IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 5, 2018, accepted February 26, 2018, date of publication March 2, 2018, date of current version March 16, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2811392

Remote Optical Cardiopulmonary Signal Extraction With Noise Artifact Removal, Multiple Subject Detection & Long-Distance

ALI AL-NAJI[1,2], (Member, IEEE), AND JAVAAN CHAHL[1,3], (Member, IEEE)
[1]School of Engineering, University of South Australia, Mawson Lakes, SA 5095, Australia
[2]Electrical Engineering Technical College, Middle Technical University, Baghdad 10022, Iraq
[3]Joint and Operations Analysis Division, Defence Science and Technology Group, Melbourne, VIC 3207, Australia

Corresponding author: Ali Al-Naji (ali_abdulelah_noori.al-naji@mymail.unisa.edu.au)

ABSTRACT Most existing imaging technologies are prone to noise artifact resulting from movement of subjects, movement of the camera, facial expressions, talking, skin tone, and lighting variations. In addition, these technologies have been limited to detecting physiological signs from a single subject at a time. Still, another challenge facing these technologies is that they have only been demonstrated over a limited range to a subject. This paper proposes an efficient remote imaging system with a new, low-complexity method for noise artifact removal to extract the cardiopulmonary signal from a number of subjects (up to six people) and at long-distances (up to 60 m). The experimental results of the proposed system showed a strong agreement, high correlation, and low noise level with the reference measurements and outperformed the conventional measurement methods such as independent component analysis and principle component analysis in both the stationary and non-stationary scenarios.

INDEX TERMS Imaging photoplethysmography (iPPG), cardiopulmonary signal, video magnification, signal decomposition technique, blind source separation (BSS), canonical correlation analysis (CCA).

I. INTRODUCTION

Camera imaging technologies have been extensively used as remote physiological signs monitoring systems in last decade. These technologies can be divided into two main categories: color-based methods, sometime called imaging photoplethysmography (iPPG), which rely on optical features passing through or reflecting from the epidermis due to cardiorespiratory activity; and motion-based methods, which rely on the mechanical activity of the cardiovascular and respiratory systems at certain regions of the human body. Each method, however, has its merits and dismerits under different assumptions, leading to several challenges for the research community, including noise artifacts, multiple detection and long-distance. This paper is focusing on the lack of systems for extracting the cardiopulmonary signal when the challenges of noise artifact, multiple detection and long-distance are main considerations. The major studies that used camera imaging-based technologies and the main challenges are reviewed and summarised in Table 1.

Based on the studies listed in Table 1, the cardiopulmonary signal could be remotely acquired from different types of camera sensors (e.g. video camera, Webcam, time of flight camera, Kinect and drone), using different methods for noise artifact removal with varying degrees of the noise immunity; however, extraction of the cardiopulmonary signal was mostly affected by noise artifacts caused by a range of effects, including subject movement (e.g. head rotation, walking, facial expressions, talking and blinking), skin tone and lighting variations. In addition, these studies were limited to extracting the cardiopulmonary signal of one subject at a time. Furthermore, all studies listed in these tables were limited to detecting the cardiopulmonary signal at short distances. Therefore, the aim of this paper is to propose an efficient remote imaging system based on both color and motion methods to extract the cardiopulmonary signal under different scenarios (stationary and non-stationary scenarios) with proposing a new method for noise artifact removal based on a combination of a signal decomposition technique and blind source separation (BSS). Meanwhile, the proposed system in this paper was able to detect the cardiopulmonary signal from multiple subjects (up to six people) and at long distances (up to 60m) under different assumptions. Thus,

TABLE 1. Main challenges of the camera imaging-based technologies.

	Activity	Sensor type	Principle	The used technique	Noise artifact	No. of subjects	Detection range (m)
Nakajima et al. [1]	Cardiac and respiratory	CCD	Motion	OF	*	1	2
Nakajima et al. [2]	Respiratory	CCD	Motion	OF	*	1	2
Parra & Da Costa [3]	Cardiac	TV camera	Motion	OF	*	1	~1
Wiesner and Yaniv [4]	Respiratory	Webcam	Motion	PCA	**	1	~1
Verkruysse et al. [5]	Cardiac and respiratory	Digital camera	Color	Three channel analysis	*	1	~1
Poh et al. [6, 7]	Cardiac and respiratory	Webcam	Color	ICA	**	1, 3	0.5
Lewandowska et al. [8]	Cardiac	Webcam	Color	PCA	*	1	1
Pursche et al. [9]	Cardiac	Webcam	Color	ICA	**	1	0.5
Scalise et al. [10]	Cardiac	Webcam	Color	ICA	**	1	0.2
Sun et al. [11, 12]	Cardiac and respiratory	CMOS+ Webcam	Color	ICA	**	1	0.35-0.4
Penne et al. [13] Falie et al [14] Yu et al [15]	Respiratory	ToF camera	Motion	Depth image sequence analysis	***	1	0.6-1, 1, 1.25
Kwon et al. [16]	Cardiac	Cell phone	Color	ICA	**	1	0.3
Bousefsaf et al. [17]	Cardiac	Webcam	Color	Single channel analysis + skin detections	***	1	1
Yu et al [18] Xia, & Siochi [19]	Respiratory	Kinect	Motion	Depth image sequence	***	1	1.4, 2
Bernacchia et al. [20]	Cardiac and respiratory	Kinect	Motion	Depth info + ICA	***	1	1.2
Alnaji et al. [21]	Respiratory	Kinect	Motion	Frame subtraction	***	1	2.5
De Haan & Jeanne [22]	Cardiac	CCD	Color	CHRO	***	1	~1
De Haan & Van Leest [23]	Cardiac	CCD	Color	PBV	***	1	~1
Li et al. [24]	Cardiac	iSight camera	Color	Single channel analysis	***	1	0.35, 0.5
McDuff et al. [25]	Cardiac and respiratory	Digital camera	Color	Single/Multiple channel analysis	**	1	3
Balakrishnan et al. [26]	Cardiac	Digital camera	Motion	Feature extraction + PCA	*	1	1
Shan & Yu [27]	Cardiac	Cell phone	Motion	Feature extraction + ICA	*	1	0.4
Irani et al. [28]	Cardiac	Webcam	Motion	Feature extraction + PCA	**	1	~1
Feng et al. [29]	Cardiac	Webcam	Color	ICA	**	1	0.75
Tarassenko et al. [30]	Cardiac and respiratory	CCD	Color	RGB channel analysis + AR model	**	1	1
Hsu et al. [31]	Cardiac	CMOS	Color	SVR ICA/ CHRO	***	1	~1
Feng et al. [32]	Cardiac	Webcam	Color	Multiple channel analysis + GRD	***	1	0.75
Chen et al. [33] Lin et al. [34]	Cardiac	Digital camera	Color	Single channel analysis +EEMD	***	1	0.1-.025, 0.1
Wiede et al. [35]	Cardiac	CCD	Color	ICA/PCA	**	1	0.5-2
Cheng et al. [36]	Cardiac	Webcam	Color	JBSS + EEMD	***	1	0.5
He et al. [37]	Cardiac	Digital camera	Motion	Video magnification	*	1	~0.5
Prathosh et al. [38]	Respiratory	CMOS	Motion	Selective ensemble aggregation	**	1	~2
Alnaji & Chahl [39-41]	Cardiac and respiratory	Digital camera	Motion	Frame subtraction	**	1	0.5-1
Wang et al. [42]	Cardiac	CCD	Color	2SR	***	1	2
Wang et al. [43]	Cardiac	CCD	Color	POS	***	1	~1.5
Xu et al. [44]	Cardiac	Webcam	Color	PLS+MEMD	***	1	0.5
Alnaji et al. [45]	Cardiac and respiratory	Drone	Color	Video magnification+BSS	***	1	3

Note: AR= auto-regressive; CCD= charge-coupled device; CHRO= chrominance; CMOS= complementary metal-oxide-semiconductor; EEMD= ensemble empirical mode decomposition; GRD= green/red difference; ICA= independent component analysis; JBSS= joint blind source separation; MEMD= multivariate empirical mode decomposition; OF= optical flow; PBV= blood volume pulse vector; PCA= principle component analysis; PLS= partial least squares; POS= plane orthogonal to skin; SVR= support vector regression; 2SR= spatial subspace rotation; ToF= time of flight.

the proposed system with the noise artifact removal method may be of value for upcoming remote monitoring systems in biomedical and clinical applications.

This paper is divided into five main sections. Section I provides an introduction to the proposed system and describes the related work in a table. Section II is about the methods and

procedures of the proposed system, including the subjects, validation methods and the details of the system framework. Section III contains the experimental results of the proposed system in different scenarios under three assumptions: presence of noise artifacts, multiple detection and long-distance detection. Finally, discussion and conclusion are made in section IV and V respectively.

II. METHODS AND PROCEDURES

A. ETHICS CONSIDERATION

The research procedure described in this paper adhered to the ethical tenets of the Declaration of Helsinki (Finland 1964) and all procedures involving human subjects were granted by the University of South Australia Human Research Ethics Committee (Adelaide, South Australia, Protocol number: 0000034901). A written informed consent had been obtained from all subjects after a full explanation of the research procedure before commencing the experiment.

B. SUBJECTS

The paper was conducted using three groups of subjects. The first group (G_1) included 15 subjects (10 males and 5 females) aged between 2 and 40 years and with different skin tones. The video data was collected using a hovering unmanned aerial vehicle (UAV) (3DR solo, 10MP, 5.4 mm GoPro Lens). The experiment was performed at various times of the day under different lighting conditions in outdoor and indoor environments at a viewing distance of 3 m between the subject and UAV camera. The subjects were instructed to breathe normally and to partake in two scenarios during videoing, including a stationary scenario and a non-stationary scenario (e.g. face rotation, facial expression, talking and blinking). The second group (G_2) included 20 subjects (12 males and 8 females) between the ages of 6 and 50 years with different skin tone. The video data was captured with a digital camera (Nikon D5300) with a lens (18-55mm) and located at a distance of 3 m away from the subjects in both outdoor and indoor environments at 60Hz and 30Hz frame rates with different resolutions. The subjects were also instructed to breathe normally when doing the two scenarios during videoing (stationary and non-stationary scenarios). The third group (G_3) included 10 subjects (8 males and 2 females) within the ages of 18–40 years with different skin tones. The video data was recorded by a digital camera (Nikon D5300 with a lens (200–500 mm) with 60 fps and a pixel resolution of 1080 × 720. The video data were captured over different days at various distances (approximately between 50 m and 60 m) and different lighting conditions in two different scenarios (stationary scenario and non-stationary scenarios). All video data was saved in AVI format on a computer for processing and analysis.

C. VALIDATION METHODS

The proposed system was validated by comparing its experimental results with the results obtained from the reference methods. The reference measurement methods, including pulse oximeter (Rossmax SA210) and respiratory belt transducer (Piezo MLT1132), were applied to the subject body for validation purposes.

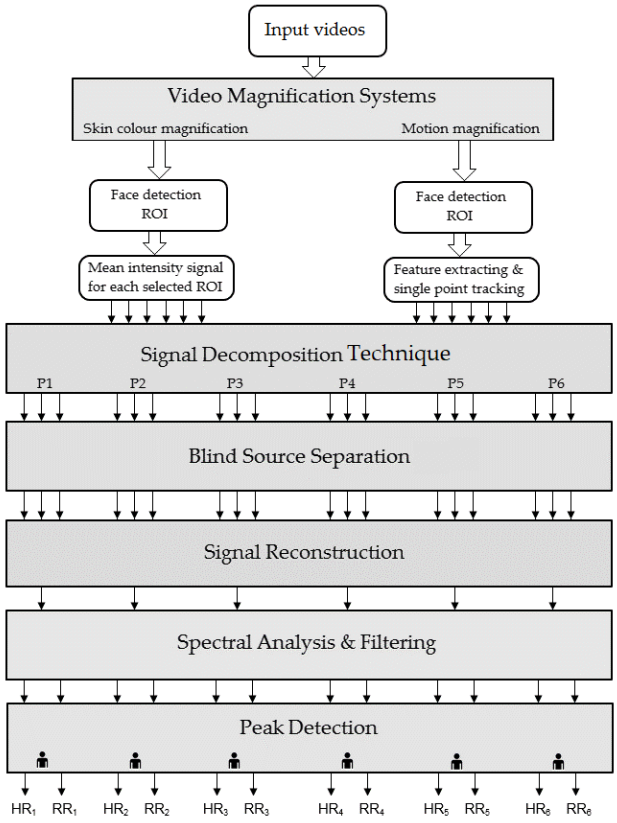

FIGURE 1. Block diagram of the proposed system including the noise artifact removal method.

D. SYSTEM FRAMEWORK

The overall proposed system is composed of several processing techniques: video magnification systems, face detection, signal decomposition, blind source separation, signal reconstruction, spectral analysis & filtering and peak detection as shown in Fig. 1.

1) VIDEO MAGNIFICATION TECHNIQUES

a: SKIN COLOR MAGNIFICATION

Skin color variations in the subject's face caused by the cardiac activity are not visible to recognise; therefore, developing video magnification technique was used to amplify these variations before data analysis. Some modifications have been done on the standard video magnification technique [46] to suit the proposed system, including wavelet pyramid decomposition and an elliptic band-pass filter. In this magnification system, wavelet pyramid decomposition, an elliptic band-pass filter (0.15 Hz- 3 Hz) and the Lanczos resampling method [47] were used to remove motion artifact, avoid motion magnification during the magnification process, and to reduce the processing time. The G component

was only selected and magnified 15x since this component has been reported to contain the strongest cardiac information signal [5]–[7], [48].

b: MOTION MAGNIFICATION

Head motion generated by blood circulation between the heart to the head via the carotid arteries is very small (about <0.5mm), and hard or impossible to visually detect; therefore, the magnification system proposed in [49] was used to amplify it before data analysis. In this magnification system, the image sequences of the video data were converted from RGB color space to YC_bC_r color space (instead of YIQ) to separate the color data from the intensity data. Only the intensity component (Y) was used and downsized to reduce the processing time. The Y component was then decomposed into different spatial frequency bands using a wavelet pyramid, followed by a Chebyshev Type I band-pass filter with a band of 0.15 Hz- 3 Hz to extract frequency bands of interest. The filtered signal was then magnified by 15x. The magnified signal was then reconstructed and added back to the original signal.

2) ROI DETECTION AND SIGNAL ACQUISITION

Based on the video obtained from the color magnification system, an efficient face detection method, proposed by Liao *et al.* [50], was applied to increase system accuracy and performance because this method is robust to challenges associated with unconstrained faces (e.g. faces in a crowd, face rotation, inclined or angled faces). A number of ROIs were then detected according to the number of subjects in the input video. Also, the regions of the eyes and mouth for each ROI were eliminated to reduce the noise artifact resulting from blinking and talking during the measurements. For each ROI, the time-series signal, $i_c(t)$, was calculated by averaging of all the image pixel intensities in the selected region as follows:

$$i_c(t) = \frac{\sum_{x,y \in ROI} I(x, y, t)}{|ROI_1|} \tag{1}$$

where $I(x, y, t)$ is the pixel intensity at a location (x, y) and time (t), and $|ROI_1|$ is the size of the ROI for the first subject. The acquired signals from six subjects are denoted by $i_{c1}(t)$, $i_{c2}(t), \ldots i_{c6}(t)$ respectively.

Based on the magnified video obtained from [49], Liao's method [50] was also applied for detecting faces and only a small region on the forehead (rectangle region) was selected as the interested region. A single point in the centre of the rectangle was then chosen as a feature point which has two components in both the vertical and horizontal axis. The Lucas-Kanade tracker was used only to track its vertical component over time which corresponds to head motion and thus obtains a time-series signal, $i_m(t)$, as follows:

$$i_m(t) = \{V_1(t), V_2(t), \ldots.., V_N(t)\} \tag{2}$$

where $V(t)$ represents the raw signals obtained from y-axis (vertical trajectories), and N, the number of frames. The acquired signals from six people are denoted by $i_{m1}(t)$, $i_{m2}(t), \ldots i_{m6}(t)$ respectively.

3) IMPROVED SIGNAL DECOMPOSITION TECHNIQUE

Signal decomposition techniques are widely used for biomedical signal processing tasks to decompose a temporal signal into a set of amplitude and frequency components, while the most interesting components need to be selected to represent the original signal. The empirical mode decomposition (EMD) is one of the most common signal decomposition techniques that is used to remove the noise artifacts from biomedical signals [51]–[53]. The EMD is a time-frequency analysis technique for adaptively decomposing a given nonlinear and/or non-stationary signal into a set of IMFs. Later, a noise-assisted version, called EEMD, was proposed by Wu and Huang [54] to remove the mode mixing problem associated with the EMD. However, EEMD still has some limitations related to residual noise, reconstruction error and modes for different realizations of signal plus noise [55]. Another improved version of the EEMD, the CEEMD technique was proposed in [56] to solve the problems associated with the EEMD. In this paper, a recent improved version, namely the CEEMD technique with adaptive noise, was used to efficiently decompose the signal of interest into a set of IMFs with less noise and more physical meaning than other techniques. In addition, this improved technique outperformed other signal decomposition techniques in terms of reduction of the amount of the residual noise from the modes and spurious modes overlapping.

The flowchart diagram of the signal decomposition based on the improved CEEMD technique with adaptive noise is shown in Fig. 2.

where $x^{(i)}$ is a combination of the acquired signal $x(t)$ and the added white Gaussian noise $\omega^{(i)}$ $[i = 1, 2, \ldots L$ (no. of realizations)], the coefficient β allows the selection of SNR at each stage, the operator $E[\blacksquare]$ produces the m^{th} modes obtained by EMD to generate number of residues r and modes IMF and the operator $M\{\blacksquare\}$ produces the local means of each realization of the signal. The signal of interests can be recovered by the following expression: [55]

$$x(t) = \sum_{m=1}^{N} IMF_m + r_N \tag{3}$$

where IMF_m is the m^{th} mode of the signal of interest, r_N is the final residue obtained from $(r_N(t) - r_{N+1}(t))$ and N is the total number of modes.

An example of the generation of eight IMFs of the signal of interest based on improved CEEMD technique with adaptive noise with 100 realizations and 100 iterations is given in Fig. 3.

4) BLIND SOURCE SEPARATION BSS

BSS is an increasingly popular data analysis technique with a broad range of applications, which is used to extract a set of source signals from their mixtures without prior

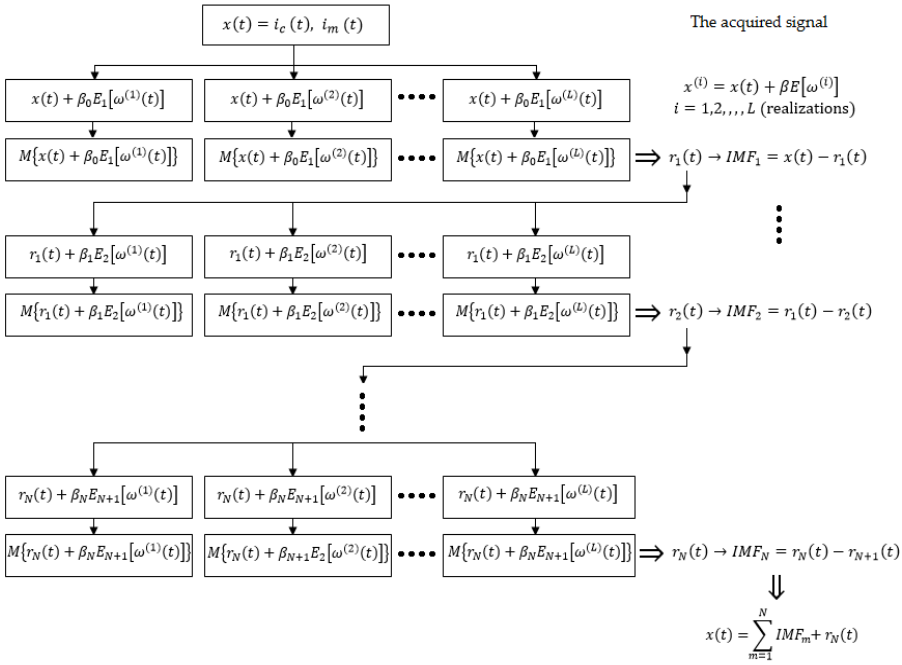

FIGURE 2. Flowchart diagram of the improved signal decomposition technique.

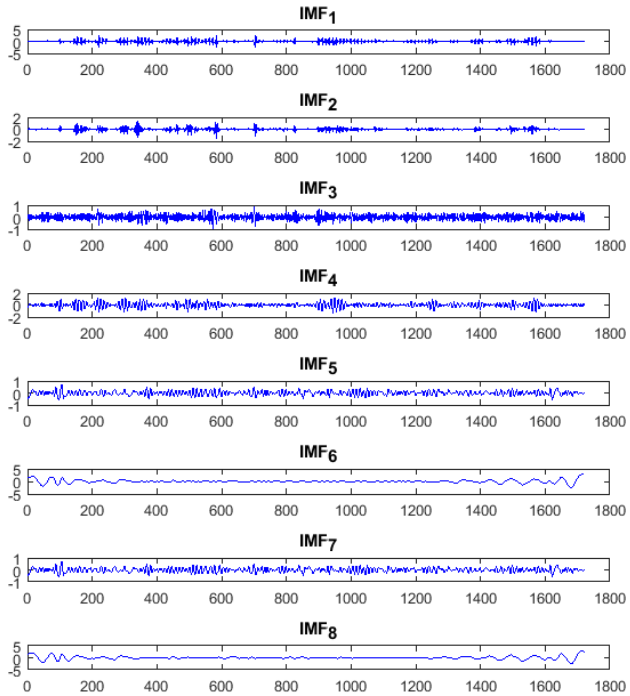

FIGURE 3. An example of eight IMF components of the improved CEEMD technique with adaptive noise.

knowledge of the characteristics of the source signals or mixing process. The canonical correlation analysis (CCA.) technique [57], [58] was used in this study as a BSS to separate a number of mixed signals since this technique can achieve better performance for BSS than the independent component analysis (ICA) technique with less computational complexity [52], [59], [60]. Also, CCA technique is an effective signal processing technique that can be used to remove noise artifacts from the biomedical signals [61]–[64].

To demonstrate how CCA technique operates as the BSS, let j and k be two multi-dimensional random signals with N mixtures. Consider the linear combinations of these signals, known as the canonical variates as follows [65]:

$$j = W_{j_1} j_1 + W_{j_2} j_2 + \ldots, W_{j_N} j_N = W_j^T j, \tag{4}$$

$$k = W_{k_1} k_1 + W_{k_2} k_2 + \ldots, W_{k_N} k_N = W_k^T k \tag{5}$$

where $W_j = [W_{j_1}, W_{j_2}, \ldots . W_{j_N}]^{\mathrm{T}}$ and $W_k = [W_{k_1}, W_{k_2}, W_{k_N}]^{\mathrm{T}}$ are weighting vectors of j and k that maximize the correlation between j and k by solving the following maximization problem:

$$\begin{aligned} \rho &= cor\,(j,k) = \max_{W_j, W_k} cor(j,k) = \frac{E[jk]}{\sqrt{E\left[j^2\right] E[k^2]}} \\ &= \frac{E[\left(W_j^T j\right)\left(W_k^T k\right)]}{E\sqrt{E\left[(W_j^T j)^2\right] E[(W_k^T k)^2]}} \\ &= \frac{W_j^T \mathrm{C}_{jk} W_k}{\sqrt{(W_j^T \mathrm{C}_{jj} W_j)(W_k^T \mathrm{C}_{kk} W_k)}} \end{aligned} \tag{6}$$

where C_{jj} and C_{kk} are the non-singular within-set covariance matrices of j and k, respectively, and C_{jk} is the between-sets covariance matrix. The maximization problem with respect

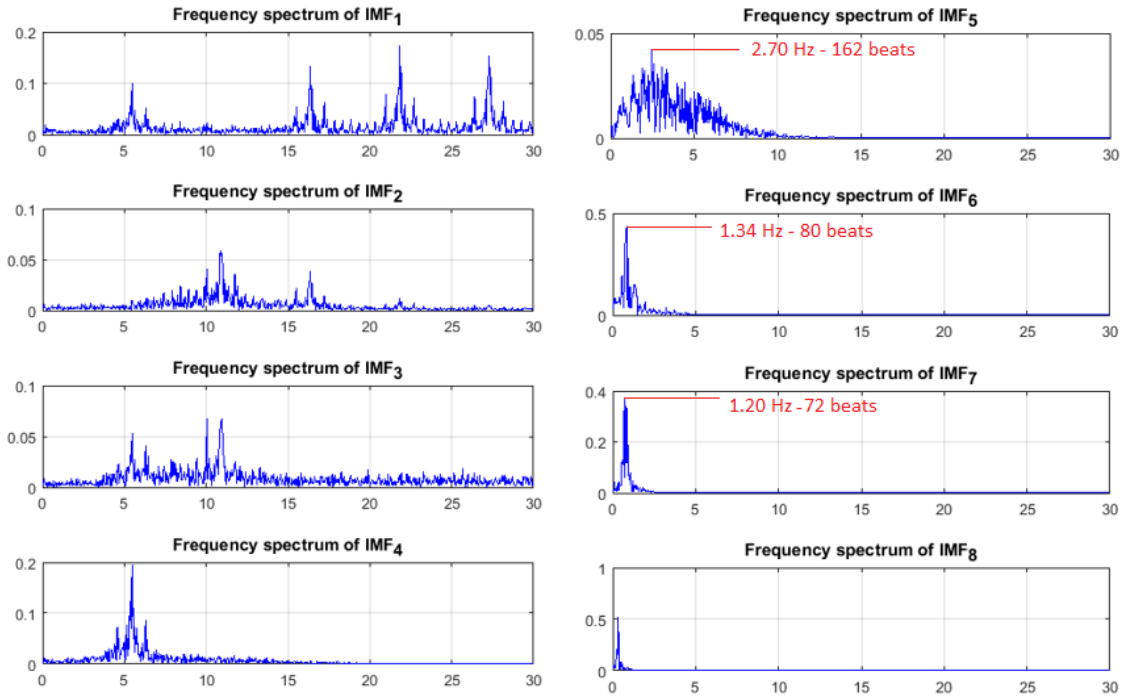

FIGURE 4. The frequency spectra of decomposed IMF components.

to W_j and W_k can be solved by the following equation:

$$\begin{bmatrix} C_{jj} & 0 \\ 0 & C_{kk} \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_{jk} \\ C_{kj} & 0 \end{bmatrix} \begin{pmatrix} \hat{W}_j \\ \hat{W}_k \end{pmatrix} = \rho \begin{pmatrix} \hat{W}_j \\ \hat{W}_k \end{pmatrix} \quad (7)$$

After manipulating Eq. (7), a complete description of the canonical correlations can be expressed as follows:

$$\begin{cases} C_{jj}^{-1} C_{jk} \hat{W}_k = \rho \hat{W}_j, \\ C_{kk}^{-1} C_{kj} \hat{W}_j = \rho \hat{W}_k. \end{cases} \quad (8)$$

The K estimates of the sources signals, $z_i(t), i = 1, 2..K$, can be obtained by

$$z_i(t) = \hat{W}_{j_i}^T j(t) \approx s_i(t) \quad (9)$$

To select the best IMFs that should be used for estimating cardiopulmonary signal, the frequency spectral analysis of each component has been done using FFT as shown in Fig. 4.

Fig. 4 shows the frequency spectra of the decomposed IMF components using the improved CEEMD technique with adaptive noise. Clearly, the frequency bands of the components (IMF5, IMF6 and IMF7) have the best cardiorespiratory frequency bands that fall within a frequency range of 0.15-3 Hz, with the maximum frequency spectra of 2.7 Hz, 1.34 Hz and 1.2 Hz corresponding to 162 beats, 80 beats and 72 beats respectively, while other spectra fall outside this range; Therefore, these components were chosen as inputs to the CCA technique for extracting the cardiopulmonary signal as shown in Fig. 5.

5) SIGNAL RECONSTRUCTION

The acquired signal (a) was decomposed into a multichannel signal (A) using the improved CEEMD technique. The IMF components falling outside frequency bands of interest were removed, while the IMF components within frequency bands of interest were selected as inputs with the un-mixing matrix W of the CCA technique. The original multichannel signal $\tilde{A}$ was then reconstructed without undesirable IMF components (artifact components) using the inverse of the un-mixing matrix W^{-1}. After that, the processed signal $\tilde{a}$ without the noise artifacts generating from the effects of subjects' movement, camera motion, skin tone and lighting variations could be found by adding the wanted IMF components in the $\tilde{A}$ matrix.

6) SPECTRAL ANALYSIS & FILTERING

Spectral analysis method based on the FFT was used to extract the cardiopulmonary frequencies, followed by two temporal filters of 5th order Butterworth band-pass at 0.5-3Hz corresponding to the heart pulse range (30-180 beats/min), and 0.15-0.5Hz corresponding to the breathing range (9-30 breaths/min). The inverse FFT was then applied to the filtered signals to get the time-series cardiac and respiratory signals respectively.

7) PEAK DETECTION

The peak detection method based on the MATLAB built-in function '*findpeaks*' was used to determine the periodicity of peaks, peaks locations and number of peaks of the

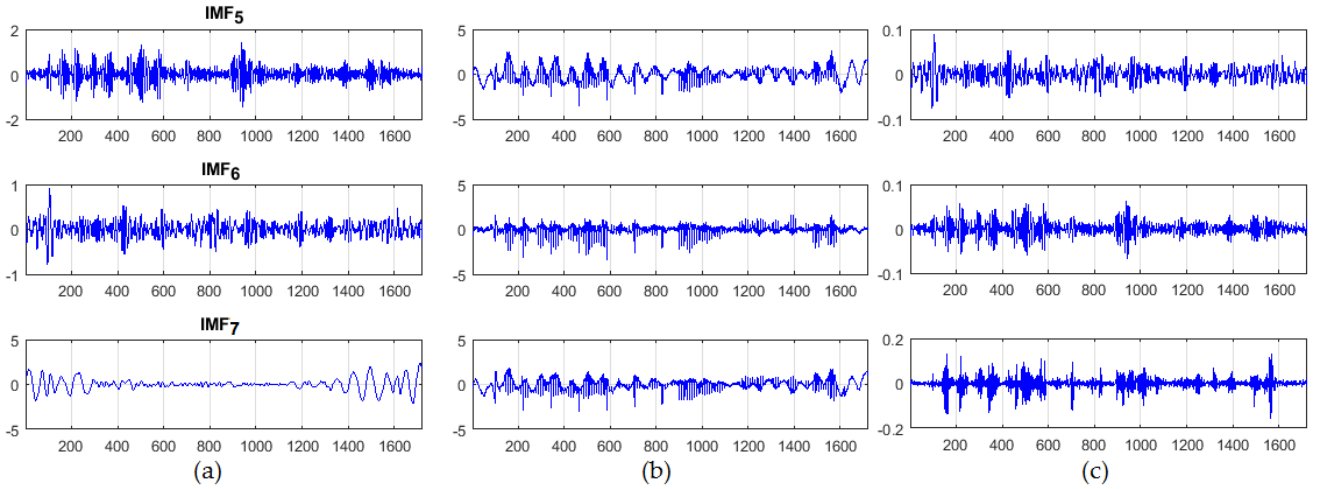

FIGURE 5. The CCA technique (a) The selected IMF components (S) (b) Transformation [$x = W.S$], where W is an un-mixing matrix and (c) CCA outputs [$y = CCA(x)$].

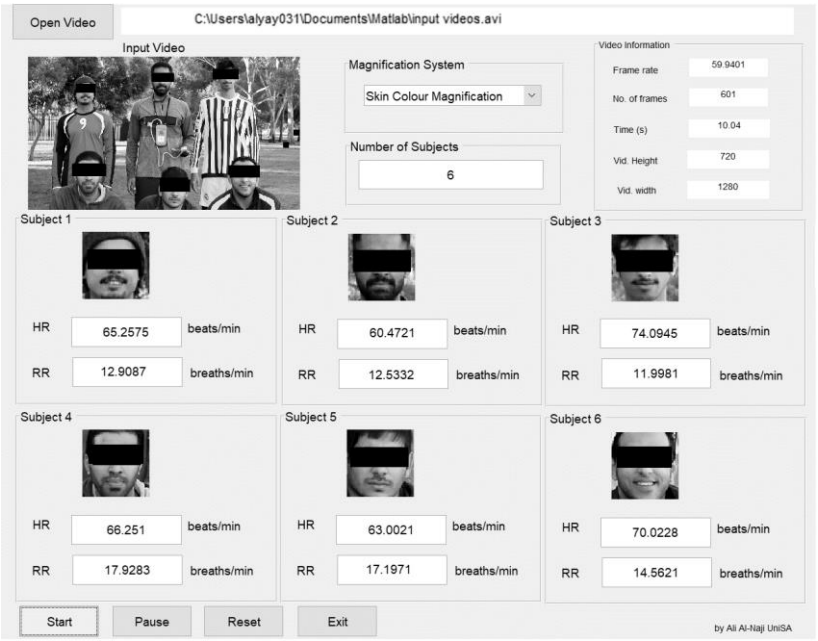

FIGURE 6. The GUI operator panel of the proposed system.

acquired signals. The measured value (M_v) of the heart and respiratory rates per minute can be calculated using the following equation:

$$M_v = \frac{60p}{t} \tag{10}$$

where p is the number of peaks of the acquired signal, t is the length of the video signal recording in seconds. In addition, once the peaks and their locations (locs) were determined, the total cycle length (Tc) between the two consecutive peaks could be determined using, Tc = mean (diff (locs)), which represents heart rate variability and the respiratory cycle.

E. GRAPHICAL USER INTERFACE

A graphical user interface (GUI) is a pictorial platform to a software system that enables a user to perform interactive tasks with graphical objects. In this paper, the GUI model was implemented in the MATLAB environment to provide an easy tool that allows a user to load video data, select the magnification type, and execute the proposed system and configurations. The GUI operator panel of the proposed system is shown in Fig. 6.

It can be seen from Fig. 6 that the proposed GUI allows for the user to see video information, the number of subjects in each input video, select the magnification type and

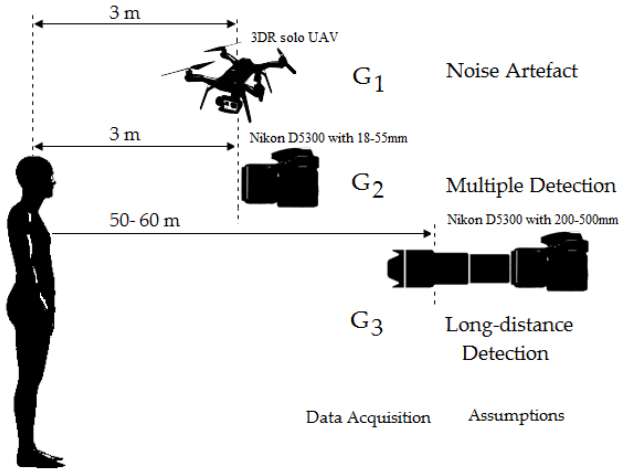

FIGURE 7. The experimental setup and data acquisition of the proposed system.

recognize the subjects' faces to enable the user to distinguish the HR and RR readings for each subject.

III. EXPERIMENTAL RESULTS

This section presents the experimental results of the proposed system under three assumptions: noise artifact, multiple detection, and long-distance detection and compares the measurements obtained from the proposed system based on both color and motion analysis against the measurements obtained from the conventional measurement methods such as ICA-based method [6], [7] and PCA-based method [8] in the stationary and non-stationary scenarios. The quantitative experimental results of the proposed system were evaluated and validated against the measurements obtained from the reference methods. The Bland-Altman method [66], Pearson's correlation coefficient (PCC), Spearman's correlation coefficient (SCC), Kendall's correlation coefficient (KCC), root mean square error (RMSE) and mean absolute error (MAE) were performed on the data collected to determine the limits of agreement, level of correlation, and error rate. The experimental setup and data acquisition based on three assumptions are shown in Fig. 7.

A. NOISE ARTIFACTS

In this section, the immunity of the proposed system to noise artifact has been investigated when different noise conditions mix with the cardiopulmonary signal. Some examples of the cardiopulmonary signal obtained from different conditions are shown in Fig. 8.

The video data obtained from the UAV in G_1 were used in the experimental results, where each subject was instructed to carry out a stationary and non-stationary scenario with normal breathing during videoing.

1) MEASUREMENTS OF CARDIAC ACTIVITY

a: STATIONARY SCENARIO

In the stationary scenario, all subjects were instructed to stay as still as possible in front of the UAV, not to talk and to not make any facial expressions while breathing normally.

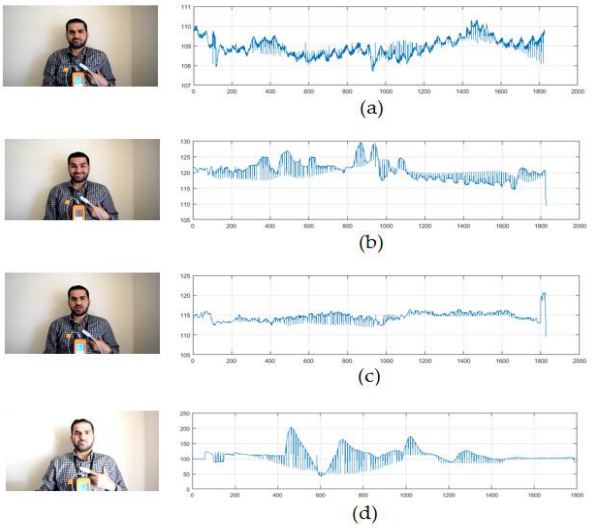

FIGURE 8. The cardiopulmonary signals for subject's face in different conditions (a) stationary subject (b) different facial expressions, (c) talking, and (d) different lighting conditions.

The Bland-Altman method and correlation coefficients of the stationary scenario under noise artifacts assumption for HR results obtained by the reference measurement method and HR predicted by the color-based method, motion-based method, ICA-based method, and PCA-based method are shown in Fig. 9.

Using the data obtained from the proposed system based on skin color analysis as shown in Fig. 9 (a), the mean bias, lower and upper limit of agreement (agreement range) between the predicted HR and the measured HR were 0.069, −0.52 to +0.66 beats/min with correlation coefficients of 0.9995, 0.999 and 0.9869 for the PCC, SCC and KCC respectively. For the data with the head motion analysis as shown in Fig. 9 (b), the mean bias and agreement range were 0.089, −1.2 to +1.4 beats/min with PCC of 0.9978, SCC of 0.9963 and KCC of 0.9657. When ICA was evaluated [Fig. 9 (c)], the mean bias and agreement range were 0.3 and −2.3 to +2.9 beats/min with correlation coefficients of 0.9902, 0.9883 and 0.9302, whereas, the statistical values were 0.34 beats/min of a mean bias, −3.2 to +3.8 beats/min of agreement range, 0.9825 of PCC, 0.981 of SCC and 0.9023 of KCC [Fig. 9 (d)] when PCA-based method was used instead.

b: NON-STATIONARY SCENARIO

In this scenario, all subjects were instructed to move and rotate their faces, talk, blink and make some facial expressions. The limits of agreement of all remote measuring systems are shown in Fig. 10.

As shown in Fig. 10 (a), the Bland-Altman statistics based on skin color analysis were 0.14 beats/min of a mean bias and −1.3 to +1.6 beats/min of agreement range with PCC of 0.9971, SCC of 0.9952 and KCC of 0.9581. The Bland-Altman plot for the HR measurements from the head motion as shown Fig. 10 (b) showed that the mean bias and agreement range were 0.2 and −1.9 to +2.3 beats/min with

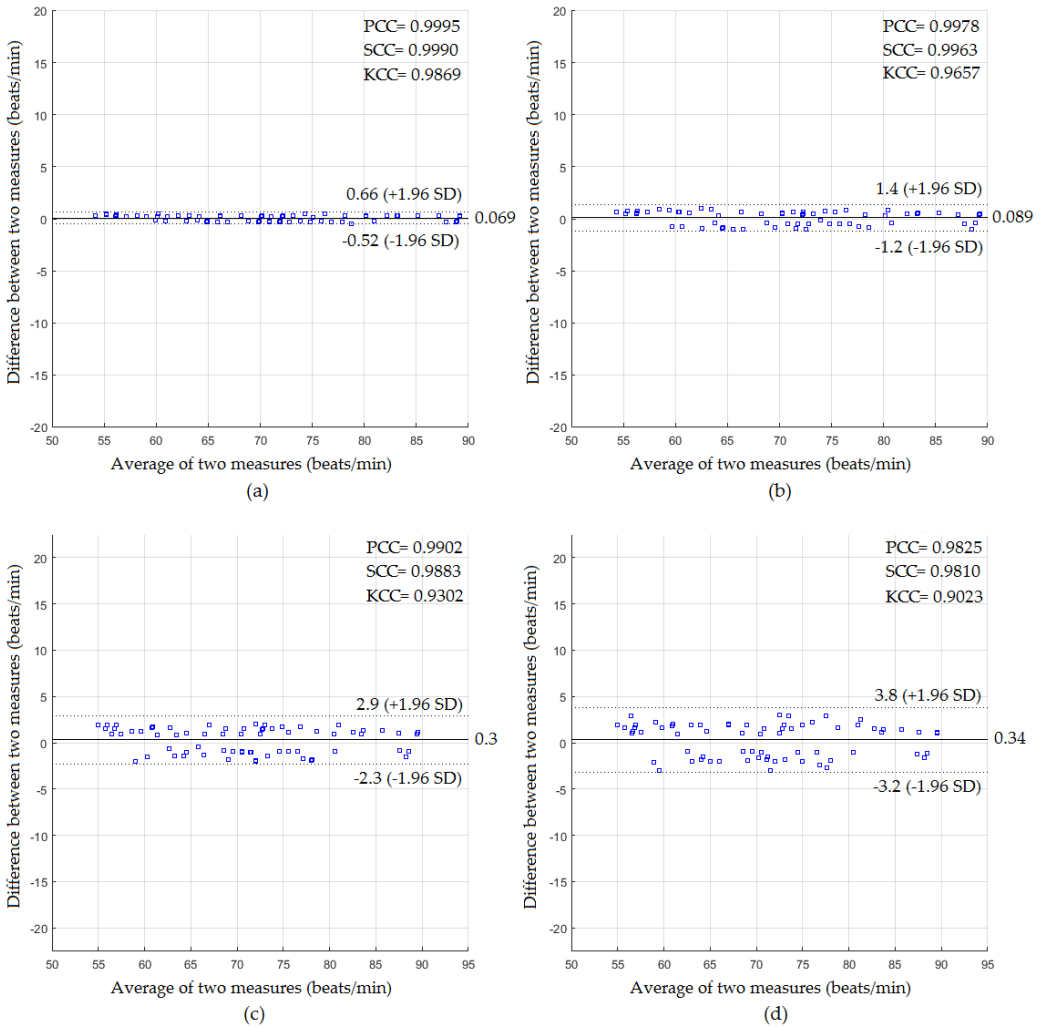

FIGURE 9. Bland-Altman plots for HR measurements in the stationary scenario under the noise artifact assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

PCC of 0.9939, SCC of 0.9915 and KCC of 0.9422. Using ICA [see Fig. 10 (c)], the mean bias and agreement range were 0.41 and −4 to +4.9 beats/min respectively with PCC of 0.9717, SCC of 0.9701 and KCC of 0.8705, whereas when PCA-based method was used instead, the Bland-Altman statistics were 0.41 beats/min, −4.6 to +5.4 beats/min with PCC of 0.9637, SCC of 0.9593 and KCC of 0.8429 [see Fig. 10 (d)].

The error rates of the HR measurements were evaluated using RMSE and MAE for both proposed scenarios as shown in Fig. 11.

2) MEASUREMENTS OF RESPIRATORY ACTIVITY

a: STATIONARY SCENARIO

In the stationary scenario, the statistical values based on Bland-Altman method for all remote measuring systems against the reference method are shown in Fig. 12.

Fig. 12 shows the statistical values of the Bland-Altman plots for the RR measurements in the stationary scenario. The Bland-Altman plot shown in Fig. 12 (a) revealed a strong agreement when skin color analysis was used. The mean bias and agreement range were 0.068 and −0.29 to +0.42 breaths/min with PCC of 0.999, SCC of 0.9963 and KCC of 0.9675. The Bland-Altman plot based on head motion analysis shown in Fig. 12 (b) showed statistical values of 0.15 breaths/min, −0.83 to +1.1 breaths/min, 0.9925, 0.9916 and 0.9438 for the mean bias, agreement range, PCC, SCC and KCC respectively. Based on ICA as shown in Fig. 12 (c), the mean bias and agreement range were 0.45, −1.9 to +2.8 breaths/min, while the PCC, SCC and KCC were 0.9578, 0.9642 and 0.8637 respectively. Based on PCA as shown in Fig. 12 (d), the mean bias and agreement range were 0.58, −2.6 to +3.8 breaths/min, while the PCC, SCC and KCC were 0.9248, 0.937 and 0.8041 respectively.

b: NON-STATIONARY SCENARIO

In this scenario, the statistical values of the RR measurements for all remote measuring systems are shown in Fig. 13.

Fig. 13 (a) indicated 0.13 breaths/min of a mean bias, −0.67 to +0.93 breaths/min of agreement range with

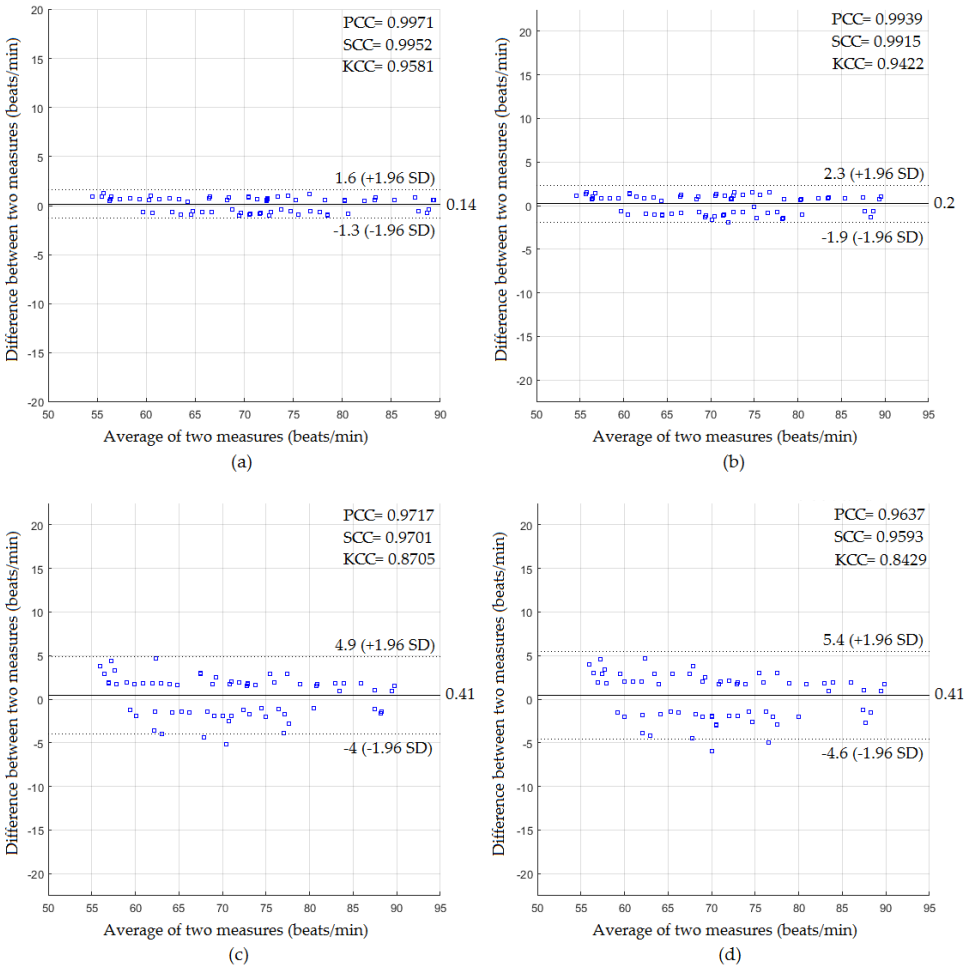

FIGURE 10. Bland-Altman plots for HR measurements in the non-stationary scenario under the noise artifact assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

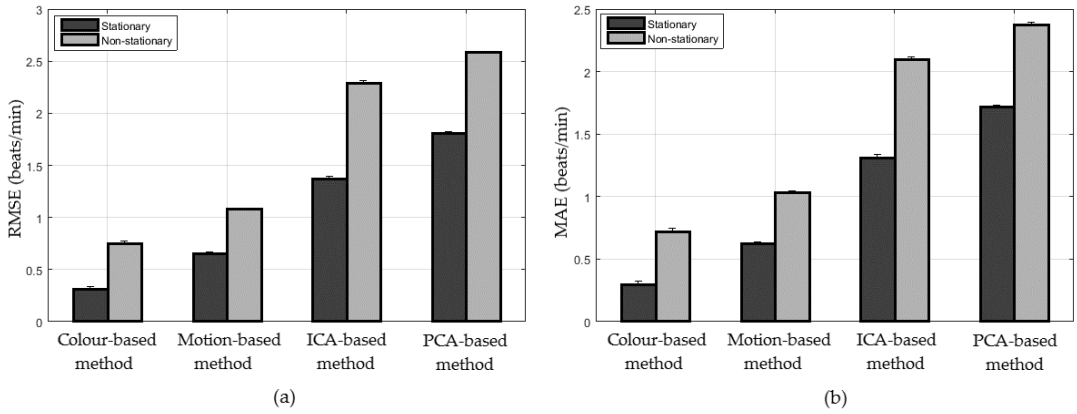

FIGURE 11. Error rates of various HR measuring methods under noise artifact assumption using (a) RMSE, (b) MAE.

0.9949 of PCC, 0.9942 of SCC and 0.9557 of KCC, whereas, Fig. 13 (b) showed 0.18 breaths/min of a mean bias, −0.92 to +1.3 breaths/min of agreement range with 0.995 of PCC, 0.9876 of SCC and 0.9243 of KCC. Based on ICA as shown in Fig. 13 (c), the statistical values were 0.71 breaths/min of a mean bias, −2.8 to +4.2 breaths/min of agreement range, 0.9093 of PCC, 0.921 of SCC and 0.7748 of KCC, whereas, when PCA was used, the statistical values were

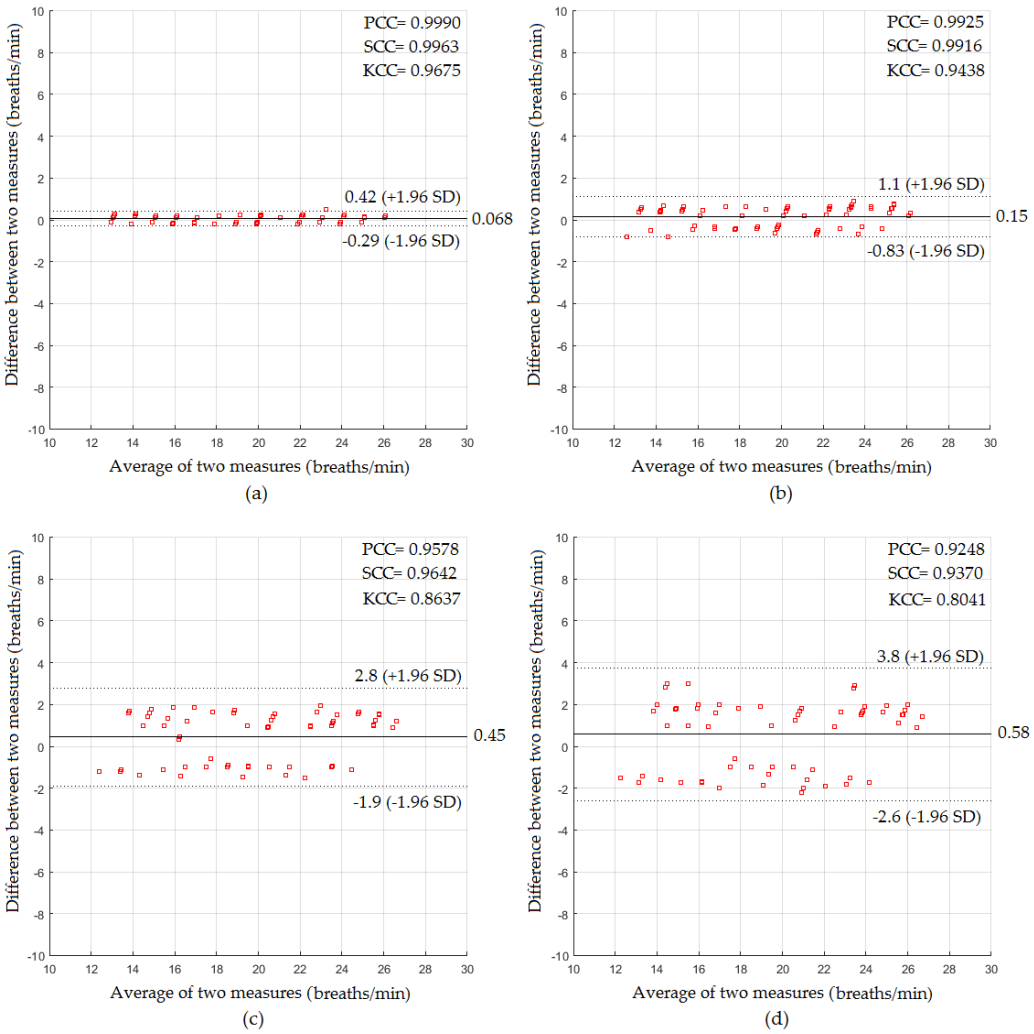

FIGURE 12. Bland-Altman plots for RR measurements in the stationary scenario under the noise artifact assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

1 breaths/min, −2.9 to +4.9 breaths/min of agreement range, 0.8951 of PCC, 0.9065 of SCC and 0.7447 of KCC as shown in Fig. 13 (d).

The error rates of the RR measurements were also evaluated using RMSE and MAE as shown in Fig. 14.

B. MULTIPLE DETECTION

In this section, the performance of the proposed system to detect the cardiopulmonary signal from multiple subjects (up to six people) is investigated. The video data from G_2 were used in the experimental results, where all subjects in the video session were instructed to perform two scenarios. The first scenario was stationary, where all subjects were instructed to stay as still as possible in front of the camera, not to talk and not to make any facial expressions while breathing normally. The second scenario was non-stationary, where all subjects were instructed to move and rotate their faces, talk, blink and make some facial expressions while breathing normally.

1) MEASUREMENTS OF CARDIAC ACTIVITY

a: STATIONARY SCENARIO

In the stationary scenario, the statistical values based on the Bland-Altman method of the HR measurements for all remote measuring systems against the reference method is shown in Fig. 15.

The statistical values of the color-based method shown in Fig. 15 (a) revealed a mean bias of 0.082 beats/min, agreement range of −0.51 to +0.68 beats/min, PCC of 0.9994, SCC of 0.9986 and KCC of 0.9864, whereas, the statistical values of the motion-based method shown in Fig. 15 (b) revealed a mean bias of 0.11 beats/min, agreement range of −1.3 to +1.5 beats/min, PCC of 0.9967, SCC of 0.9958 and KCC of 0.9632. Using ICA-based method as shown in Fig. 15 (c), the statistical values were 0.62 beats/min of a mean bias, −4.6 to +5.8 beats/min of agreement range, 0.949 of PCC, 0.9657 of SCC and 0.8655 of KCC, whereas, when PCA-based method was used, the statistical values were 0.77 beats/min, −6 to +7.5 beats/min of agreement range,

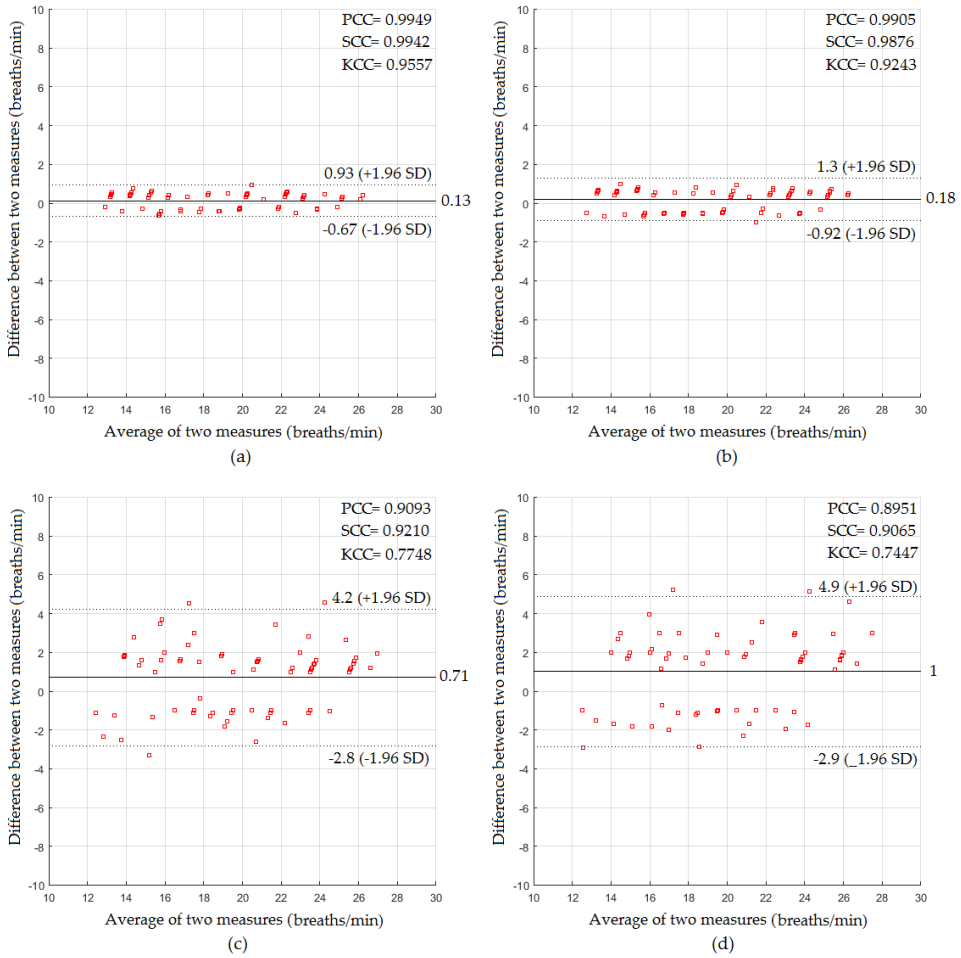

FIGURE 13. Bland-Altman plots for RR measurements in the non-stationary scenario under the noise artifact assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

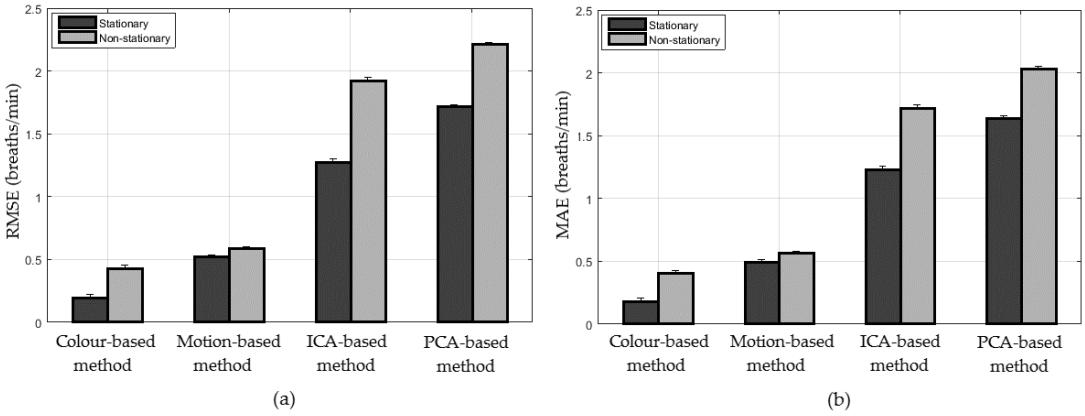

FIGURE 14. Error rates of various RR measuring methods under noise artifact assumption using (a) RMSE, (b) MAE.

0.913 of PCC, 0.9355 of SCC and 0.7971 of KCC as shown in Fig. 15 (d).

b: NON-STATIONARY SCENARIO

In this scenario, the agreement of the HR measurements is shown in Fig. 16.

The Bland-Altman plot [(Fig. 16 (a)] showed the statistical values were 0.29 beats/min, −1.9 to +2.5 beats/min, 0.9913, 0.99 and 0.9334 for the mean bias, agreement range, PCC, SCC and KCC respectively when the color-based method was used, whereas, Fig. 16 (b) showed that the statistical values of the motion-based method were 0.52 beats/min,

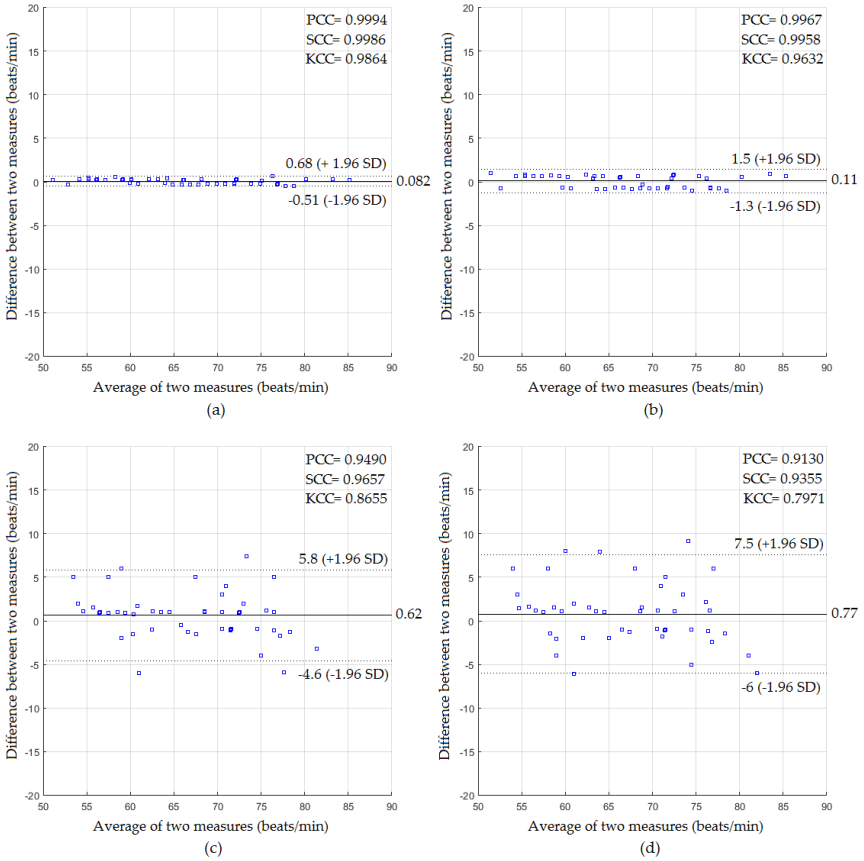

FIGURE 15. Bland-Altman plots for HR measurements in the stationary scenario under the multiple detection assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

−2.8 to +3.8 beats/min, 0.98, 0.9809 and 0.9039 respectively. The statistical values based on ICA were 1.8 beats/min, −6.4 to +9.9 beats/min, 0.8729, 0.8905 and 0.7116 as shown in Fig. 16 (c), whereas they were 1.9 beats/min, −7.7 to +11 beats/min, 0.825, 0.8484 and 0.6557 based on PCA as shown in Fig. 16 (d).

A performance comparison of the HR measurements using the error rates (RMSE and MAE) is shown in Fig. 17.

2) MEASUREMENTS OF RESPIRATORY ACTIVITY

a: STATIONARY SCENARIO

In the stationary scenario, the statistical values based on the Bland-Altman method of the RR measurements are shown in Fig. 18.

The statistical values of the color-based method shown in Fig. 18 (a) revealed a mean bias of 0.087 breaths/min, agreement range of −0.52 to +0.69 breaths/min, PCC of 0.9942, SCC of 0.9918 and KCC of 0.9501, whereas, the statistical values of the motion-based method shown in Fig. 18 (b) revealed a mean bias of 0.18 breaths/min, agreement range of −1.1 to +1.4 breaths/min, PCC of 0.9775, SCC of 0.9566 and KCC of 0.8604. Using ICA-based method as shown in Fig. 18 (c), the statistical values were 0.23 breaths/min of a mean bias, −3 to +3.5 breaths/min of agreement range, 0.8724 of PCC, 0.8236 of SCC and 0.6782 of KCC, whereas, when PCA-based method was used, the statistical values were 0.39 breaths/min, −4.4 to 5.2 breaths/min, 0.7696, 0.6769 and 0.5325 for the mean bias, agreement range, PCC, SCC and KCC respectively, as shown in Fig. 18 (d).

b: NON-STATIONARY SCENARIO

In this scenario, the statistical values based on the Bland-Altman method of the RR measurements are shown in Fig. 19.

Fig. 19 (a) showed a mean bias of 0.18 breaths/min, agreement range of −2.1 to +2.4 breaths/min, PCC of 0.9383, SCC of 0.9062 and KCC of 0.7768, whereas, Fig. 19 (b) showed a mean bias of 0.21 breaths/min, agreement range of −2.6 to +3 breaths/min, PCC of 0.9072, SCC of 0.8651 and KCC of 0.7204. Based on ICA and PCA, the statistical values were 0.58 breaths/min, −4.5 to +5.6 breaths/min, 0.6908, 0.6768 and 0.5291 for the ICA-based method as shown in Fig. 19 (c) and 0.7 breaths/min, −5.8 to +7.1 breaths/min, 0.5882, 0.5434 and 0.4121 for the PCA-based method as shown in Fig. 19 (d).

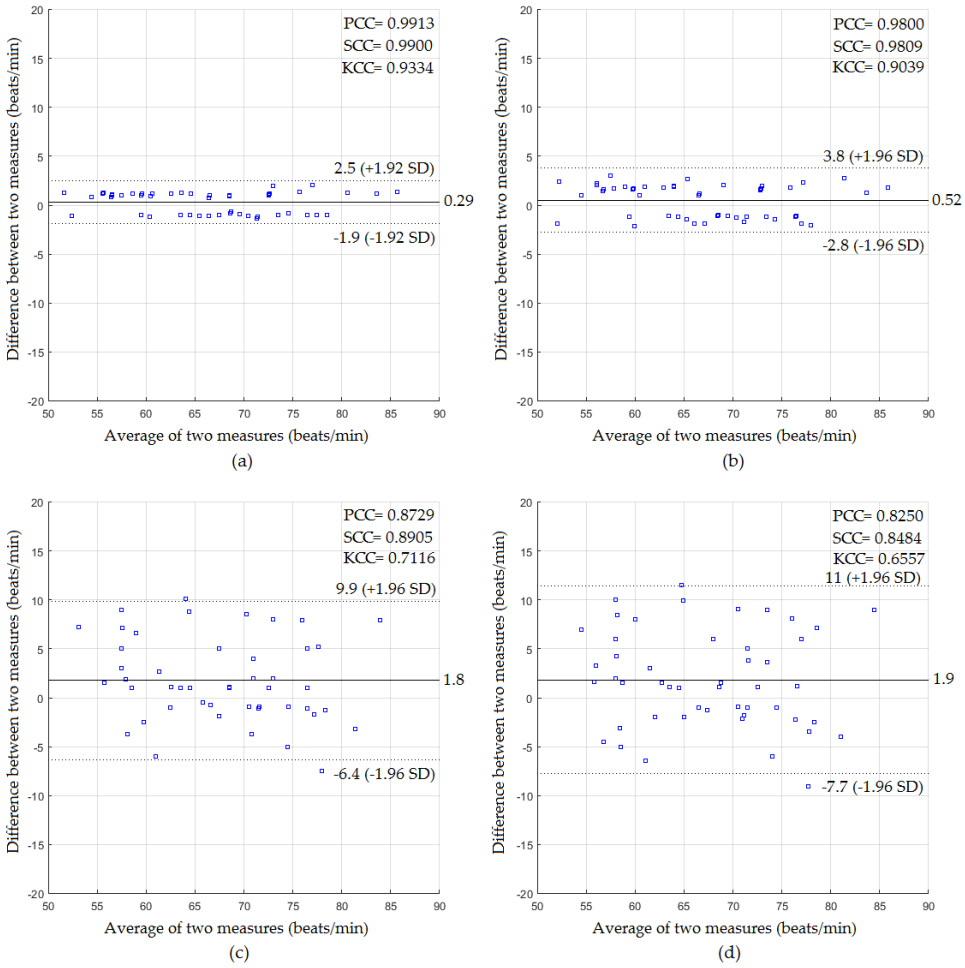

FIGURE 16. Bland-Altman plots for HR measurements in the non-stationary scenario under the multiple detection assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

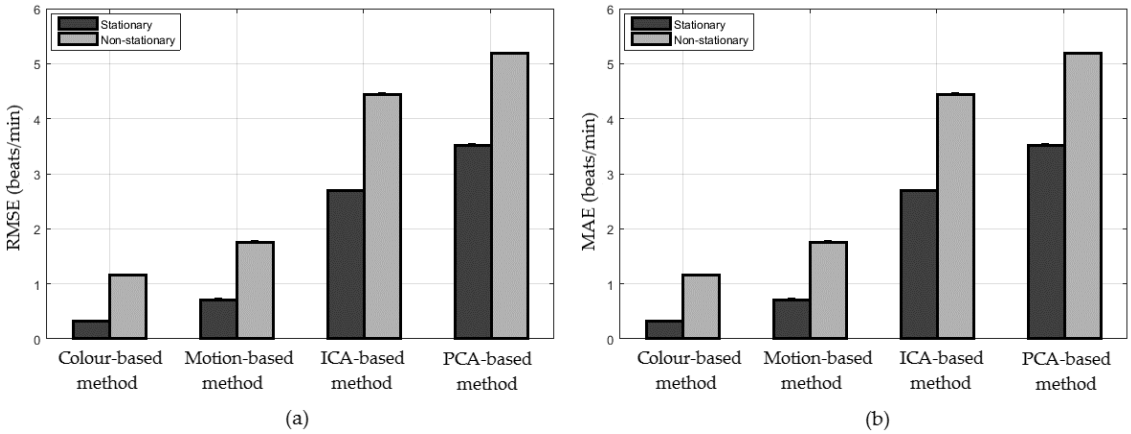

FIGURE 17. Error rates of various HR measuring methods under multiple detection assumption using (a) RMSE, (b) MAE.

The error rates of the RR measurements for the multiple detection were also investigated using RMSE and MAE as shown in Fig. 20.

C. LONG-DISTANCE DETECTION

In this section, the performance of the proposed system to detect the cardiopulmonary signal at long-distances of up

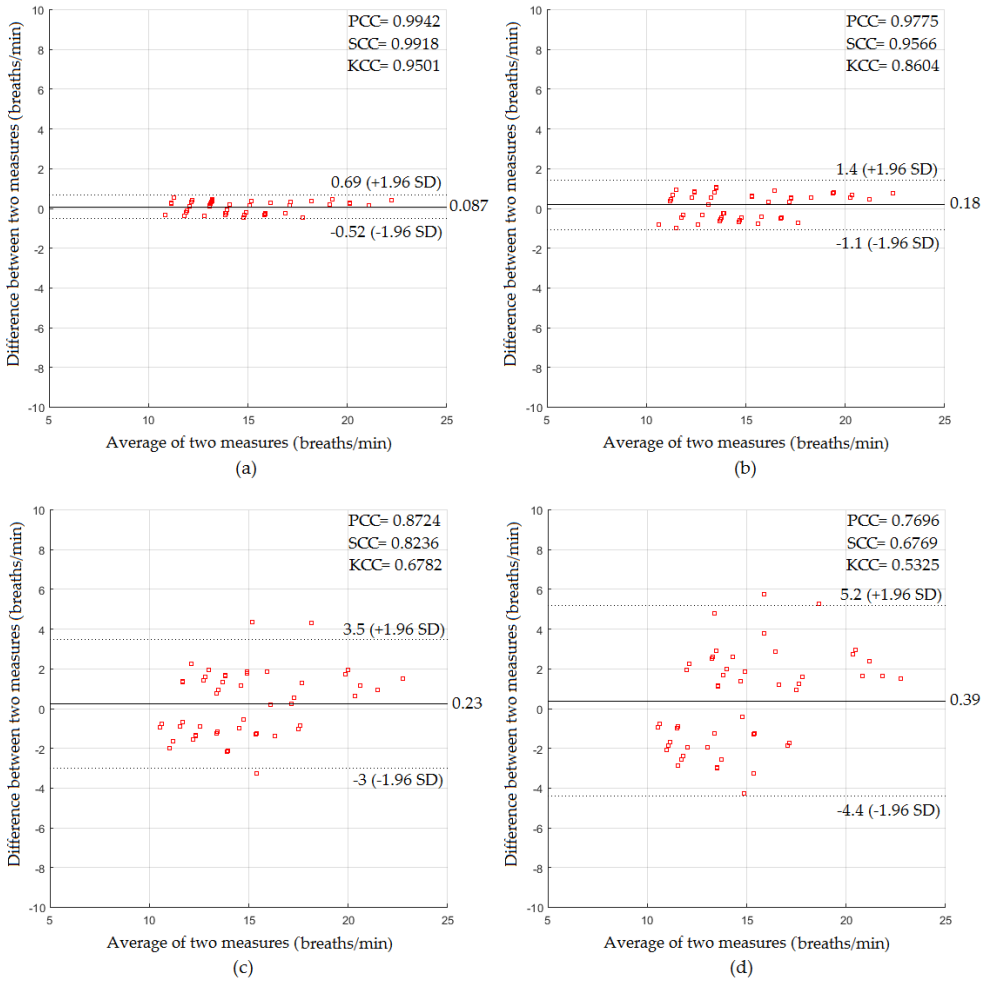

FIGURE 18. Bland-Altman plots for RR measurements in the stationary scenario under the multiple detection assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

to 60 m is investigated. The video data from G_3 were used in the experimental results and divided into two scenarios: the stationary scenario where all subjects were instructed to stay still, not to talk and not to make any facial expressions, and the non-stationary scenario where all subjects were instructed to move naturally toward the camera for approximately 10 m at a walking pace and instructed to rotate their faces while walking.

1) MEASUREMENTS OF CARDIAC ACTIVITY

a: STATIONARY SCENARIO

In the stationary scenario, the statistical values of the HR measurements from the remote measuring systems and the reference methods are shown in Fig. 21.

As shown in Fig. 21 (a), the proposed system based on skin color method for the stationary scenario had a mean bias of 0.56 beats/min, agreement range of -0.81 to $+1.9$ beats/min and correlation coefficients of 0.9973, 0.9964 and 0.9706 for the PCC, SCC and KCC respectively, whereas, when the motion-based method was used, the proposed system had a mean bias of 0.65 beats/min, agreement range of -1.2 and $+2.5$ beats/min and correlation coefficients of 0.9953, 0.9924 and 0.9478 for the PCC, SCC and KCC respectively as shown in Fig. 21 (b). When the ICA-based method was alternatively used on the acquired data, the statistical values were 0.85 beats/min (mean bias), -4 to $+5.7$ beats/min (agreement range) and coefficients of 0.964, 0.9669 and 0.8664 as shown in Fig. 21 (c), whereas the statistical values were 0.97 beats/min (mean bias), -4.9 to $+6.8$ beats/min (agreement range) and coefficients of 0.948, 0.9516 and 0.8252 when PCA-based method was used instead as shown in Fig. 21 (d).

b: NON-STATIONARY SCENARIO

In this scenario, the statistical values of the HR measurements are shown in Fig. 22.

As shown in Fig. 22 (a), the proposed system based on the skin color method had a mean bias of 0.81 beats/min, agreement range of -1.4 to $+3$ beats/min, yielding correlation coefficients of 0.9929, 0.9893 and 0.9374 for

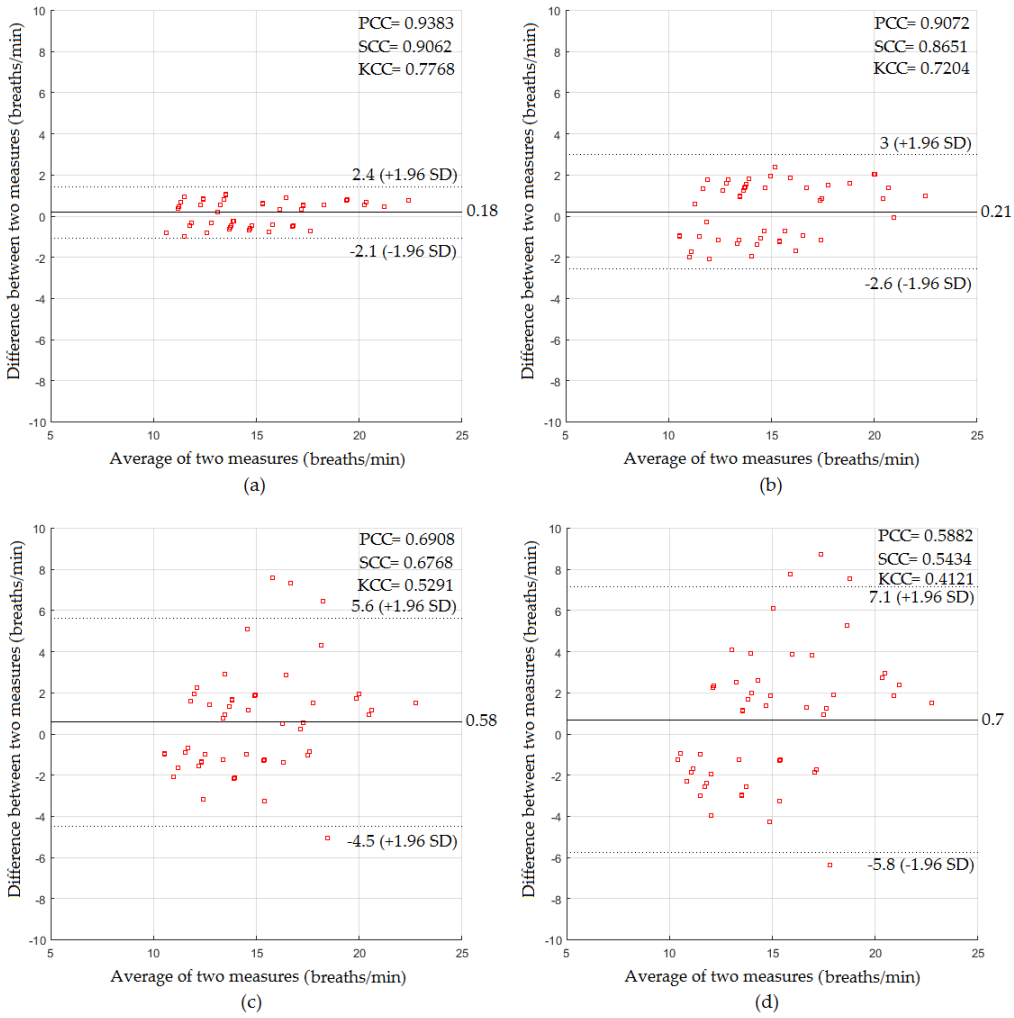

FIGURE 19. Bland-Altman plots for RR measurements in the non-stationary scenario under the multiple detection assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

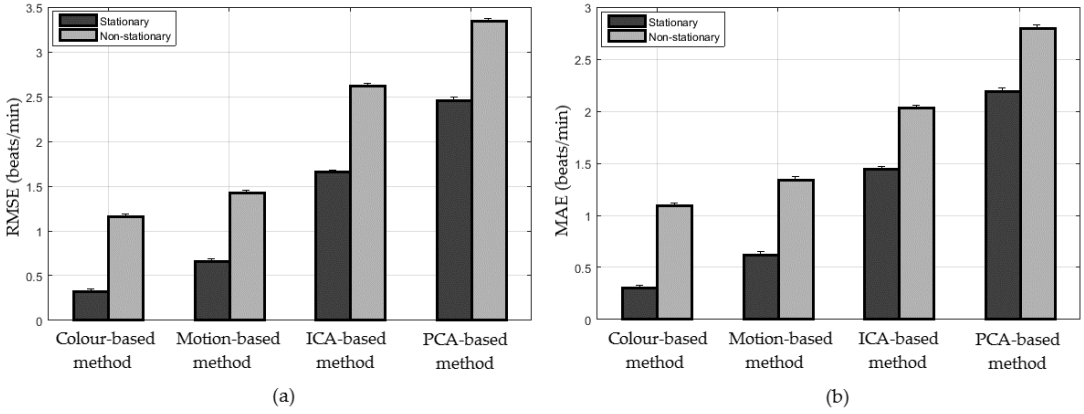

FIGURE 20. Error rates of various RR measuring methods under multiple detection assumption using (a) RMSE, (b) MAE.

the PCC, SCC and KCC respectively, whereas, when the motion-based method was used, the proposed system had a mean bias of 1.1 beats/min, agreement range of −1.8 to +3.9 beats/min, yielding correlation coefficients of 0.9889, 0.985 and 0.9212 as shown in Fig. 22 (b). When the ICA-based method was alternatively used on the acquired

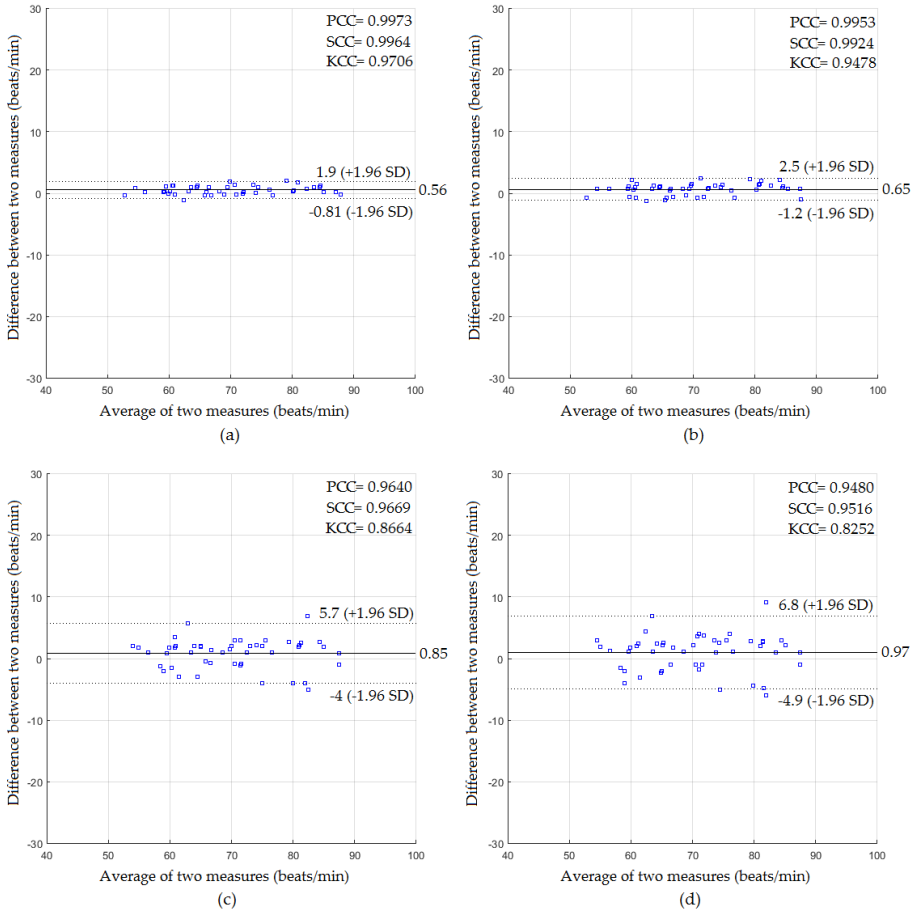

FIGURE 21. Bland-Altman plots for HR measurements in the stationary scenario under the long-distance assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

data, the statistical values were 1.6 beats/min (mean bias), −6.7 to +10 beats/min (agreement range) and coefficients of 0.8873, 0.8966 and 0.7398 as shown in Fig. 22 (c), whereas, the statistical values were 1.9 beats/min (mean bias), −7.7 to +11 beats/min (agreement range), and coefficients of 0.8531, 0.8709 and 0.6903 when PCA-based method was used instead, as shown in Fig. 22 (d).

The error rates of the HR measurements obtained from long-distances in both scenarios were performed and evaluated using RMSE and MAE as shown in Fig. 23.

2) MEASUREMENTS OF RESPIRATORY ACTIVITY

a: STATIONARY SCENARIO

In the stationary scenario, the statistical values of the RR measurements from the remote measuring methods and the reference method are shown in Fig. 24.

As shown in Fig. 24 (a), the mean bias and agreement range between the color-based method and the reference method were 0.25 and −0.97 to 1.5 breaths/min with PCC of 0.9827, SCC of 0.9851 and KCC of 0.9238, whereas, Fig. 24 (b) had a mean bias and agreement range of 0.43, −1.3 to 2.2 breaths/min respectively with PCC of 0.9652, SCC of 0.9601 and KCC of 0.8649 when the motion-based method was used. Using ICA-based method, a mean bias and agreement range were 0.82 and −2.6 to +4.3 breaths/min with PCC of 0.8809, SCC of 0.8616 and KCC of 0.7264 as shown in Fig. 24 (c), whereas, with the PCA-based method, the statistical values were 1 breaths/min of mean bias, −3.5 to +5.6 breaths/min of agreement range, 0.8126 of PCC, 0.7826 of SCC and 0.6358 of KCC as shown in Fig. 24 (d).

b: NON-STATIONARY SCENARIO

In this scenario, the statistical values of the RR measurements are shown in Fig. 25.

As shown in Fig. 25 (a), the mean bias and agreement range between the color-based method and the reference method were 0.42 and −1.7 to +2.6 breaths/min with PCC of 0.9494, SCC of 0.9423 and KCC of 0.8271, whereas Fig. 25 (b) had a mean bias and agreement range of 0.7 and −1.9 to +3.3 breaths/min respectively with PCC of 0.9268, SCC of 0.9094 and KCC of 0.7745 when the motion-based method was used. Using ICA-based method, a mean bias and agreement range were 0.78, and −4.7 to +6.3 breaths/min with

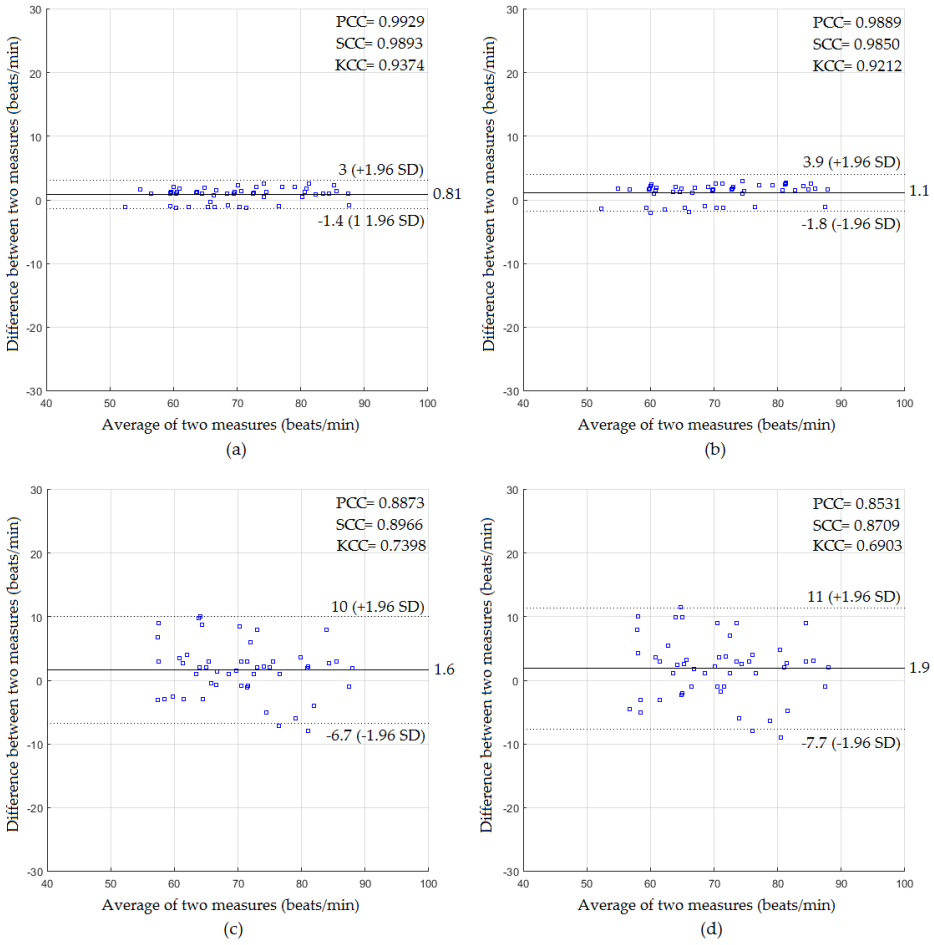

FIGURE 22. Bland-Altman plots for HR measurements in the non-stationary scenario under the long-distance assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

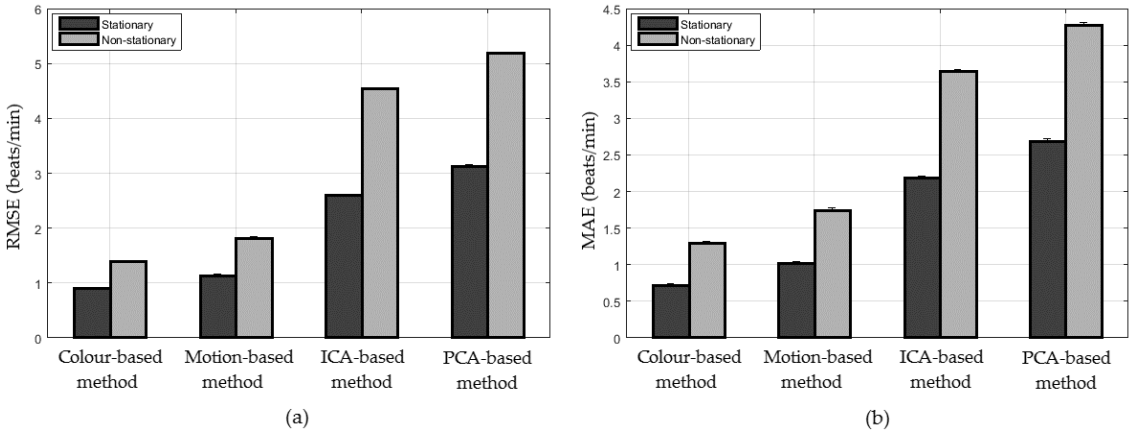

FIGURE 23. Error rates of various HR measuring methods under long-distance assumption using (a) RMSE, (b) MAE.

PCC of 0.7052, SCC of 0.7057 and KCC of 0.5567 as shown in Fig. 25 (c), whereas, with the PCA-based method, the statistical values were 0.98 breaths/min (mean bias), −5.6 to +7.5 breaths/min (agreement range), 0.6171, 0.571 and 0.4561 for PCC, SCC and KCC respectively as shown in Fig. 25 (d).

The error rates of the RR measurements obtained from long-distances in both scenarios were performed

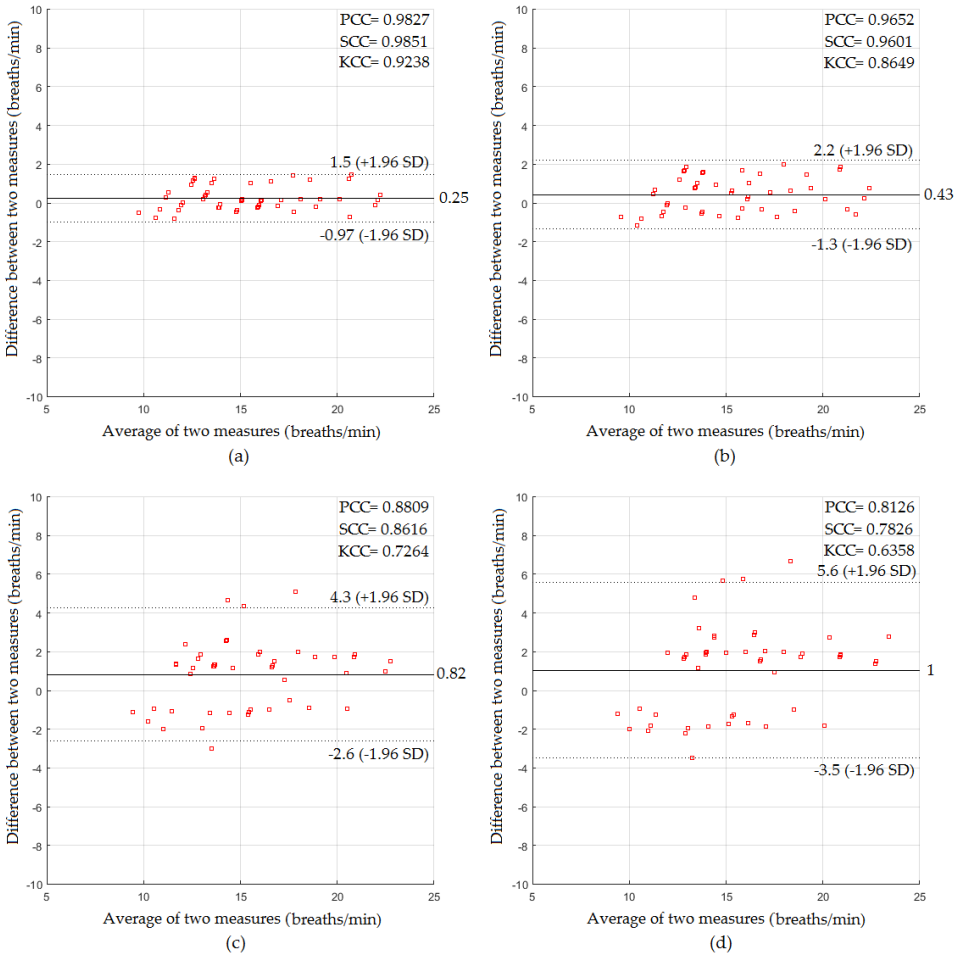

FIGURE 24. Bland-Altman plots for RR measurements in the stationary scenario under the long-distance assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

and evaluated using RMSE and MAE as shown in Fig. 26.

IV. DISCUSSION

The challenges of removing noise artifacts from the cardiopulmonary signal, detecting the cardiopulmonary signal from multiple subjects and detecting the cardiopulmonary signal from long-distance have been addressed in this study. The proposed system in this paper also relied on the methods based on color and motion analysis to extract cardiopulmonary signal when noise and motion artifacts sensitivity cannot be solved by a single method. The proposed system is feasible and has high tolerance of noise artifacts for both methods (color and motion) when the subjects were in stationary and non-stationary scenarios. The estimated HR and RR results obtained from the proposed system based on color analysis under three assumptions: noise artifact, multiple detection and long-distance detection had slightly better statistics outcomes (PCC, SCC, KCC, RMSE and MAE) than when the motion-based method was used instead in comparison with those obtained from conventional reference methods in both stationary and non-stationary scenarios. The proposed system based on both analysis methods also outperformed the traditional measurement methods (ICA and PCA) in terms of agreement, correlation, noise level and computational time.

For the noise artifacts assumption, the proposed system with skin color analysis under the stationary scenario showed an excellent agreement with the reference methods with the statistical values (PCC = 0.9995, SCC = 0.999, KCC = 0.9869, RMSE = 0.31 beats/min and MAE = 0.29 beats/min for HR measurements and PCC = 0.999, SCC = 0.9963, KCC = 0.9675, RMSE = 0.19 breaths/min and MAE = 0.18 breaths/min for RR measurements). The proposed system with motion analysis could also extract the cardiopulmonary signal with a very good agreement (PCC = 0.9978, SCC = 0.9963, KCC = 0.9657, RMSE = 0.65 beats/min and MAE = 0.62 beats/min for HR measurements and PCC = 0.9925, SCC = 0.9916, KCC = 0.9438, RMSE = 0.52 breaths/min and MAE = 0.49 breaths/min for

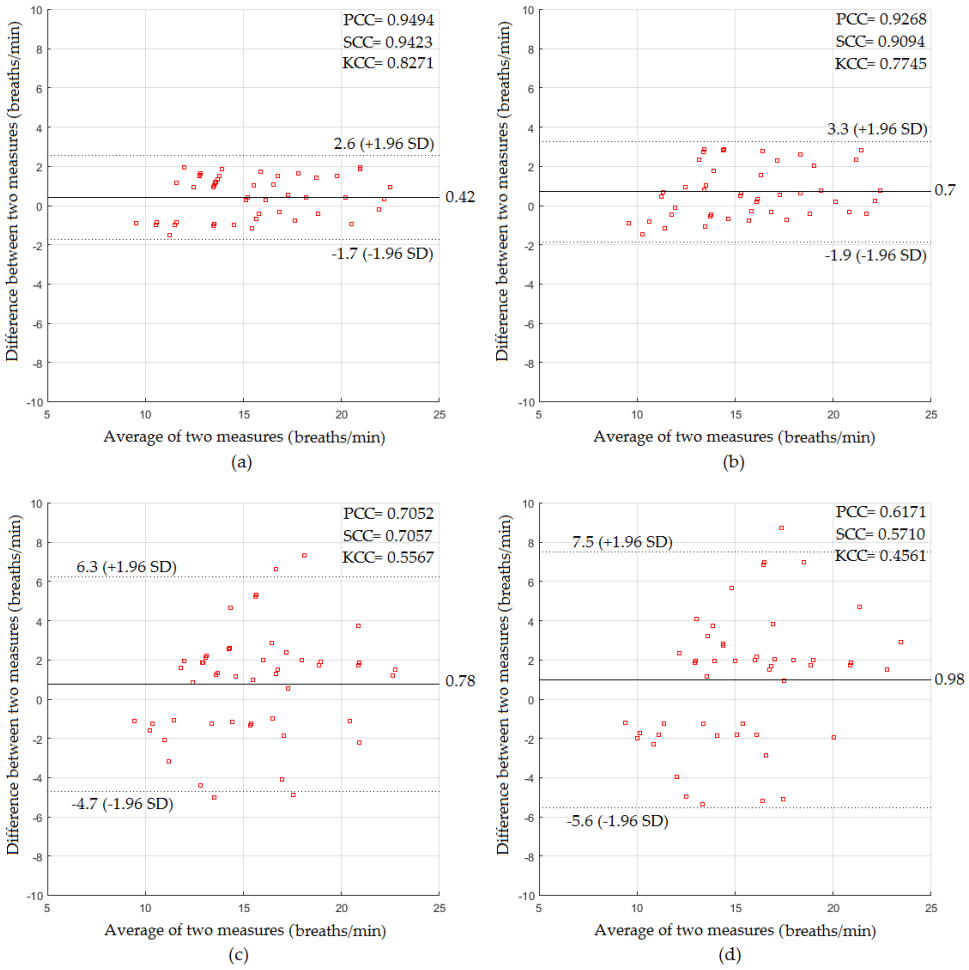

FIGURE 25. Bland-Altman plots for RR measurements in the non-stationary scenario under the long-distance assumption using (a) color-based method, (b) motion-based method, (c) ICA-based method, and (d) PCA-based method.

RR measurements), which were better statistical outcomes than when ICA-based method and PCA-based method were used instead. ICA under the stationary scenario had statistical outcomes (PCC = 0.9902, SCC = 0.9883, KCC = 0.9302, RMSE = 0.37 beats/min and MAE = 1.31 beats/min for HR measurements and PCC = 0.9578, SCC = 0.9642, KCC = 0.8637, RMSE = 1.27 breaths/min and MAE = 1.23 breaths/min for RR measurements), whereas, PCA-based method had statistical outcomes (PCC = 0.9825, SCC = 0.981, KCC = 0.9023, RMSE = 1.8 beats/min and MAE = 1.71 beats/min for HR measurements and PCC = 0.9248, SCC = 0.937, KCC = 0.8041, RMSE = 1.71 breaths/min and MAE = 1.64 breaths/min for RR measurements). The estimated results under the non-stationary scenario also exhibited very good statistical values (PCC = 0.9971, SCC = 0.9952, KCC = 0.9581, RMSE = 0.74 beats/min and MAE = 0.72 beats/min for HR measurements and PCC = 0.9949, SCC = 0.9942, KCC = 0.9557, RMSE = 0.42 breaths/min and MAE = 0.4 breaths/min for RR measurements) with respect to skin color analysis and statistical outcomes (PCC = 0.9939, SCC = 0.9915, KCC = 0.9422, RMSE = 1.07 beats/min and MAE = 1.03 beats/min for HR measurements and PCC = 0.9905, SCC = 0.9876, KCC = 0.9243, RMSE = 0.58 breaths/min and MAE = 0.56 breaths/min for RR measurements) with respect to head motion analysis, whereas, ICA-based method and PCA-based method under non-stationary scenario might fail in extracting HR and RR under the noise artifacts assumption with lower correlation levels and higher RMSE and MAE.

For the multiple detection assumption, the proposed system with skin color analysis under the stationary scenario also showed a very good agreement with the statistical values (PCC = 0.9994, SCC = 0.9986, KCC = 0.9864, RMSE = 0.31 beats/min and MAE = 0.29 beats/min for HR measurements and PCC = 0.9942, SCC = 0.9918, KCC = 0.9501, RMSE = 0.32 breaths/min and MAE = 0.3 breaths/min for RR measurements). The proposed system with the head motion analysis could also extract the cardiopulmonary signal with good agreement (PCC = 0.9967, SCC = 0.9958, KCC = 0.9632, RMSE = 0.7 beats/min and

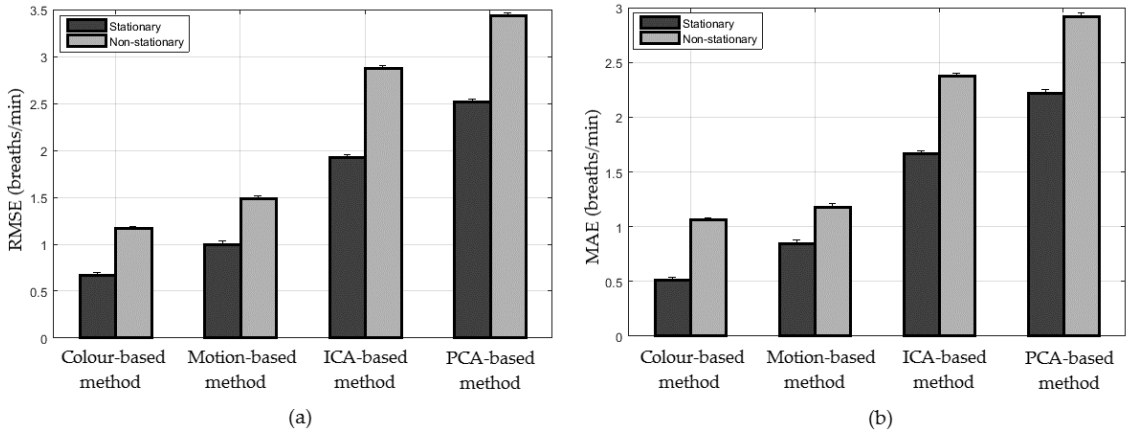

FIGURE 26. Error rates of various RR measuring methods under long-distance assumption using (a) RMSE, (b) MAE.

MAE = 0.69 beats/min for HR measurements and PCC = 0.9775, SCC = 0.9566, KCC = 0.8604, RMSE = 0.65 breaths/min and MAE = 0.61 breaths/min for RR measurements), which were better statistical outcomes than both ICA and PCA. ICA-based method under the stationary scenario had statistical outcomes (PCC = 0.949, SCC = 0.9657, KCC = 0.8655, RMSE = 2.69 beats/min and MAE = 2.02 beats/min for HR measurements and PCC = 0.8724, SCC = 0.8236, KCC = 0.6782, RMSE = 1.65 breaths/min and MAE = 1.44 breaths/min for RR measurements), whereas, PCA-based method had statistical outcomes (PCC = 0.913, SCC = 0.9355, KCC = 0.7971, RMSE = 3.51 beats/min and MAE = 2.76 beats/min for HR measurements and PCC = 0.7696, SCC = 0.6769, KCC = 0.5325, RMSE = 2.45 breaths/min and MAE = 2.18 breaths/min for RR measurements). The estimated results under the non-stationary scenario also exhibited good and acceptable statistical values (PCC = 0.9913, SCC = 0.99, KCC = 0.9334, RMSE = 1.14 beats/min and MAE = 1.12 beats/min for HR measurements and PCC = 0.9383, SCC = 0.9062, KCC = 0.7768, RMSE = 1.16 breaths/min and MAE = 1.09 breaths/min for RR measurements) with respect to skin color analysis and statistical outcomes (PCC = 0.98, SCC = 0.9809, KCC = 0.9039, RMSE = 1.74 beats/min and MAE = 1.68 beats/min for HR measurements and PCC = 0.9072, SCC = 0.8651, KCC = 0.7204, RMSE = 1.42 breaths/min and MAE = 1.33 breaths/min for RR measurements) with respect to the head motion analysis, whereas, ICA-based method and PCA-based method under the non-stationary scenario have failed in extracting HR and RR from multiple subjects with weaker agreement, lower correlation coefficients and higher RMSE and MAE.

For the long-distance detection assumption, the proposed system with skin color analysis under the stationary scenario showed a very good agreement with the statistical values (PCC = 0.9973, SCC = 0.9964, KCC = 0.9706, RMSE = 0.89 beats/min and MAE = 0.71 beats/min for HR measurements and PCC = 0.9827, SCC = 0.9851, KCC = 0.9238, RMSE = 0.67 breaths/min and MAE = 0.51 breaths/min for RR measurements). The proposed system with the head motion analysis could also extract the cardiopulmonary signal with good agreement (PCC = 0.9953, SCC = 0.9924, KCC = 0.9478, RMSE = 1.12 beats/min and MAE = 1.01 beats/min for HR measurements and PCC = 0.9652, SCC = 0.9601, KCC = 0.8649, RMSE = 0.99 breaths/min and MAE = 0.84 breaths/min for RR measurements), which outperformed the statistical outcomes obtained by the ICA-based method and the PCA-based method. ICA-based method under the stationary scenario had statistical outcomes (PCC = 0.964, SCC = 0.9669, KCC = 0.8664, RMSE = 2.59 beats/min and MAE = 2.18 beats/min for HR measurements and PCC = 0.8809, SCC = 08616, KCC = 0.7264, RMSE = 1.92 breaths/min and MAE = 1.66 breaths/min for RR measurements), whereas, the PCA-based method had lower statistical outcomes (PCC = 0.948, SCC = 0.9516, KCC = 0.8252, RMSE = 3.11 beats/min and MAE = 2.68 beats/min for HR measurements and PCC = 0.8126, SCC = 0.7826, KCC = 0.6358, RMSE = 2.51 breaths/min and MAE = 2.21 breaths/min for RR measurements). The estimated results under the non-stationary scenario also exhibited good and acceptable statistical outcomes (PCC = 0.9929, SCC = 0.9823, KCC = 0.9374, RMSE = 1.38 beats/min and MAE = 1.29 beats/min for HR measurements and PCC = 0.9494, SCC = 0.9423, KCC = 0.8271, RMSE = 1.16 breaths/min and MAE = 1.05 breaths/min for RR measurements) with respect to the skin color analysis and statistical outcomes (PCC = 0.9889, SCC = 0.985, KCC = 0.9212, RMSE = 1.8 beats/min and MAE = 1.74 beats/min for HR measurements and PCC = 0.9268, SCC = 0.9094, KCC = 0.7745, RMSE = 1.48 breaths/min and MAE = 1.18 breaths/min for RR measurements) with respect to the head motion analysis, whereas, ICA-based method and PCA-based method under the non-stationary scenario have failed in extracting HR and RR from long-distance with weaker agreement, lower correlation coefficients and higher RMSE and MAE.

The computational time of the proposed method to remove noise artifacts using both improved signal decomposition

technique and the BSS against the traditional measurement methods (ICA and PCA) has also been evaluated in this paper by using the MATLAB environment (2017b). The test was running on a PC with Microsoft Windows 10 (64 bits), i7-2.6 GHz of CPU and 8.00 GB of RAM. The mean computational time for the noise artifact removal method with 200 iterations for 30 second cardiopulmonary signal was 1.01 s, whereas it takes about 0.80 s and 0.75 s when ICA and PCA were used instead. The computational time is acceptable for noise artifact removal from the acquired signal, making the proposed method more suitable for real-time applications.

V. CONCLUSION

In conclusion, this paper explored the feasibility of extracting the cardiopulmonary signal from video data captured by different types of camera sensors (UAV and digital camera) under three assumptions: noise artifacts, multiple detection and long-distance. The proposed system used the magnification process to magnify the imperceptible variations caused by cardiopulmonary activity, including skin color and head motion to extract the cardiopulmonary signal followed by a new, robust and fast noise removal method based on a combination of improved signal decomposition technique and BSS-based CCA technique to remove the noise artifacts generating from the subject's movement, camera motion, skin tone and lighting conditions. The experimental evaluation of the proposed system has been conducted using three video data sources and showed a strong agreement, high correlation and low noise level compared with the reference measurements. Also, the proposed system outperformed the conventional measurement methods such as ICA-based method and PCA-based method in both the stationary and non-stationary scenarios.

REFERENCES

[1] K. Nakajima, A. Osa, and H. Miike, "A method for measuring respiration and physical activity in bed by optical flow analysis," in *Proc. 19th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Oct. 1997, pp. 2054–2057.

[2] K. Nakajima, Y. Matsumoto, and T. Tamura, "Development of real-time image sequence analysis for evaluating posture change and respiratory rate of a subject in bed," *Physiol. Meas.*, vol. 22, no. 3, pp. 21–28, Aug. 2001.

[3] J. E. Parra and G. da Costa, "Optical remote sensing of heartbeats," *Proc. SPIE*, vol. 4368, pp. 113–121, Aug. 2001.

[4] S. Wiesner and Z. Yaniv, "Monitoring patient respiration using a single optical camera," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, Aug. 2007, pp. 2740–2743.

[5] W. Verkruysse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Exp.*, vol. 16, no. 26, pp. 21434–21445, 2008.

[6] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7–11, Jan. 2011.

[7] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Exp.*, vol. 18, no. 10, pp. 10762–10774, 2010.

[8] M. Lewandowska, J. Rumiński, T. Kocejko, and J. Nowak, "Measuring pulse rate with a webcam—A non-contact method for evaluating cardiac activity," in *Proc. Federated Conf. Comput. Sci. Inf. Syst. (FedCSIS)*, 2011, pp. 405–410.

[9] T. Pursche, J. Krajewski, and R. Moeller, "Video-based heart rate measurement from human faces," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2012, pp. 544–545.

[10] L. Scalise, N. Bernacchia, I. Ercoli, and P. Marchionni, "Heart rate measurement in neonatal patients using a webcamera," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, May 2012, pp. 1–4.

[11] Y. Sun, C. Papin, V. Azorin-Peris, R. Kalawsky, S. Greenwald, and S. Hu, "Use of ambient light in remote photoplethysmographic systems: Comparison between a high-performance camera and a low-cost webcam," *J. Biomed. Opt.*, vol. 17, no. 3, pp. 370051–3700510, 2012.

[12] Y. Sun, S. Hu, V. Azorin-Peris, S. Greenwald, J. Chambers, and Y. Zhu, "Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise," *J. Biomed. Opt.*, vol. 16, no. 7, p. 077010, 2011.

[13] J. Penne, C. Schaller, J. Hornegger, and T. Kuwert, "Robust real-time 3D respiratory motion detection using time-of-flight cameras," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 3, no. 5, pp. 427–431, Nov. 2008.

[14] D. Falie, L. David, and M. Ichim, "Statistical algorithm for detection and screening sleep apnea," in *Proc. Int. Symp. Signals, Circuits Syst. (ISSCS)*, 2009, pp. 1–4.

[15] M.-C. Yu, H. Wu, J.-L. Liou, M.-S. Lee, and Y.-P. Hung, "Breath and position monitoring during sleeping with a depth camera," in *Proc. Int. Conf. Health Informat. (HEALTHINF)*, 2012, pp. 12–22.

[16] S. Kwon, H. Kim, and K. S. Park, "Validation of heart rate extraction using video imaging on a built-in camera system of a smartphone," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, San Diego, CA, USA, Aug. 2012, pp. 2174–2177.

[17] F. Bousefsaf, C. Maaoui, and A. Pruski, "Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 568–574, 2013.

[18] M.-C. Yu, J.-L. Liou, S.-W. Kuo, M.-S. Lee, and Y.-P. Hung, "Noncontact respiratory measurement of volume change using depth camera," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2012, pp. 2371–2374.

[19] J. Xia and R. A. Siochi, "A real-time respiratory motion monitoring system using KINECT: Proof of concept," *Med. Phys.*, vol. 39, no. 5, pp. 2682–2685, 2012.

[20] N. Bernacchia, L. Scalise, L. Casacanditella, I. Ercoli, P. Marchionni, and E. P. Tomasini, "Non contact measurement of heart and respiration rates based on Kinect," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, Jun. 2014, pp. 1–5.

[21] A. Al-Naji, K. Gibson, S.-H. Lee, and J. Chahl, "Real time apnoea monitoring of children using the microsoft Kinect sensor: A pilot study," *Sensors*, vol. 17, no. 2, p. 286, 2017.

[22] G. de Haan and V. Jeanne, "Robust pulse rate from chrominance-based rPPG," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 10, pp. 2878–2886, Oct. 2013.

[23] G. de Haan and A. Van Leest, "Improved motion robustness of remote-PPG by using the blood volume pulse signature," *Physiol. Meas.*, vol. 35, no. 9, p. 1913, 2014.

[24] X. Li, J. Chen, G. Zhao, and M. Pietikainen, "Remote heart rate measurement from face videos under realistic situations," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 4264–4271.

[25] D. McDuff, S. Gontarek, and R. W. Picard, "Improvements in remote cardiopulmonary measurement using a five band digital camera," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 10, pp. 2593–2601, Oct. 2014.

[26] G. Balakrishnan, F. Durand, and J. Guttag, "Detecting pulse from head motions in video," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2013, pp. 3430–3437.

[27] L. Shan and M. Yu, "Video-based heart rate measurement using head motion tracking and ICA," in *Proc. 6th Int. Congr. Image Signal Process. (CISP)*, 2013, pp. 160–164.

[28] R. Irani, K. Nasrollahi, and T. B. Moeslund, "Improved pulse detection from head motions using DCT," in *Proc. Int. Conf. Comput. Vis. Theory Appl. (VISAPP)*, 2014, pp. 118–124.

[29] L. Feng, L.-M. Po, X. Xu, and Y. Li, "Motion artifacts suppression for remote imaging photoplethysmography," in *Proc. 19th Int. Conf. Digital Signal Process. (DSP)*, 2014, pp. 18–23.

[30] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. Clifton, and C. Pugh, "Non-contact video-based vital sign monitoring using ambient light and auto-regressive models," *Physiol. Meas.*, vol. 35, no. 5, pp. 807–831, 2014.

[31] Y. Hsu, Y.-L. Lin, and W. Hsu, "Learning-based heart rate detection from remote photoplethysmography features," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2014, pp. 4433–4437.

[32] L. Feng, L. M. Po, X. Xu, Y. Li, and R. Ma, "Motion-resistant remote imaging photoplethysmography based on the optical properties of skin," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 5, pp. 879–891, May 2015.

[33] D.-Y. Chen *et al.*, "Image sensor-based heart rate evaluation from face reflectance using Hilbert-Huang transform," *IEEE Sensors J.*, vol. 15, no. 1, pp. 618–627, Jan. 2015.

[34] K.-Y. Lin, D.-Y. Chen, and W.-J. Tsai, "Face-based heart rate signal decomposition and evaluation using multiple linear regression," *IEEE Sensors J.*, vol. 16, no. 5, pp. 1351–1360, Mar. 2016.

[35] C. Wiede, J. Richter, and G. Hirtz, "Signal fusion based on intensity and motion variations for remote heart rate determination," in *Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST)*, Oct. 2016, pp. 526–531.

[36] J. Cheng, X. Chen, L. Xu, and Z. J. Wang, "Illumination variation-resistant video-based heart rate measurement using joint blind source separation and ensemble empirical mode decomposition," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 5, pp. 1422–1433, Sep. 2017.

[37] X. He, R. A. Goubran, and X. P. Liu, "Wrist pulse measurement and analysis using Eulerian video magnification," in *Proc. IEEE-EMBS Int. Conf. Biomed. Health Inform. (BHI)*, Las Vegas, NV, USA, Feb. 2016, pp. 41–44.

[38] A. P. Prathosh, P. Praveena, L. K. Mestha, and S. Bharadwaj, "Estimation of respiratory pattern from video using selective ensemble aggregation," *IEEE Trans. Signal Process.*, vol. 65, no. 11, pp. 2902–2916, Jun. 2017.

[39] A. Al-Naji and J. Chahl, "Contactless cardiac activity detection based on head motion magnification," *Int. J. Image Graph.*, vol. 17, no. 1, p. 1750001, 2017.

[40] A. Al-Naji and J. Chahl, "Non-contact heart activity measurement system based on video imaging analysis," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 31, no. 2, p. 1757001, 2017.

[41] A. Al-Naji and J. Chahl, "Remote respiratory monitoring system based on developing motion magnification technique," *Biomed. Signal Process. Control*, vol. 29, pp. 1–10, Aug. 2016.

[42] W. Wang, S. Stuijk, and G. de Haan, "A novel algorithm for remote photoplethysmography: Spatial subspace rotation," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 9, pp. 1974–1984, Sep. 2016.

[43] W. Wang, A. C. den Brinker, S. Stuijk, and G. de Haan, "Algorithmic principles of remote PPG," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 7, pp. 1479–1491, Jul. 2017.

[44] L. Xu, J. Cheng, and X. Chen, "Illumination variation interference suppression in remote PPG using PLS and MEMD," *Electron. Lett.*, vol. 53, no. 4, pp. 216–218, 2017.

[45] A. Al-Naji, A. G. Perera, and J. Chahl, "Remote monitoring of cardiorespiratory signals from a hovering unmanned aerial vehicle," *Biomed. Eng. Online*, vol. 16, no. 1, p. 101, 2017.

[46] H.-Y. Wu, M. Rubinstein, E. Shih, J. V. Guttag, F. Durand, and W. T. Freeman, "Eulerian video magnification for revealing subtle changes in the world," *ACM Trans. Graph.*, vol. 31, no. 4, p. 65, 2012.

[47] B. Madhukar and R. Narendra, "Lanczos resampling for the digital processing of remotely sensed images," in *Proc. Int. Conf. VLSI, Commun., Adv. Devices, Signals Syst. Netw. (VCASAN)*, 2013, pp. 403–411.

[48] A. Al-Naji, K. Gibson, S. H. Lee, and J. Chahl, "Monitoring of cardiorespiratory signal: Principles of remote measurements and review of methods," *IEEE Access*, vol. 5, pp. 15776–15790, 2017.

[49] A. Al-Naji, S.-H. Lee, and J. Chahl, "Quality index evaluation of videos based on fuzzy interface system," *IET Image Process.*, vol. 11, no. 5, pp. 292–300, 2017.

[50] S. Liao, A. K. Jain, and S. Z. Li, "A fast and accurate unconstrained face detector," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 211–223, Feb. 2016.

[51] X. Zhang and P. Zhou, "Filtering of surface EMG using ensemble empirical mode decomposition," *Med. Eng. Phys.*, vol. 35, no. 4, pp. 537–542, 2013.

[52] X. Chen, A. Liu, J. Chiang, Z. J. Wang, M. J. McKeown, and R. K. Ward, "Removing muscle artifacts from EEG data: Multichannel or single-channel techniques?" *IEEE Sensors J.*, vol. 16, no. 7, pp. 1986–1997, Apr. 2016.

[53] M. Blanco-Velasco, B. Weng, and K. E. Barner, "ECG signal denoising and baseline wander correction based on the empirical mode decomposition," *Comput. Biol. Med.*, vol. 38, no. 1, pp. 1–13, Jan. 2008.

[54] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, 2008.

[55] M. A. Colominas, G. Schlotthauer, and M. E. Torres, "Improved complete ensemble EMD: A suitable tool for biomedical signal processing," *Biomed. Signal Process. Control*, vol. 14, pp. 19–29, Nov. 2014.

[56] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2011, pp. 4144–4147.

[57] W. Liu, D. P. Mandic, and A. Cichocki, "Analysis and online realization of the CCA approach for blind source separation," *IEEE Trans. Neural Netw.*, vol. 18, no. 5, pp. 1505–1510, Sep. 2007.

[58] Y.-O. Li, T. Adali, W. Wang, and V. D. Calhoun, "Joint blind source separation by multiset canonical correlation analysis," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 3918–3929, Oct. 2009.

[59] L. Zou, X. Chen, A. Servati, S. Soltanian, P. Servati, and Z. J. Wang, "A blind source separation framework for monitoring heart beat rate using nanofiber-based strain sensors," *IEEE Sensors J.*, vol. 16, no. 3, pp. 762–772, Feb. 2016.

[60] K. T. Sweeney, S. F. McLoone, and T. E. Ward, "The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 97–105, Jan. 2013.

[61] D. Safieddine *et al.*, "Removal of muscle artifact from EEG data: Comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches," *EURASIP J. Adv. Signal Process.*, vol. 2012, no. 1, p. 127, 2012.

[62] Y.-P. Lin, Y. Wang, and T.-P. Jung, "A mobile SSVEP-based brain-computer interface for freely moving humans: The robustness of canonical correlation analysis to motion artifacts," in *Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2013, pp. 1350–1353.

[63] J. Gao, C. Zheng, and P. Wang, "Online removal of muscle artifact from electroencephalogram signals based on canonical correlation analysis," *Clin. EEG Neurosci.*, vol. 41, no. 1, pp. 53–59, 2010.

[64] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2583–2587, Nov. 2006.

[65] M. Borga and H. Knutsson, "A canonical correlation approach to blind source separation," Dept. Biomed. Eng., Linköping Univ., Linköping, Sweden, Tech. Rep. LiU-IMT-EX-0062, 2001.

[66] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Int. J. Nursing Stud.*, vol. 47, no. 8, pp. 931–936, 2010.

ALI AL-NAJI received the bachelor's degree in medical instrumentation engineering techniques from the Electrical Engineering Technical College, Middle Technical University, Baghdad, Iraq, in 2005, and the M.S. degree from the Electrical and Electronic Engineering Department, Technology University, Baghdad, Iraq, in 2008. He is currently pursuing the Ph.D. degree with the School of Engineering, Electrical and Information Engineering, University of South Australia. His research interests include biomedical engineering, computer vision systems, and microcontroller applications.

JAVAAN CHAHL received the Ph.D. degree from The Australian National University. He joined the Defence Science and Technology Group (DST Group), as a Research Scientist, in 1999. In 2011, he joined RMIT University, as a founding Professor of unmanned aerial vehicles. In 2012, he became a Chair of sensor systems a Joint appointment between DST Group and the University of South Australia. He is a member of the Institute of Engineers Australia and was elected as a Fellow of the Royal Aeronautical Society in 2014.