main challenges of this method are the limited rPPG signal strength and motion artifacts (MAs). To obtain a clean signal while preserving photoplethysmography features, this study proposes algorithms to eliminate the noise and distortions of the extracted rPPG signal. To obtain the best point for the rPPG signal extraction, the highest green channel difference in two consecutive frames is calculated. Mexican hat wavelet transform decomposition is used for MA distortion elimination. Furthermore, we propose a recursive baseline-wander removal algorithm with an adaptive window to effectively remove baseline drift. The peak detection accuracy is enhanced by using an adaptive sliding-window size based on the detected fast Fourier transform beat frequency. Our proposed algorithm was validated using a range of HR values from seven subjects. The results of the algorithm validation showed a real-time operation and accuracy improvements of more than 37.5% for peak detection in the worst case. The proposed method can measure HR remotely from a 0.4 m distance without any additional sensors, achieving a mean absolute error ± standard deviation of 3.58 ± 2.4.

Keywords Remote Photoplethysmography · Heart rate · Wavelet transforms · Motion artifacts · Baseline wander

✉ Mohamed Atef
    moh_atef@uaeu.ac.ae

    Asmaa Hosni
    asmaaf@uaeu.ac.ae

1   College of Information Technology, United Arab Emirates University, 15551 Al Ain, Abu Dhabi, UAE

2   Assiut University, Assiut 71515, Egypt

3   Electrical Engineering Department, United Arab Emirates University, 15551 Al Ain, Abu Dhabi, UAE

1 Introduction

Heart rate (HR) is a major physiological indicator of considerable significance for personal health monitoring, which can be measured using photoplethysmography (PPG) signals. A PPG sensor consists of a light transmitter for sending light that passes through the tissue. The absorption of the light depends on the blood volume changes in the circulatory system. Then, a light detector is used to receive and sense the reflected/transmitted light. The received light contains information related to changes in blood volume and HR [23].

PPG signals can be sensed from the wrist, fingertip, forehead, or earlobe. Smartwatches with optical sensors can be used to extract PPG signals from the wrist. Moreover, smartphone cameras can measure PPG signals from the fingertip, wrist, and forehead. However, contact between the PPG sensor/camera and the skin is required to obtain a high-quality PPG signal. While the contact-based PPG method can obtain high-quality signals, it is difficult to use as a universal sensor for different parts of the body [10, 23].

In recent years, many researchers have used smartphone cameras as sensors for monitoring HR, oxygen saturation level, blood glucose, and blood pressure using recorded PPG videos [11, 28, 22, 27]. Moreover, robots can collect information during interactions with humans and estimate their HR with a tolerance of 5 bpm [1]. The PPG technique extracts the PPG signal from the skin color changes in a recorded video and then calculates the HR according to the number of peaks detected in a certain period. A smartphone camera can extract PPG signals from any part of the human body in a noncontact manner [15, 17, 6, 24] (Fig. 1). The availability of smartphones makes them a convenient computing platform for medical diagnoses.

Different factors negatively affect the quality of the PPG signals. Motion artifacts (MAs) are one such factor; MAs induce unwanted in-band variations in the acquired PPG signals. Furthermore, PPG signals extracted from smartphone cameras are exposed to MAs more than PPG signals acquired from contact PPG sensors. MAs occur because of the subject's breathing or movement, causing misplacement of the measurement position during video recording. The MA levels are greater in amplitude than PPG signals and are considered a major source of distortion for the PPG signal, which creates a challenge for accurate PPG measurements [21].

Several MA-removal techniques have recently been proposed for PPG signals. Hardware-based MA detection and removal approaches use extra sensors, such as accelerometers, gyroscopes, or infrared light-emitting diodes to generate reference signals to remove MAs [26, 18, 13]. However, hardware-based MA reduction methods require additional sensors to be connected to smartphones, which is inconvenient for wearable systems and increases their cost and design complexity.

By contrast, many software methods have been developed to detect and remove MAs, such as filtering techniques, cubic spline interpolation, and wavelet transform [25, 12, 21]. Algorithm-based methods for MA reduction, such as time-domain processing, frequency-domain techniques, combining time and frequency techniques, and blind source separation methods, have been proposed in [5]. Low-frequency MAs can be removed by using a high-pass filter. Because the MA associated with PPG signals is time-varying, time-frequency approaches, such as wavelet transforms, can be effective in removing MA components from PPG signals [29, 20]. However, the method proposed in [29] requires a wrist-worn accelerometer. In addition, the techniques in [29, 20] were applied to PPG signals acquired directly with a finger-connected pulse oximeter and not to remote PPG (rPPG) signals.
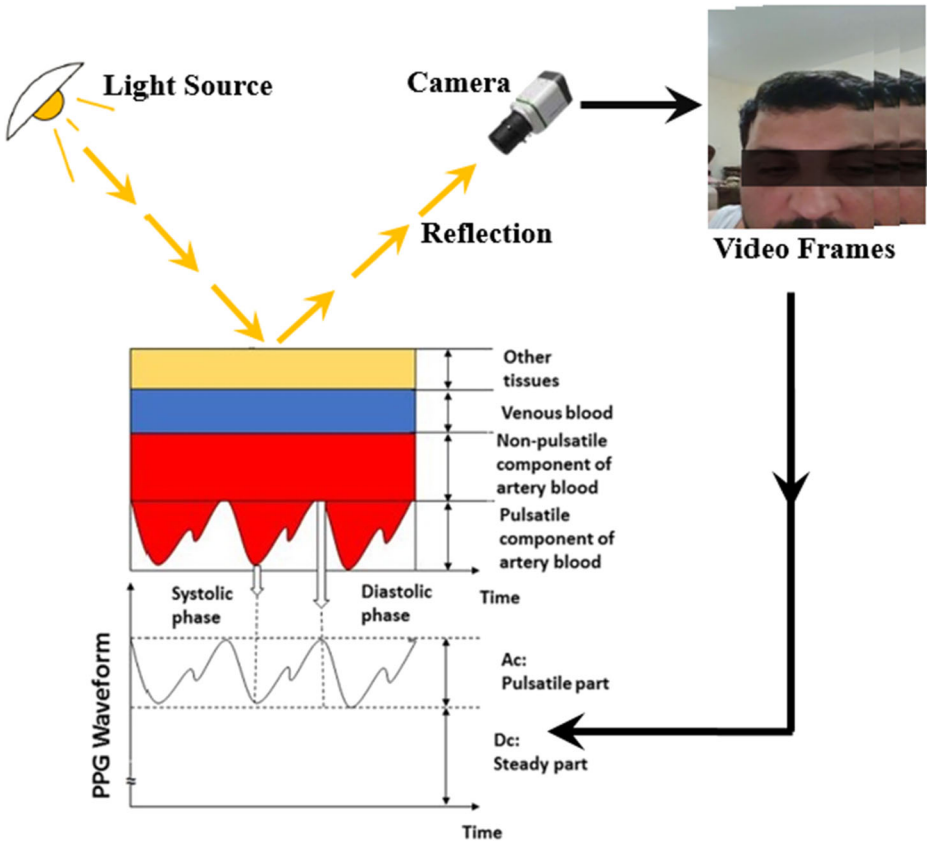
Fig. 1 Smartphone camera used to sense PPG signals in a contactless manner

In this paper, we present an MA-removal method for contactless smartphone PPG signals. It utilizes the best point selection and adaptive multistage baseline-wander removal algorithms to enhance the quality of weak noisy PPG signals. The proposed algorithm uses wavelet filtering for MA cancelation and adapts four parameters for peak detection, segmentation, and normalization. These parameters are the average signal value, peak-to-peak time interval, successive peak values, and PPG segment time. These parameters are used to enhance the PPG signal quality. Therefore, the PPG signal features and shapes remain the same with minimum interference. The experimental results indicate that the proposed algorithm is effective in extracting extremely weak contactless rPPG signals that are more susceptible to MA interference.

Our main contributions include the following:

1) An algorithm to select the best point for rPPG signal extraction.
2) A multistage recursive baseline-wander removal (RBLWR) algorithm to effectively remove baseline drift.
3) Adaptive peak detection and segmentation based on fast Fourier transform (FFT).

The remainder of this article is organized as follows: Section 2 presents an overview of the proposed system and method. Section 3 describes the main steps and best point selection. Section 4 explains the proposed motion artifact and baseline-wander removal. The HR

calculation is described in Section 5. The results and discussion are presented in Section 6. Finally, Section 7 concludes the paper.

2 Proposed method and measurement setup

The videos were recorded and processed according to the following steps, and a flowchart of the entire process is shown in Fig. 2. In the first step, a 30 s video of the face was recorded at 60 fps using a Samsung S20 camera. Videos were recorded for seven subjects (three men, two women, and two children). The video was then transferred to a laptop for further processing and HR calculations using Python. The entire face and forehead were detected from a distance of 0.4 m. Then, the best point was selected and checked to determine if the chosen pixel has a maximum green color value change in two consecutive frames. Subsequently, the PPG signal was extracted from the video frames.

The PPG perfusion ratio (AC/DC) enables the determination of the relationship between the color channels used. Based on the perfusion ratio values, it can be seen that the green is the
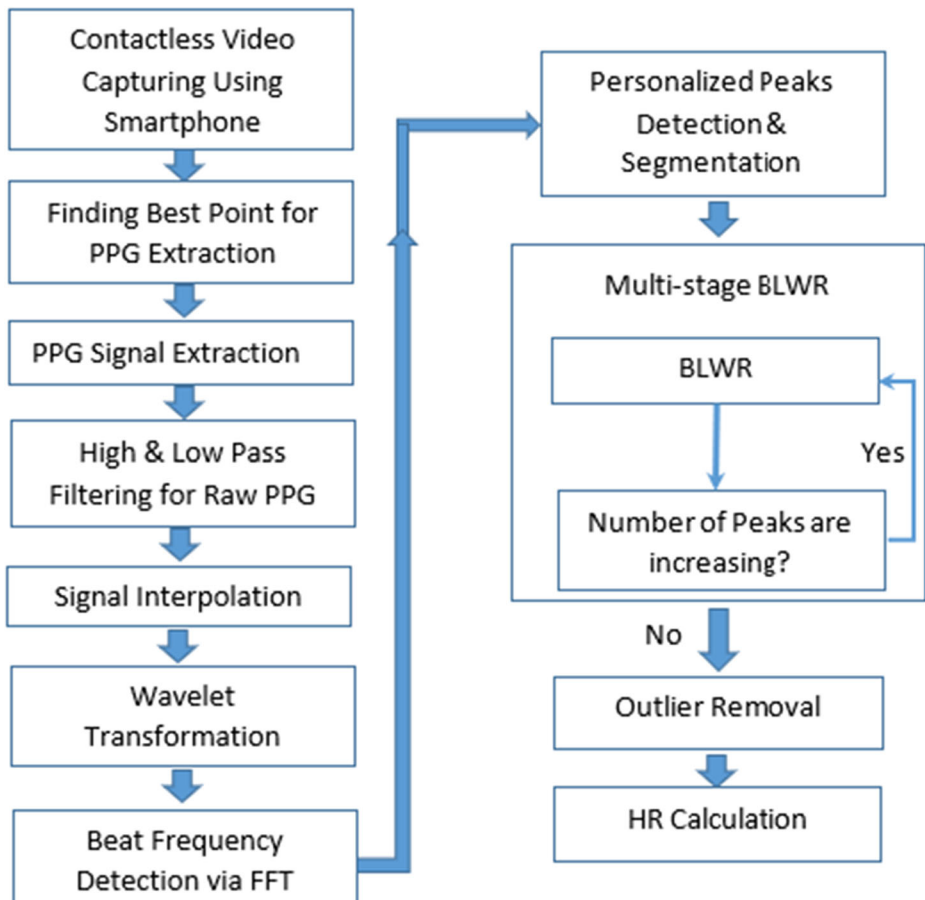

Fig. 2 Flowchart of the proposed HR monitoring based on rPPG

most pulsatile channel [4]. The PPG signal obtained from the green channel is the strongest signal that reflects and preserves the blood volume variation during cardiac cycles. In addition, the quality of the video recording is affected by face position, motion, and surrounding light intensity. Therefore, we evaluated the recorded video by monitoring the strength of the green channel frame by frame. The strength of each color was evaluated to provide a suitable threshold for accepting or rejecting the acquired video. A total of 1800 video frames from the first 5 s of the seven subjects were used to calculate the appropriate threshold level.

A 5 Hz low-pass filter (LPF) to remove high-frequency noise and 0.4 Hz high-pass filter (HPF) to remove slow-motion artifacts and breathing effects were then applied. Subsequently, the filtered PPG signal was interpolated using a cubic spline to improve the accuracy of peak detection as the sampling rate was increased from 60 to 240 samples/s. An accurate HR variability analysis requires high sampling rates.

Then, a continuous wavelet transform (CWT) was applied to filter the slow components related to MAs. Subsequently, an FFT was used to extract the beat frequency. Then, initial peak detection was applied to find peaks and valleys and to create PPG segments based on the FFT-extracted beat frequency. The proposed RBLWR method based on segment normalization in conjunction with recursive peak detection was applied. Before computing the HR, outliers were removal by removing segments with irregular pulse-to-pulse interval (PPI), pulse width, and amplitude features.

Finally, the HR was calculated from the clean PPG signal. Simultaneously with recording of the video to generate the PPG signal, reference HR was recorded using the FDA-approved BEURER PO30 Pulse Oximeter [3]. This was performed to validate the calculated HR results of the proposed algorithm.

3 Preprocessing and best point selection

The PPG signal extraction and preprocessing steps were as follows:

1. Skin detection
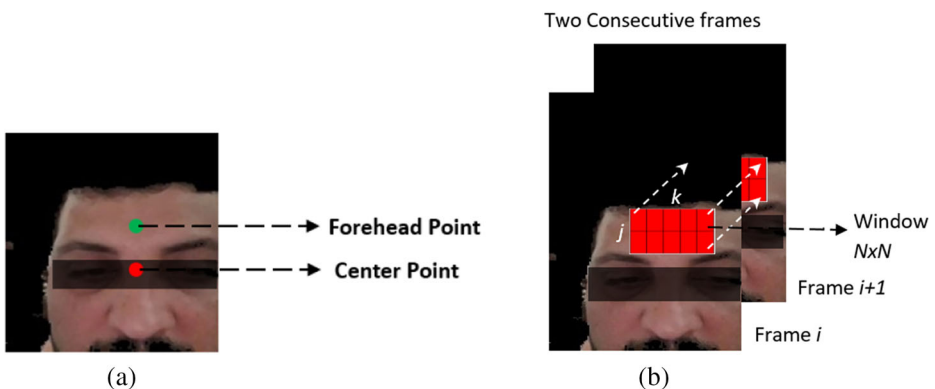2. Best point selection around the center of gravity

Fig. 3 Skin detection and the center-of-gravity pixel detection. Best pixel calculations based on maximum color change in two consecutive frames

3. Spatial smoothing around the selected best point
4. *Spatiotemporal* extracting of the green channel
5. Out-of-band noise removal using LPF and HPF

The recorded video was converted into frames, and a skin detection algorithm based on skin tone was then applied to one frame. The skin detection algorithm was adapted based on a threshold that recognizes skin regions using the RGB-HSV-YCbCr model [19]. The skin detection algorithm returned the skin portion of the image (primarily the face), and the center of gravity pixel was automatically detected from the forehead, as shown in Fig. 3.

The center of gravity of the connected region of the detected face is computed as follows:

$$p_c(x,y) = \left(\frac{1}{\varnothing} \sum_{x,y \in \varnothing} x, \frac{1}{\varnothing} \sum_{x,y \in \varnothing} y \right), \tag{1}$$

where x and y are the horizontal and vertical pixel indices, respectively, and $\varnothing$ is the total number of pixels in the detected region. Because the forehead is found in the upper one-fourth of the face, the forehead pixel p_f is computed as

$$p_f(x,y) = p_c \left(x, \ y - \frac{\varnothing}{4} \right) \tag{2}$$

To generate a strong PPG signal, we checked whether the chosen pixel had a maximum green color value change in two consecutive frames. Within a window W of neighborhood pixels, let us assume that p_f in Eq. (2) is computed at the ith frame, as shown in Fig. 4. The best point p_{best} for generating the PPG signal is then computed as

$$p_{best}(x,y) = argma \left[x_{p_f \in W} \left(p_{f_i} - p_{f_{i+1}} \right) \right] \tag{3}$$
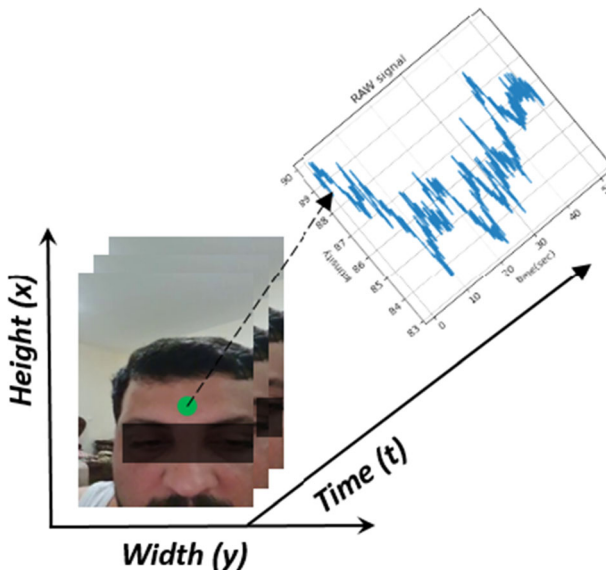

Fig. 4 Extracting PPG signal from the video frames

The absorption of light by the blood depends on the change in the forehead blood volume. Changes in light intensity due to absorption were recorded in the video. Consequently, the pixel intensity p_f in a fixed location had different values for each frame i. The PPG signal was extracted from the temporal changes in the pixel intensity in the adjacent frames. The green channel from each frame was selected and processed. The acquired video was then converted into an image sequence (frames). A 30 s video was recorded for each subject. The frame rate was 60 fps; thus, a video of 30 s contained M = 30 × 60 = 1800 images.

Each frame in the image sequence was transformed into data points by averaging the selected pixels according to a certain window size. For any i_{th} frame, assuming neighboring pixels in the window of size N × N and the pixel color value of $P_i(j, k)$, the pixel values of all points were superimposed and averaged.

The average pixel value P_i of the ith frame was obtained using the following formula:

$$P_i = \frac{\sum_{j=1}^{N} \sum_{k=1}^{N} P_i(j,k)}{N^2} \quad , \quad i = 1, 2, ..M \tag{4}$$

Using this method, the 2D image for the ith frame was converted to data points; then, M frames of the image were converted to M data points, namely, $P_1, P_2, P_3, .. P_M$. Finally, the PPG signal was obtained by sorting the data points across time. The time of sample I is t_i, and can be calculated using the following equation:

$$t_i = \frac{i}{f_s} \quad , i = 1, 2, ..M \tag{5}$$

where f_s is the sampling rate, which is equal to the number of frames per second multiplied by the interpolation factor, that is 60 × 4 = 240 samples/s.

Interpolation was used to increase the number of data points of the PPG signal so that the interpolated PPG signal would have a higher sampling rate. The extracted rPPG samples from the video frames versus the sample time are plotted in Fig. 4.

4 Motion artifact and baseline-wander removal

4.1 Wavelet transformation

PPG extracted from video frames may experience various sources of motion interference and baseline drift. Therefore, further filtering of the pulse wave is necessary. Removing low-frequency components using the HPF to eliminate the baseline wander results in the loss of some useful features in the PPG signal. The wavelet filter was not bound to a fixed cutoff frequency. Hence, it could approximate where the drift frequencies that vary with time increase or decrease below the cutoff frequency threshold. The application of wavelet filters is considered one of the most effective methods for the MA filtering of biological signals [20].

The Mexican hat wavelet was selected as the mother wavelet because it is similar in shape to the PPG signal. By trying different wavelet functions, the Mexican hat mother wavelet with scale 8 yielded the highest signal quality after the reconstruction. We selected a scale of 8, and the signal was stretched by a factor of 8, making it more sensitive to lower frequencies in the signal. Increasing the scale value resulted in more sub-bands, and consequently, in more degrees of freedom for noise suppression. Increasing the scale value requires more computational complexity,

which results in longer processing time. The first two levels have low-frequency content and artifacts related to motion. Next, reconstruction was performed by summing the scaled weighted sub-signals over the selected scales.

4.2 Peak detection and normalization

A sliding-window algorithm was adopted to find the PPG signal peaks and valleys. However, owing to baseline drift, the signal peaks and valleys could not be detected. If the peaks and valleys were incorrectly detected, the subsequent peaks and valleys had a high probability of being detected incorrectly. Therefore, peaks and valleys must be detected accurately. Consequently, we proposed a new algorithm to remove the baseline wander, after which the PPG signal is returned to its normal time base and segmented into single pulsed periods.

The peaks and valleys of the PPG signal were detected correctly using an adaptive sliding window with a recursive mechanism. The pseudocode for the RBLWR is shown below, and it lists the entire process of the algorithm. For periodic segmentation, correct localization of all peaks and valleys is required. After conducting several experiments, we empirically selected a window size of five periods, which achieved the best performance in terms of accuracy and real-time operation. The period length ($1/f_o$) was obtained by extracting the beat frequency(f_o) using the FFT for the PPG signal. One advantage of the RBLWR algorithm is the correction of incorrectly detected peaks and valleys using outlier removal. It is not expected that the HR will exceed 200 bpm, even under strenuous exercise. Therefore, the minimum time spacing between two consecutive peaks was set to be greater than $0.4 \times 1/f_o$ to remove the false peaks.

Outlier removal was implemented by removing segments with irregular features. As the HR limits are 50–200 bpm, segments with width of more than 1.2 s or less than 0.3 s were removed as outliers.

Algorithm Recursive Baseline-Wander Removal (RBLWR)

Input: Reconstructed Wavelet
Output: PPG segments after baseline-wander removal.
1. FFT to get the beat frequency f_o
2. Set a window size = $5 \times 1/f_o$
3. Set a minimum spacing between two consecutives maximums to be greater than $0.3 \times 1/f_o$.
4. Slide the window to find all maximums over the entire PPG signal length.
5. Find each minimum between two maximums obtained in step 4.
6. Define segment n between two minimums and one maximum.
7. Calculate the new normalized sample values inside segment n based on Equation 6.
8. Repeat steps 4 to 7 until obtaining a total fixed number of maximums and minimums.
9. Remove outlier segments with width of more than 1.2 s or less than 0.3 s or peaks lower than the window average.

The new normalized sample values inside segment n were calculated based on the following equation [7]:

$$PPG_n = \frac{PPG_0 - \min(PPG_0)}{\max(PPG_0) - \min(PPG_0)} \tag{6}$$

PPG_n indicates the normalized sequence of a single beat, whereas PPG_0 represents the original sampled real-amplitude sequence. Two functions, *min* and *max,* returned the minimum and maximum values of the sequence, respectively.

After noise cancellation and signal normalization, the outlier segments caused by occasional incorrect measurements were removed.

The outlier removal algorithm has the following steps:

a) For all n PPG segments, we measure PPI and pulse interval.
b) If the corresponding PPI and pulse interval values for a certain segment are larger or smaller than a certain threshold, the segment is removed from the sequence as an outlier.

5 Heart rate calculation

To extract the HR from the rPPG signal, we first calculated the PPI of PPG. PPG peaks were detected using a traditional peak detection algorithm [2] and used to calculate the PPI time series (Fig. 5).

The mean instantaneous pulse rate within the one-minute time window was calculated as

$$HR = \frac{60}{\overline{PPI}} \tag{7}$$

where $\overline{PPI}$ is the mean of the PPIs.

To evaluate the performance of our system, we compared the reference HR values from BEURER PO30 with the HR values obtained using the proposed method.

6 Results and discussion

The effects of the extracted raw PPG signal from video frames using the forehead skin center of gravity and selected best point around the center of gravity are compared in Table 1. Videos of seven subjects (three men, two women, and two children) were recorded and then
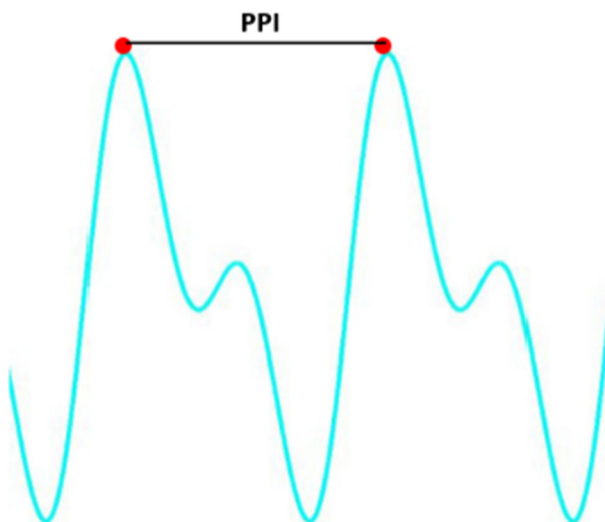

Fig. 5 Extracting PPI from PPG signal

Table 1 Measured HR using the proposed algorithm compared with the reference HR

	Reference HR	HR without best point	HR Error	HR with best point	HR Error
Subject #1	110	107	2.7%	107	2.7%
Subject #2	109	93	14.6%	106	2.8%
Subject #3	83	86	3.6%	85	2.5%
Subject #4	85	91	7%	91	7%
Subject #5	92	91	1%	93	1%
Subject #6	95	97	2.1%	97	2.1%
Subject #7	93	100	7%	100	7%
MAE±SD		5.43±4.66		3.58±2.4	

transferred to a laptop for further processing and HR calculation using Python. The entire face and forehead were detected from a distance of 0.4 m. The results in Table 1 indicate that the HR measurement error improved from 14.6% to 2.8% for Subject #2 and from 3.6% to 2.5% for Subject #3 when the proposed best point selection method was used. The highest HR error measured using the proposed algorithm for rPPG signal was 7%. This improvement can be attributed to the best point method having a stronger signal intensity.

The measured HR Error introduced in Eq. (8) is considered a metric for performance improvements:

$$HR\ Error\% = \left| \frac{Refrence\ HR - Meaured\ HR}{Refrence\ HR} \right| x100\% \tag{8}$$

where the reference HR is measured using BEURER PO30 Pulse Oximeter.

The mean absolute error (MAE) and standard deviation (SD) were calculated using Eqs. (9) and (10), respectively.

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \frac{|y_i - \widehat{y}_i|}{y_i} \tag{9}$$

$$SD = \sqrt{\frac{\sum_{i=1}^{N} (y_i - MAE)^2}{N}} \tag{10}$$

Here, N is the number of measurements, and y_i is the measured HR in each subject.

Furthermore, the proposed algorithm was also applied to the MPSC-rPPG dataset [9]. The MPSC-rPPG dataset contains HD RGB videos recorded at 30 fps for 5 min of a volunteer sitting still on a chair 3 ft. away from the camera. Concurrently, a wristwatch tracked the wrist PPG sampled at 64 Hz from the volunteer's dominant hand. The dataset consists of videos of seven volunteers: one woman and six men. All subjects were healthy and had no known medical conditions, with HR ranging from 50 bpm to 95 bpm [9].

Table 2 shows the HR measured using the proposed algorithm for the MPSC-rPPG dataset. The results shown in Table 2 indicate that the maximum HR measurement error was 4.36 in Subject #2. The proposed algorithm shows an HR MAE ± SD of 2.23 ± 1.58 for the MPSC-rPPG dataset.

Table 2 Measured HR using the proposed algorithm with the MPSC-rPPG dataset

|	Reference HR (bpm)	HR using proposed Alg. (bpm)	HR Error
Subject #1	96.44	96.14	0.3
Subject #2	82.2	86.56	4.36
Subject #3	87.62	88.8	1.18
Subject #4	81.83	80.61	1.22
Subject #5	91.6	88.1	3.5
Subject #6	93.2	89.46	3.74
Subject #7	82.52	83.84	1.32
MAE±SD		2.23±1.58	

The detected PPG signals are susceptible to external noise sources, such as ambient light noise, and 60–50 Hz powerline interference and its harmonics. Hence, the extracted PPG signal was denoised by an LPF to reject frequency components higher than 5 Hz and by an HPF to remove frequencies lower than 0.4 Hz, as shown in Fig. 6. The high-frequency noise was removed by the LPF, and part of the baseline wander was removed using the HPF.

By using wavelet decomposition and reconstruction, the MAs were eliminated. Comparing Fig. 7a for the filtered signal and Fig. 7b for the wavelet-constructed signal, it can be seen that the amplitude affecting the low-frequency noise is well corrected using the Mexican hat wavelet with a scale of 8.

To find the peaks, a search window related to the PPG period must be set. The period length was obtained by analyzing the frequency of the periodic signal using FFT and selecting the beat frequency, as shown in Fig. 8.

The normalized PPG signal after the first normalization using Eq. (6) is shown in Fig. 9a. The normalization was repeated until all peaks and valleys were found, as shown in Fig. 9b. The final PPG signal after outlier removal is shown in Fig. 10b. Of the total samples, 2% were removed from the original signal compared with Fig. 10a. A zoomed-in PPG segment, which demonstrates that the features are preserved after wavelet reconstruction and recursive normalization, is shown in Fig. 11.

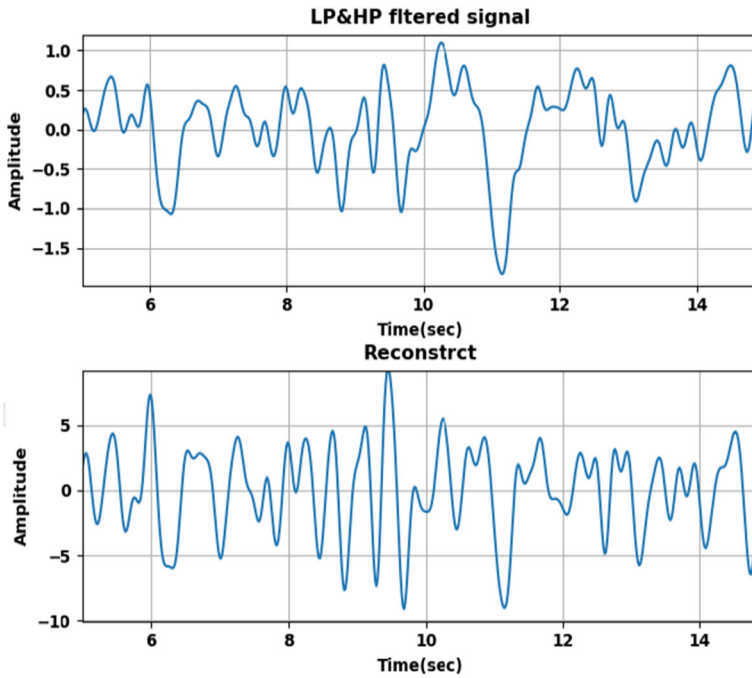**Fig. 6** Filtered PPG using 5 Hz LPF and 0.4 Hz HPF
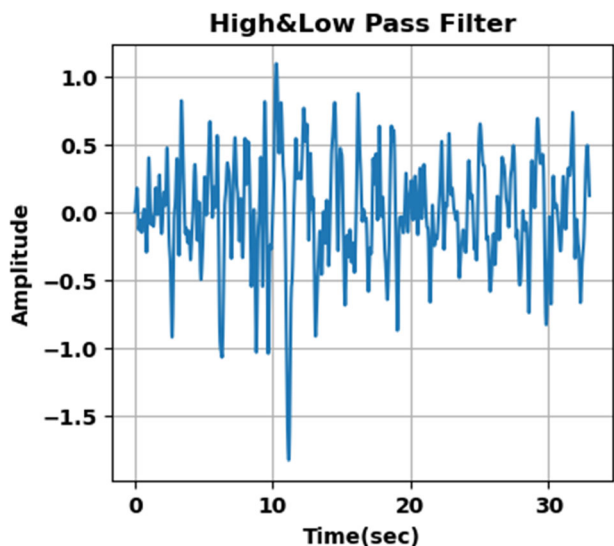
Fig. 7 a The input signal after LPF and HPF filtering, and **b** the reconstructed signal after wavelet transformation

To measure the accuracy of the proposed method, the detection error was defined by Eq. (11) [2]:

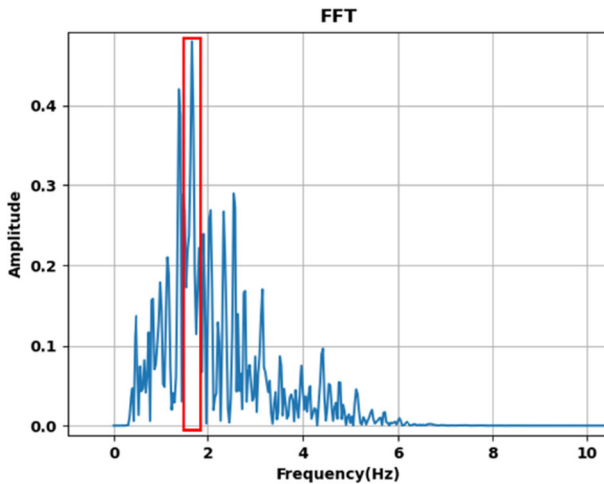$$Error\% = \frac{FP + FN}{Total} \; x \; 100 \; \% \tag{11}$$
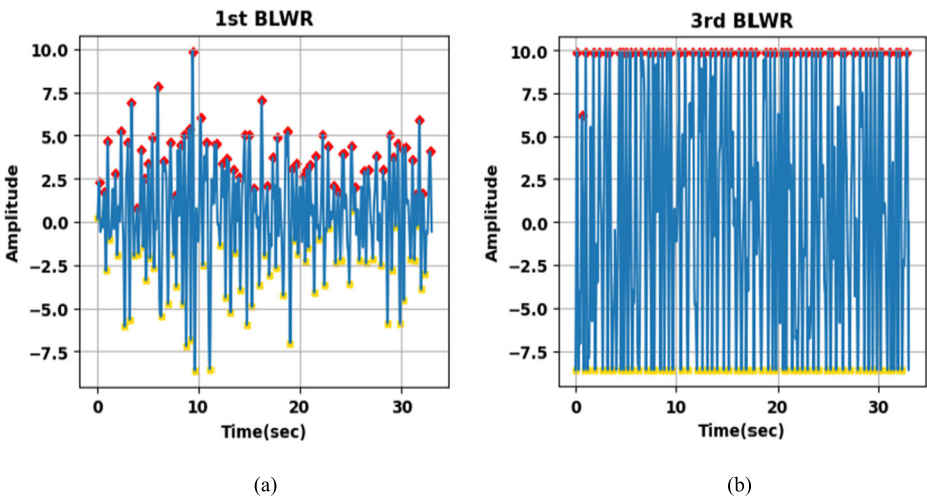

Fig. 8 FFT for filter PPG

Fig. 9 PPG signal after recursive normalization (a) first trial and (b) after finding all peaks

where *FN* and *FP* denote the number of false negatives (undetected peaks) and false positives (erroneously detected peaks), respectively. The sum of *FN* and *FP* was counted as the total detection error. The denominator is the total number of peaks labeled visually.

The measured HR values and their related error percentages using the proposed algorithm compared with the reference HR are listed in Table 3. It also summarizes the accuracy improvements of using Mexican hat for peak detection. Before using the wavelet transformation, the detection error ranged from 47% to 74%. After the wavelet reconstruction of the signal, the highest detection error was 9.5%, a 37.5% accuracy improvement. These improvements were due to the removal of MAs from the reconstructed signal.

One of the main advantages of the proposed system is its low computational time. The algorithm has low computational complexity because the main computations are performed for
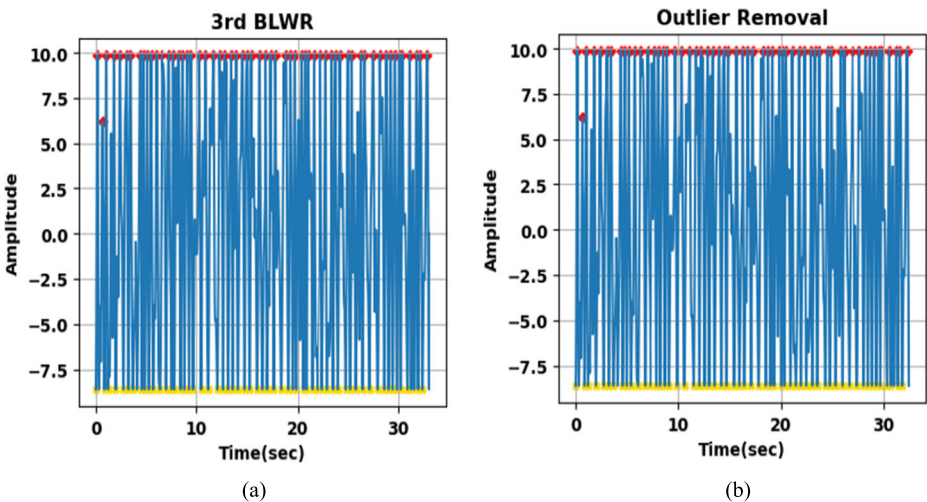

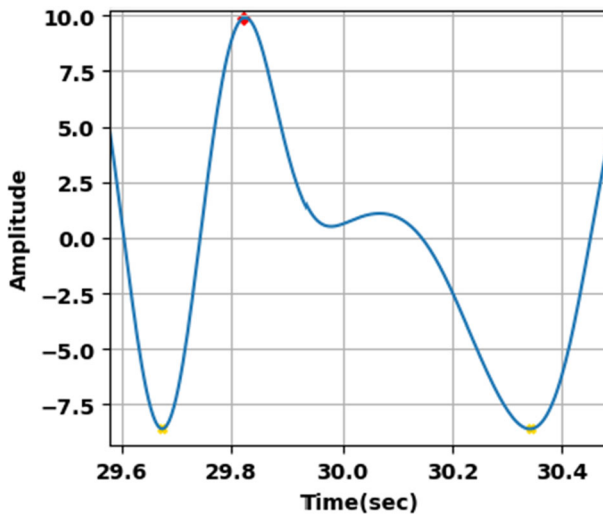Fig. 10 **a** PPG after BLWR and **b** PPG after outlier removal

Fig. 11 One PPG segment after wavelet reconstruction and recursive normalization

the 1D rPPG signal, not 2D, as in other methods. Thus, our system can be used in real-time applications, such as wearable devices. Table 4 shows the time consumed to generate the rPPG signal and the processing time for our algorithm for two different video lengths. In general, the total time complexity of our algorithm is low, proving that it can be used in real-time applications. The most time is consumed by generating the rPPG signal. In future work, we will focus on enhancing this step using parallel processing.

Table 5 shows the comparison of rPPG state-of-the-art algorithms in terms of MAE and SDs. The work presented in [14] used spatiotemporal homomorphic filtering (ASTHF) and achieved a MAE ± SD of 4.81 ± 3.39 with the camera at 0.5 m distance from the target. Nevertheless, a major drawback of ASTHF is that is not robust to MAs [14].

Principal component analysis and discrete cosine transform were used in [8] with a 0.5 m distance to achieve a comparable error of 4.69 ± 3.43. However, this method is not applicable for real-time HR measurements because of its dependency on the temporal stability of the facial point trajectory [8].

The HR was measured from a larger distance of 1 m in [24] using Soft Signature (SoftSig), and from 1.5 m in [16] using a deep recursive convolutional network. However, the methods introduced in [16, 24] required an additional near-infrared sensor or multi-imager system,

Table 3 Detected peaks after LPF and HPF filtering and after wavelet reconstruction

	Total Reference Peaks	Peaks After LPF&HPF	Error %	Peaks After Wavelet	Error %
Subject#1	60	16	73%	62	3.3%
Subject #2	58	15	74%	60	3.4%
Subject #3	21	11	47%	23	9.5%
Subject #4	69	25	63%	71	2.8%
Subject #5	94	37	60%	94	0%
Subject#6	108	43	60%	107	0.1%
Subject #7	86	20	76%	89	3.4%
MAE±SD				3.21±3.15	

Table 4 Processing times for different video lengths

Video Length (s)	Generating rPPG (s)	Proposed Algorithm (s)	Total Time (s)
36	1.24	0.1	1.34
64.8	1.95	0.18	2.13

Table 5 Performance comparison of the state-of-the-art rPPG algorithms in terms of MAE and SD

	Distance	Motion Reference Sensor	Processing Method	MAE±SD
[14]	0.5 m	No	ASTHF	4.81±3.39
[8]	0.4 m	No	PCA+DCT	4.69±3.43
[24]	1 m	Near-Infrared sensor	SoftSig	5.23
[16]	1.5 m	Multi-Imager +16 mm Lens	DRCN	3.08
[30]	0.5 m	PFF database	CNN	4.65
Proposed	0.4 m	No	CWT+RBLWR	3.58±2.4

respectively. The work in [30] achieved MAE of 4.65 for 0.5 m distance using a convolutional neural network.

The proposed method based on CWT and RBLWR works from a 0.4 m distance without any additional sensors, achieving an MAE ± SD of 3.58 ± 2.4. These MAE ± SD metrics were calculated based on the results of the seven subjects, as summarized in Table 1. As filtering and MA removal are applied to 1D data points, not to the 2D image frames, real-time HR measurement is achieved.

The background light fluctuations influence the accuracy of the results. In addition, accuracy decreases as the target distance increases to more than 0.4 m. In future work, we will study the effect of changing the distance and the background light on the accuracy of the proposed method.

7 Conclusion

A recursive normalization method with an adaptive window was proposed to reduce the influence of MAs on the rPPG signal. The amplitude variations due to low-frequency MA noise were well corrected using a Mexican hat wavelet with a scale of 8. Before using the wavelet transformation, the detection error ranged from 47% to 74%. After the wavelet reconstruction of the signal, the highest detection error was 9.5%. These improvements were due to the removal of MAs from the reconstructed signal. The PPG signal was recursively normalized until all the peaks and valleys were detected. Outlier removal was applied to the PPG, with 2% of the total samples removed from the original signal. The highest HR error measured using the proposed algorithm for rPPG signal was 7%. The proposed method shows high performance and real-time operation compared with state-of-the-art methods. HR can be measured remotely from a 0.4 m distance without any additional sensors, achieving an MAE ± SD of 3.4 ± 2.2. Our experiments were performed indoors under fixed light conditions. We plan to study the effect of varying the surrounding light intensity on the rPPG signal and HR accuracy.

The next step is to design an Android app to calculate the HR from recorded video based on the proposed algorithm. The proposed algorithm can run in real time at 1.34 s using 2-core 2.5 GHz CPU, with a CPU usage of 30%. Energy consumption should be minimized for implementation in

battery-powered devices, such as smartphones. When using Samsung S20 with a 2.73 GHz Octa-Core processor, sufficient computational power remains available for additional processing.

Data availability The datasets generated during and/or analyzed during the current study are not publicly available but are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest.

References

1. Abate AF, Bisogni C, Cascone L, Castiglione A, Costabile G, Mercuri I (2020) Social robot interactions for social engineering: opportunities and open issues. In Intl Conf on cyber science and technology congress (DASC/PiCom/CBDCom/CyberSciTech), Calgary, AB, Canada
2. Atef M, Xiyan L, Lian G, Lian Y (2016) PTT based continuous time non-invasive blood pressure system. In: IEEE 59th International Midwest Symposium on Circuits and Systems (MWSCAS), Abu Dhabi
3. Beurer Art.- No. 454.71 FDA, [Online]. Available: https://fda.report/GUDID/10852547004211. Accessed 10 12 2021
4. Blackford EB et al (2018) Remote spectral measurements of the blood volume pulse with applications for imaging photoplethysmography. Proc SPIE 10501:1–8
5. Christinaki E et al. (2014) Comparison of blind source separation algorithms for optical heart rate monitoring. In: 4th International Conference on Wireless Mobile Communication and Healthcare (MOBIHEALTH), Athens, Greece
6. Demirezen H, Eroglu Erdem C (2021) Heart rate estimation from facial videos using nonlinear mode decomposition and improved consistency check. SIViP 15:1415–1423
7. Duan K, Qian Z, Atef M, Wang G (2016) A feature exploration methodology for learning based cuffless blood pressure measurement using photoplethysmography. In: 8th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA
8. Haque MA, Irani R, Nasrollahi K, Moeslund TB (2016) Heartbeat rate measurement from facial video. IEEE Intell Syst 31(3):40–48
9. Hasan Z, Ramamurthy SR, Roy N (2021) MPSC-rPPG Datase. IEEE Dataport
10. Helleptte QL, Van N (2021) PPG sensors for the new Normal: a review. In: 2021 18th International SoC Design Conference (ISOCC), Jeju Island, Korea
11. Jean Effil N, Rajeswari R (2022) Wavelet scattering transform and long short-term memory network-based noninvasive blood pressure estimation from Photoplethysmograph signals. SIViP 16:1–9
12. Kong Y, Chon KH (2019) Heart rate tracking using a wearable Photoplethysmographic sensor during treadmill exercise. IEEE Access 7:152421–152428
13. Lee H, Chung H, Lee J (2019) Motion artifact cancellation in wearable Photoplethysmography using gyroscope. IEEE Sensors J 19(3):1166–1175
14. Liu X, Yang X, Wang D, Wong A (2021) Detecting pulse rates from facial videos recorded in unstable lighting conditions: an adaptive spatiotemporal homomorphic filtering algorithm. IEEE Trans Instrum Meas 70:1–15
15. Luguern D, Macwan R, Benezeth Y, Moser V, Dunbar LA (2021) Wavelet variance maximization: a contactless respiration rate estimation. Biomed Signal Process Control 63:102263
16. McDuff D (2018) Deep super resolution for recovering physiological information from videos. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA
17. McDuff DJ, Blackford EB, Estepp JR (2018) Fusing partial camera signals for noncontact pulse rate variability measurement. IEEE Trans Biomed Eng 65(8):1725–1739
18. Nabavi S, Bhadra S (2020) A robust fusion method for motion artifacts reduction in Photoplethysmography signal. IEEE Trans Instrum Meas 69(12):9599–9608
19. Shaik KB, Ganesan P, Kalist V, Sathish BS, Jenitha JMM (2015) Comparative Study of Skin Color Detection and Segmentation in HSV and YCbCr Color Space. In: 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)

20. Singla M, Azeemuddin S, Sistla P (2020) Accurate fiducial point detection using Haar wavelet for beat-by-beat blood pressure estimation. IEEE J Transl Eng Health Med 8:1–11
21. Tabei F, Kumar R, Phan TN, McManus DD, Chong JW (2018) A novel personalized motion and noise artifact (MNA) detection method for smartphone photoplethysmograph (PPG) signals. IEEE Access 6:60498–60512
22. Tabei F, Gresham JM, Askarian B, Jung K, Chong JW (2020) Cuff-less blood pressure monitoring system using smartphones. IEEE Access 8:1534–11545
23. Wang G, Atef M, Lian Y (2018) Towards a continuous non-invasive Cuffless blood pressure monitoring system using PPG: systems and circuits review. IEEE Circ Syst Mag 18(3):6–26
24. Wang W, Den Brinker AC, De Haan G (2019) Single-Element Remote-PPG. IEEE Trans Biomed Eng 66(7):2032–2043
25. Wu C, Chen I, Fang W (2017) An implementation of motion artifacts elimination for PPG signal processing based on recursive least squares adaptive filter. In: IEEE biomedical circuits and systems conference (BioCAS), Turin, Italy
26. Zhang Y, Song S, Vullings R, Biswas D, Simões-Capela N, van Helleputte N, van Hoof C, Groenendaal W (2019) Motion artifact reduction for wrist-worn Photoplethysmograph sensors based on different wavelengths. Sensors 19(637):1–18
27. Zhang G, Mei Z, Zhang Y, Ma X, Lo B, Chen D, Zhang Y (2020) A noninvasive blood glucose monitoring system based on smartphone PPG signal processing and machine learning. IEEE Trans Ind Inf 16(11):7209–7218
28. Zhang G, Zhang S, Dia Y, Shi B (2021) Using rear smartphone cameras as sensors. IEEE Access 9:20460–20468
29. Zhao C, Zeng W, Hu D, Liu H (2021) Robust heart rate monitoring by a single wrist-worn accelerometer based on signal decomposition. IEEE Sensors J 21(14):15962–15971
30. Zheng K, Ci K, Li H, Shao L, Sun G, Liu J, Cui J (2022) Heart rate prediction from facial video with masks sing eye location and corrected by convolutional neural networks. Biomed Signal Process Control 75:103609

Asmaa Hosni received the B.Sc. and M.Sc. degrees in electrical engineering, Computer Engineering from Assiut University, Egypt, in 2000 and 2005, respectively. She received his PhD in 2013 from Vienna University of Technology, Institute of Software Technology, Interactive Media Systems Group. Asmaa Hosni has been an Assistant Professor at Assiut University, Egypt since 2013. Her research interests are in the areas of computer vision, computer graphics and biomedical signal processing.

Mohamed Atef (M08, SM12) received the B.Sc. and M.Sc. degrees in electrical engineering, electronics and communications from Assiut University, Egypt, in 2000 and 2005, respectively. From 2006 to 2007, he got a research scholarship from the Czech Technical University in Prague, Department of Microelectronics, he worked on the improvement of quantum dot optical properties. He received his PhD in 2010 from Vienna University of Technology, Institute of Electrodynamics, Microwave and Circuit Engineering, and then worked as a post-doctoral researcher until the end of 2012. Dr. Atef visited the School of Microelectronics, Shanghai Jiao Tong University, China, from 2015 to 2017. Mohamed Atef has been an Associate Professor and Professor at Assiut University, Egypt since 2016 and 2020 respectively. In 2020 he joined Electrical Engineering Department, United Arab Emirates University, United Arab Emirates, and he is still serving there as an Associate Professor. His research interests are in the areas of optoelectronic integrated circuits, and biomedical circuits and system. He is an author of two Springer books, "Optical Communication over Plastic Optical Fibers: Integrated Optical Receiver Technology" and "Optoelectronic Circuits in Nanometer CMOS Technology" and an author and co-author of more than 80 scientific publications. Mohamed Atef served as a TPC member in many IEEE conference, as a lead editor for Sensors Journal special issue (2019–2020), TBioCAS WFISEDIT 2019 and TBioCAS ISICAS 2020 special issues, and as an associate editor for TBioCAS (2020-present). Mohamed Atef is a member of the Biomedical and Life Science Circuits and Systems Technical Committee (2018-present). He was awarded the State Encouragement Award in Advanced Technological Sciences Serving the Engineering Sciences for 2018 from the Egyptian Academy of Scientific Research and Technology (ASRT).