IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 8, 2019, accepted July 22, 2019, date of publication July 31, 2019, date of current version August 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2932204

A New Approach for Noncontact Imaging Photoplethysmography Using Chrominance Features and Low-Rank in the IoT Operating Room

HONGWEI YUE[1], XIAORONG LI[2], HONGTAO WANG[1], HUAZHOU CHEN[3], XIAOJUN WANG[4], AND KEN CAI[5]

[1]Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen 529020, China
[2]Department of Radiology, General Hospital of Southern Theater Command, PLA, Guangzhou 510010, China
[3]College of Science, Guilin University of Technology, Guilin 541004, China
[4]Industrial Training Center, Guangdong Polytechnic Normal University, Guangzhou 510665, China
[5]College of Automation, Zhongkai University of Agriculture and Engineering, Guangzhou 510225, China

Corresponding author: Ken Cai (icken@126.com)

This work was supported in part by the China Postdoctoral Science Foundation under Grant 2018T110880 and Grant 2017M620375, in part by the Natural Science Foundation of Guangdong Province under Grant 2018A030313882 and Grant 2016A030313003, in part by the National Natural Science Foundation of China under Grant 81671788 and Grant 61505037, in part by the Projects for International Scientific and Technological Cooperation under Grant 2018A05056084, in part by the Science Foundation of Wuyi University for Young Teachers under Grant 2018td01, in part by the Guangzhou Science and Technology Program under Grant 201805010001, in part by the Guangdong Provincial Science and Technology Program under Grant 2016B090927009, and in part by the Natural Scientific Foundation of Guangxi under Grant 2018GXNSFAA050045.

ABSTRACT Advancements in image-processing and medical imaging technologies and other research areas have resulted in the growing importance of surgical navigation systems in the field of minimally invasive surgery. However, sophisticated auxiliary equipment is a prerequisite for full heart-rate monitoring in operative systems. A system puts forward to address this issue with a wireless sensor network architecture comprising acquisition modules, router gateways, remote servers, and a medical monitoring center. Factors such as variations in ambient light and face shaking can also easily affect heart rate detection based on face videos, thus resulting in inaccurate estimations of heart rate from the blood volume pulse (BVP) signals. This study proposes to address this concern by employing a novel method for non-contact heart rate estimation to overcome noise interference. First, chrominance features are selected to extract BVP signals, and the low-rank and sparse matrix decomposition methods are applied to overcome the detrimental effects of noise and interference while ensuring that valuable data are preserved.Next, the data are relayed to the server via the gateway. Finally, users can log on to the health-related cloud platform and gain information regarding their health status in real time. Experimental results reveal the advantages of the proposed technique over conventional face-based heart-rate estimation methods, including the capability to decrease dependence on sophisticated auxiliary equipment, avoid direct skin contact that can cause discomfort to patients undergoing surgery, and improve the comfort of surgical operations. Moreover, the proposed heart-rate measurement technique can contribute to improve the construction of smart cities.

INDEX TERMS Smart city, photoplethysmography, chrominance features, heart-rate detection, low-rank and sparse matrices.

I. INTRODUCTION

Smart cities are a regional development concept that originated from the application of innovative technologies, including Internet of Things (IoT), electronic communications, and cloud computing. The use of IoT and big data in smart city construction can aid in enhancing services for urban dwellers and realizing the sustainable development of the cities for a long time. Numerous researchers have recently studied the public services offered smart cities, and some are

The associate editor coordinating the review of this article and approving it for publication was Chang Choi.

suggesting the use of smart cities in public domains such as smart healthcare, public administration, and intelligent transportation, among many others[1].

Smart healthcare is an integral component of smart cities and has emerged as a research hotspot in recent years. This health service system can be an effective support in the construction of smart cities through leveraging IoT technology to develop wireless and miniature medical apparatus.

In addition, surgical systems enhance the safety and accuracy during surgical operations by using local anatomy visualization to direct surgical instruments toward the lesion sites [2]. During surgery, the patient's heart rate, breathing, and blood pressure, among other vital signs, should be monitored. As a key indicator of human physiological condition, heart rate has significant applications in clinical research on cardiovascular diseases [3], physical exercise [4], and other fields. However, the requirement of direct skin contact for most heart-rate measurement apparatuses can be a source of distress for surgical patients and may adversely affect the operation. In photoplethysmography (PPG), variations in transmitted and reflected light are used to acquire the cardiovascular pulse wave [5]. PPG can greatly improve the ease of heart-rate measurement, but it requires a specific light source, such as infrared light. Moreover, a fingertip spring clip is attached to the patient being monitored in PPG. To date, even though contact-assisted devices used in clinical practice can be monitored in real time, the auxiliary equipments still occupy large amounts of space, which can adversely influence the surgery. Consequently, the development of comprehensive noncontact heart-rate detection and estimation approaches has emerged as research hotspot in the field of physiological information monitoring to minimize the drawbacks of current heart-rate measurement techniques. Many new technologies for noncontact heart-rate detection and estimation have been suggested, such as ultrasonic testing [6], thermodynamic imaging [7], and video image-based imaging PPG (iPPG) [8]. Among these techniques, remote heart rate detection and estimation based on video images can enhance patient comfort and convenience as well as lower medical expenses because it can automatically monitor the heart rate just by processing the body skin color from the collected video.

Surgical systems have likewise been applied in abdominal surgery, spinal surgery, and neurosurgery [9], [10]. They can be classified as mechanical, electromagnetic, optical, or ultrasonic systems according to their spatial positioning technologies [11]. However, these systems are incapable of gathering and reading various clinical data and image records of patients and achieving communication between operating rooms and the outside. Medical informatization is crucial element of smart cities, the development of which is bound to impact the medical industry. The IoT network links the physical domain by integrating intelligent processing and computing technologies with sensing technologies, such as sensor network, communication network, and the Internet. The advent of IoT technology has realized the interconnectedness of medical data and stimulated the growth of the healthcare information industry, resulting in the creation of intelligential services and breaking out from the numerous bottlenecks plaguing the industry.

When IoT technology and non-contact physiological signal measurement are combined, medical quality and quantity can be elevated because surgeons will be able to maintain their focus on the operations and improve their surgical precision.

The rest of this paper is divided into the following sections. Second II presents the related works. Section III discusses the theoretical background and describes the main principles of robust principal component analysis (RPCA) and the chrominance feature model. Furthermore, it provides a discussion of blood volume pulse (BVP) denoising and reconstruction. Section IV elaborates on the experimental results. Section V concludes the paper.

II. RELATED WORKS

The importance of IoT technology in medical informatization is primarily represented in the following aspects [12]. First, the application of IoT technology for online monitoring in precision medicine can help manage the medical process, which can lower incidence of medical accidents in unpredictable scenarios because of human factors. Second, the use of IoT technology in the medical process can promote the digital progress of this field, realize the dynamic observation of diseases, and perform follow-up analysis. Fig. 1 illustrates an IoT-based medical care system. In the figure, the sensors gather crucial human physiological parameters, which are processed by the processor and then transmitted to the ZigBee network.

In this work, iPPG is primarily studied to address the difficulty connected with the use of sophisticated auxiliary instruments for monitoring heart rate.

IPPG is a new photoelectric detection technology that measures the reflected light intensity after living tissue is irradiated with a photoelectric device to detect variations in blood volume in capillaries and microarteries [13]. BVP signals that contain key physiological information are obtained and widely used in the clinical measurement of human physiological parameters [14]. The skin's optical properties are generally credited to light absorption by hemoglobin, melanin, subcutaneous moisture, and other factors. Of these factors, the highest light absorption is by blood [5]. The intermitted perfusion of blood influences the intensity of reflected light when light enters subcutaneous tissue, where blood perfusion produces the AC component of the light signal and the reflected light from surrounding tissues and blood vessels creates the DC component of the light signal [15]. Specifically, subtle variations in human skin color are tightly linked to variations in blood volume in capillaries and microarteries. Therefore, disparities in light signal reflected by the vascular network can be perceived using imaging devices to capture the light. Additionally, key physiological data from variations in blood volume can be measured on the basis of human

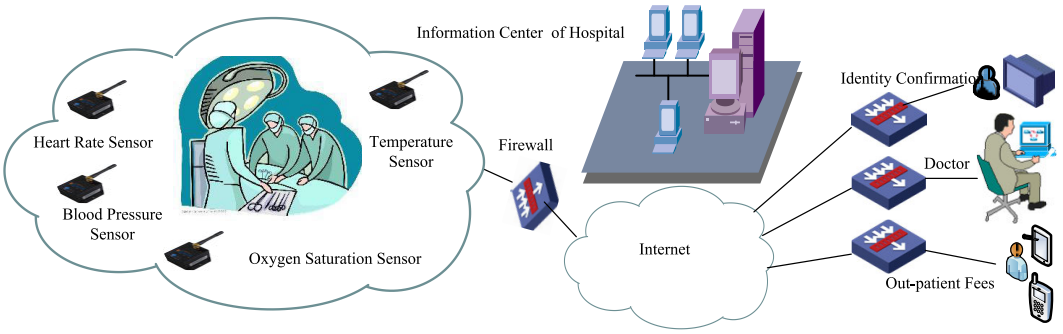


FIGURE 1. Medical care system based on IoT.

skin color. Given that the iPPG signal originates from the reflection of skin capillaries to light, the pseudo difference caused by movement will inevitably influence the signal, and the unstable light source will generate an extracted signal with a low signal-to-noise ratio.

Existing studies on iPPG have mainly focused on image quality [16], frame rate [17], region of interest (ROI) selection [18], motion pseudo difference [19], and clinical trials and applications [20]–[22]. Studies on video-based noncontact heart-rate estimation aim to obtain physiological parameters with increased accuracy. Recently, numerous video-based noncontact heart-rate measurement approaches have been proposed and innovative ideas and approaches for eliminating the interference component of the iPPG signal have been put forward. In 2011, Sun *et al.* [23] utilized the single-channel independent component analysis (ICA) method for the movement attenuation of iPPG based on a black-and-white camera. Subsequently, in 2014, an adaptive filter and the constraint ICA method were applied for iPPG signal extraction [24]. Poh *et al.* [25], [26] and Pursche *et al.* [27] adopted the ICA algorithm for denoising and blind source separation. However, interference is brought about by large non-skin areas of the face, and interference with a frequency close to that of the heart rate, such as blinking, cannot be easily eliminated by this technique. In 2016, Chwyl *et al.* [28] proposed a heart-rate detection approach based on Bayes estimation.

Other aspects of pulse-signal extraction have been improved to further improve its precision. In 2015, Lam and Kuno [29] extracted pulse signals from multipoint local regions of the face to exclude movement artifacts and then used the signals as linear blind-source signals for the source decomposition of independent information. They next acquired pulse signals with practical physiological value through a voting method. However, their method was time-consuming and had difficulties when the face moved or was obstructed because of the random selection of multipoint local regions of the face.

The development of skin models to diminish the impact of light variations is currently a research hotspot. The video-data-use mode has been extensively examined as well. For instance, a chrominance model algorithm, such as CHROM method [20], has been suggested by Haan. The author considered BVP signals as the linear combination of chrominance signals and established a standard skin color for the white-balance camera. The author then further optimized the chrominance model [30] to enhance the removal of movement interference. In 2016, Wang proposed the plane-orthogonal-to-skin (POS) algorithm [31], which can define a plane that is orthogonal to the skin-tone in a temporally normalized RGB space for pulse extraction. This algorithm exhibits improved anti-interference capability.

Noncontact video-based heart-rate estimation practically evolved from the principle of PPG [32]. Skin tissue and the ambient temperature of the patient affect blood perfusion in the skin, a condition that raises interference in the iPPG signal and, consequently, decreases estimation precision for heart rate and blood oxygen saturation. This effect likewise shows that disparities in facial movement, ambient temperature, and light around the face diminish the stability and precision of face-based heart-rate measurement based on the principle of PPG. The superposition of these interferences emphasizes the urgent need for stringent requirements for filtering non-BVP signals.

RPCA is a promising signal processing method in signal denoising technology and feature extraction. Its development was grounded on the theories of matrix completion and compressive perception [33]. RPCA can recreate low-rank and sparse representation components from gathered data, from which implicit target information can then be obtained. While video backgrounds with robust data redundancy can match low-rank components after decomposition, changing targets with rapid amplitude variations can match sparse components after decomposition. Therefore, RPCA possesses good signal separation properties and can be utilized to distinguish between foreground targets and video backgrounds. Wright effectively applied RPCA in a video target detection system by forming an observation matrix with image frames as column vectors and then decomposing the matrix via RPCA [34].

In this paper, variations in skin color are considered as superimpositions on repeated variations in blood volume by changing noises (e.g., noises caused by light, shaking, and temperature) and can be presumed as estimations by a low-rank matrix. By contrast, variations in skin color caused by noise can be presumed as sparse, with a sparse matrix

employed to estimate the noise. RPCA can decompose such information into a low-rank matrix and a sparse matrix that correspond to blood volume and noise variations, respectively. Then, the effective BVP signal can be extracted and noise interference filtered out.

Working from the above research, we propose an IoT-based heart-rate detection technique that blends low-rank and sparse decomposition to denoise BVP signals while simultaneously enhancing the ability of such signals to preserve valuable data. This method denoises and recreates chrominance features by applying the RPCA algorithm after the face image is derived via face-tracking technology and corrected. Numerous experiments show that the proposed method can be amended for analyzing and processing BVP signals and providing precise heart-rate estimations.

With the proposed noncontact heart-rate detection method, the need for complicated hardware equipment is eliminated, and the working space of surgical systems is optimized.

III. THEORETICAL BACKGROUND OF HEART-RATE ESTIMATION

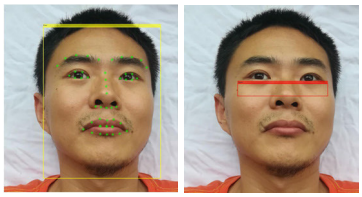
Precise estimation of BVP signals is the main problem in heart-rate estimation based on face videos. The proposed method obtains heart-rate estimation results by (1) creating a BVP signal acquisition model, (2) considering the influence of incident illumination and subcutaneous and epidermal reflection, and (3) involving chrominance features and low-rank and sparse decomposition to remove interference from movement and light.

A. FACE DETECTION AND SKIN REGION SELECTION

The PPG principle asserts that blood volume varies with the periodic diastolic and systolic function of the human heart [35]. Such variations can also alter the intensity of reflected and absorbed lights, resulting in the periodic variations in skin color. A face-detection method must be used to extract data related to the periodic variations in human skin color, remove the background region of the face, and locate the ROI.

Despite the numerous face-recognition technologies [36]–[39], face detection in unconfined environments is challenging because of the different illuminations, movements, and occlusions. To achieve accurate face location, initially, the Viola–Jones face detector is adopted to identify the rectangular face region of the first frame of the image and set the coordinates of the rectangular box. Then, a total of 49 face feature points are marked for subsequent tilt correction and tracking,, as shown in Fig. 2(a).

Skin pixels can be estimated to consist of diffuse and specular reflections. Variations in diffuse reflection caused by facial movement are smaller than those in specular reflection. Thus, in existing literature, spatial averaging of the illumination intensity of the entire facial area can remove the discontinuous variation in epidermal reflection. However, the BVP signal quality of different face regions in real environments is changes due to the influence of several



(a) (b)

FIGURE 2. Selected region of face: (a) initial face region; (b) ROI.

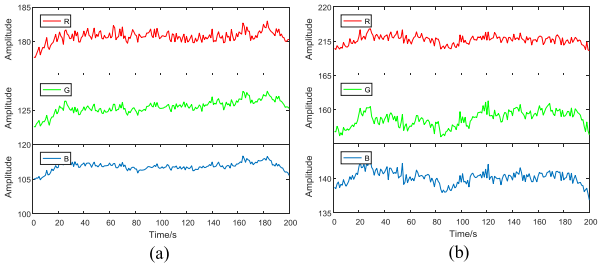


FIGURE 3. Time sequence signals of the RGB channels: (a) time sequence signals of the entire face region; (b) time sequence signals of the ROI.

factors, including facial movement, hair occlusion, and corrective glasses. Fig. 3 displays the time series of three channels obtained from the different face regions of the subject. Figs. 3(a) and 3(b) present the component obtained via spatial averaging of the entire face region and the region selected in this work(Fig. 2(b)), respectively. Interference components with complexity and strong randomness comprise the signals acquired via spatial averaging because the interferences formed through variations in illumination, uncontrolled movement of facial organs, and overall facial movement superimpose one another. In this study, the region chosen has more blood vessels and a steadier signal than the entire face area. Fig. 3 reveals that the BVP signal in the nose area undergoes considerable numerical variation. Therefore, instead of just averaging the pixels of the whole face area, the region with high signal-to-noise ratio must be chosen instead for heart-rate estimation and BVP signal extraction.

B. EXTRACTION OF CHROMINANCE FEATURES

The iPPG system gathers original video images that have pulse-wave data. Acquiring a pulse-wave signal with high signal-to-noise ratio requires an algorithm with strong anti-interference ability to achieve extraction of the original pulse signal. When the sensitive area is identified, the current acquisition technique for the 1-D iPPG signal (R, G, B) is used to compute for the mean value of this area. However, the technique cannot enhance the separation of the pulse and specular reflection components. Hence, the POS model is used to extract the chrominance features [31]. These features are robust against facial movement, which results in variations in facial brightness and color because of changes in the angle between the face and the incident light source. Chrominance features can minimize this phenomenon.

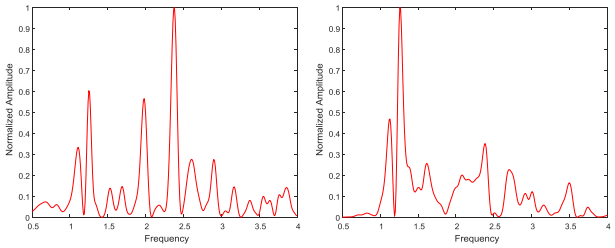


FIGURE 4. Comparison of frequency spectra: (a) spectrum of original pulse signal extracted by the G channel; (b) spectrum of original pulse signal extracted by the POS algorithm.

In the POS model, CHROM principle and BVP knowledge are used for chrominance feature extraction, which approximates a projection region on the plane orthogonal to the temporally normalized skin-tone direction and real-time tuning to refine an exact projection direction on the plane. The chrominance feature when the R, G, and B channels are merged is computed as

$$h(t) = S_1(t) + \alpha \cdot S_2(t) \tag{1}$$

where $\alpha = \sigma(S_1)/\sigma(S_2)$, $\sigma(\cdot)$ denotes the standard deviation operator,and $S_1(t) = G(t) - B(t)$, $S_2(t) = G(t) + B(t) - 2R(t)$.

The POS model defines skin color as a linear combination of pulse and specular reflection components. $S_1(t)$ and $S_2(t)$ appear in the same phase when the pulse wave is dominant. The appearance of $h(t)$ can intensify the obtained signal. $S_1(t)$ and $S_2(t)$ are inverted when the specular variation is dominant. α can pull/push the intensity of the specular variation of one signal to the same level as the other signal, that is, $\sigma(S_1) = \sigma(\alpha \cdot S_2)$. Two inverted signals then offset specular distortion.

The heart-rate estimation and iPPG signal extraction results rely on the features extracted from the face ROI. Here, the chrominance features robust to facial motion and discriminate against skin color variations are chosen. Next, the POS model and G channel data are utilized to obtain original pulse signals from the video samples of facial shaking (Fig. 4). As presented in the figure, the POS model improves the suppression of noise caused by head movement.

C. DENOISING OF BLOOD VOLUME PULSE SIGNAL

RPCA is also referred to as sparse matrix and low-rank matrix decomposition because it primarily aims to decompose the video matrix into sparse and low-rank matrices. The RPCA model can simultaneously estimate the background, separate moving targets, and conduct background estimation for videos containing moving targets without needing to input a video with a clean background as the background training sample. The low-rank matrix remains strong against degenerative factors, including data loss, noise, and slow light variations. It does not have special requirements to determine the motion mode of the foreground target. Moreover, the low-rank matrix can handle non-rigid objects and requires modification of few parameters. These characteristics help in the promotion and application of the matrix in the engineering field [40], [41].

$M \in R^{i \times j \times p}$ denotes the video sequence, where i and j refer to the size of the video, and p is the number of video frames. A novel matrix$M \in R^{(i \times j) \times p}$ can be created by converting each video frame into a column vector.

$$M = [vec(M_1), \ldots, vec(M_p)] \in R^{(i \times j) \times p} \tag{2}$$

Each frame of video sequence images is taken as the sum of the background and target prospects, that is,

$$M_1 = S_1 + e_1, \ldots, M_p = S_p + e_p \tag{3}$$

Therefore, the image background of low rank component is $A = [vec(S_1), \ldots, vec(S_p)]$, and the image foreground of sparse component is $E = [vec(e_1), \ldots, vec(e_p)]$.

Each frame image is transformed to $S_i \circ \tau_i^{-1}$ so that the video sequence data matrix approximation meets the sparse and low rank. The video sequence after transformation is expressed as

$$M_1 = (S_1 + e_1) \circ \tau_1^{-1}, \ldots, M_p = (S_p + e_p) \circ \tau_p^{-1} \tag{4}$$

Low-rank decomposition of the video sequences with $D \circ \tau^{-1}$ can be expressed as

$$\min_{A,E} rank(A) \; s.t. \; D \circ \tau = A + E, \quad \|E\|_0 \leq k \tag{5}$$

Formula (5) is converted to a Lagrangian type:

$$\min_{A,E} rank(A) + \lambda \|E\|_0 \; s.t. \; D \circ \tau = A + E \tag{6}$$

The model must be modified because the matrix decomposition in Equation (6) is an NP problem. A feasible method involves the application of the kernel norm to estimate the matrix rank and of the L_1 norm of the matrix to estimate the L_0 norm of the matrix. A solution model for the convex optimization problem then can be obtained:

$$\min_{A,E} \|A\|_* + \lambda \|E\|_1 \; s.t. \; D \circ \tau = A + E \tag{7}$$

$$\|A\|_* = \sum_{i=1}^{m} \mu_i(A) \tag{8}$$

$$\|E\|_1 = \sum_{ij} |E_{ij}| \tag{9}$$

wherein λ is a positive weight parameter to weigh the rank of matrix A and the sparsity of matrix E; $\|\cdot\|_*$ represents the kernel norm of matrix A, which is used to replace the rank of matrix A;and $\|\cdot\|_1$ represents the sum of the absolute values of the matrix elements. Minimizing the following augmented Lagrangian function with the alternating direction method of multipliers can resolve this optimization problem.

$$L(A, E, Y, \mu) = \|A\|_* + \lambda \|E\|_1 + \langle Y, X - A - E \rangle + \mu/2 \|X - A - E\|_F^2 \tag{10}$$

In the equation, Y is a Lagrangian multiplier, and $\mu > 0$ is a penalty parameter. When this model is solved, the low-rank

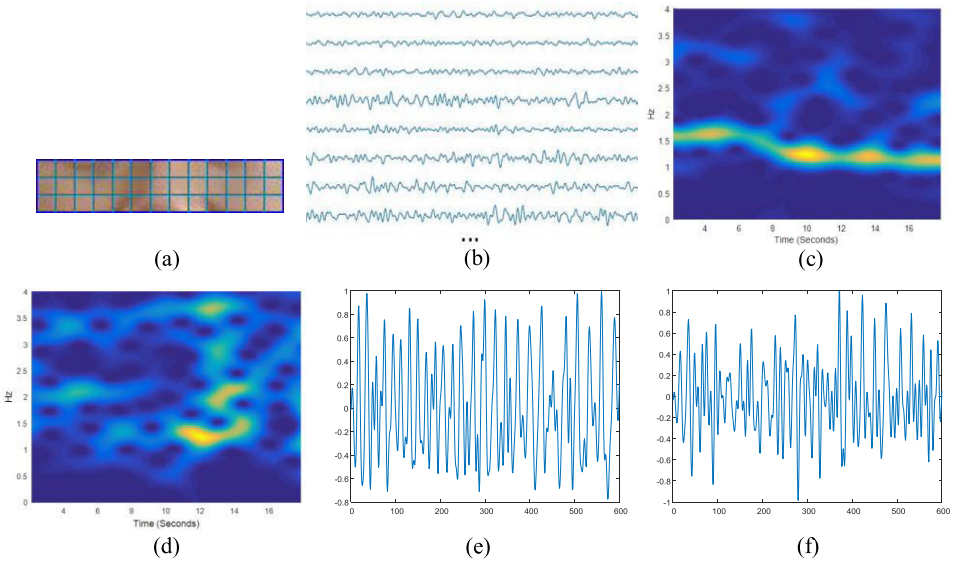


FIGURE 5. Denoising of BVP signal: (a) gridded ROI; (b) chrominance features of each grid block; (c) spectrogram of the low-rank component obtained by RPCA decomposition; (d) spectrogram of the sparse component obtained by RPCA decomposition; (e) time domain waveform of the sparse component; (f) time domain waveform of the low-rank component.

background matrix A and the sparse foreground matrix E can be decomposed.

The heart-rate signal is connected to the primary signal of the low-rank background matrix. Singular value decomposition is conducted for feature vector extraction. Here, matrix A is denoted as $A = U\Gamma V^{\mathrm{T}}$, where V_1 is the approximate heart-rate signal, and V_1 is the first column vector of V.

In a video signal, impact source (interference) and harmonic source (heartbeat) are characterized by anisotropy in the spectrum, that is, harmonic source is irregular in the frequency-axis direction and smooth and continuous in the time-axis direction. By contrast, the impact source is mostly concentrated in a short period and will form a series of vertical broadband spectra. Hence, the impact source is irregular in the time-axis direction and smooth and continuous in the frequency-axis direction. Fig. 5 presents the schematic of low-rank and sparse decomposition. First, the ROI is divided into $m \times n$ grid blocks, as shown in Fig. 5(a).Afterward, the chrominance features of each image block are calculated.If the video has p frames in total, then the RPCA can be used to conduct low-rank and sparse decomposition for the matrix $mn \times p$ in size, as shown in Fig. 5(c) and 5(d). Fig. 5 illustrates that under strong low-rank and sparse assumption, heartbeat components with repetitive rhythms are often assigned to the low-rank portion rather than to the sparse portion. The interference component features a short duration and a sparse characteristic and is thus assigned to the sparse component. Therefore, RPCA can separate the background and the foreground. In accordance with the aforementioned noise suppression method, high-frequency noise components in the original BVP wave can be removed, and the reconstructed BVP wave can be obtained. The AC components of the pulse wave can be observed well from the reconstructed BVP wave, as shown in Fig. 5(e).

IV. EXPERIMENT RESULTS AND ANALYSIS

The medical industry is a core element of urban construction. IoT technology and cloud computing improve the intelligence of medical care when building smart cities. Fig. 6 shows the operating room where the IoT-based optical navigation system in this work is located. In this navigation system, heart rate and other physiological parameters analyzed through cloud computing are relayed to the surgeon via IoT.

A. SOFTWARE ARCHITECTURE AND INTELLIGENT NODE

The application software of smart medical systems is a function of the entire system. This study designs the interaction diagram between modules of the server (Fig. 7). The gateway transmits data to the server through Socket communication. After receiving the data, the server determines the data type and format, parses and stores the information in the database, and displays the data through various charts on a web page.

As a key component of the information-gathering system, the intelligent node transmits various information gathered by the surgical site to the center server after preliminary processing. The node also receives commands from the control center and carries out suitable operations according to these commands. Furthermore, it is primarily responsible for running the operating system, web server, and application, as shown in Fig. 8.

B. SELECTION OF SLIDING WINDOW LENGTH

Videos of the subjects are recorded using a network camera. The subjects are positioned at a distance of 0.8-1 m from the camera, and the videos are recorded at the rate of 30 frames/s and with an image resolution of $1,920 \times 1,080$ pixels. The subject's whole face must be positioned within the video screen. Eighty-five subjects are grouped into three. For Group 1, data are collected as the subjects hold their breath

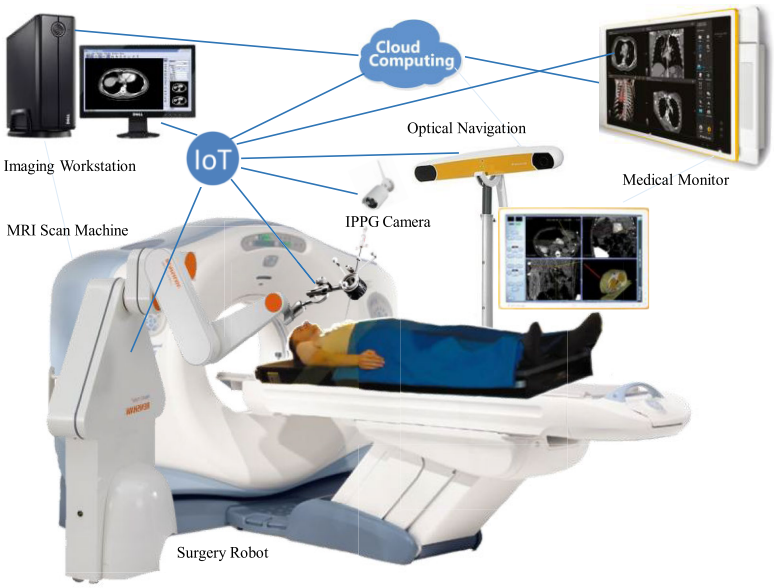


FIGURE 6. IoT operating room.

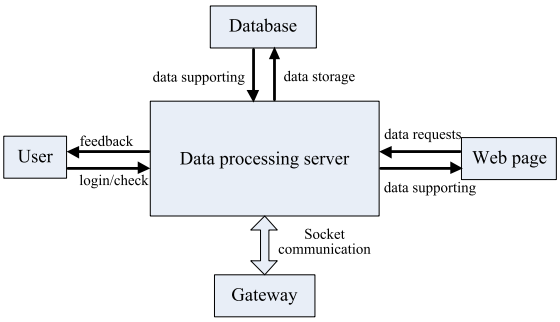


FIGURE 7. Interaction diagram between modules of the server.

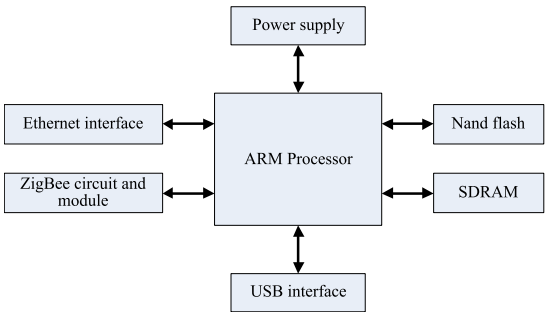


FIGURE 8. Frame structure of the intelligent node.

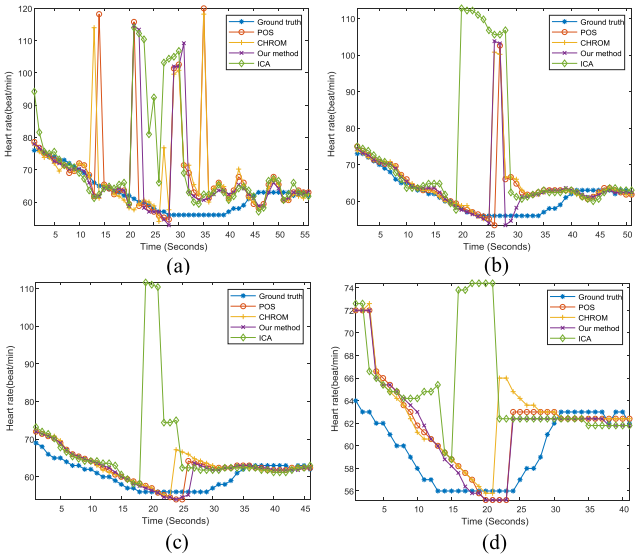


FIGURE 9. Comparison of estimated heart-rate values acquired with sliding windows of varying lengths.

and their heads immobile. The subjects can release their breath and breathe deeply when they can no longer hold their breath. For Group 2, data are collected as the subjects hold their heads immobile. For Group 3, data are collected as the subjects shake their heads intermittently. The real heart rate signals of the human body are collected via pulse oximeter as the reference value during video recording. The heart rate value is collected once per second by the pulse oximeter.

A comparison of the algorithms' robustness for the ROI in this study is conducted, in which the effect of video-recording duration on heart-rate detection in real scenes is considered. We detect heart rates for different video-recording durations and assess the real-time performances of the algorithms for each duration. The samples in Group 1 are used for testing. The video samples in Group 1 last for 60 s, and the sliding windows of 5s, 10s, 15s, and 20s with 1s increment are used. Different algorithms are used for extracting BVP signals from the samples in Group 1, and the corresponding power spectrum is obtained. Thus, the heart-rate measurement is completed. The skin region is considered as input samples for the CHROM [20], ICA [24], and POS methods to validate the performance of the proposed method. The measurement results from the finger clip-type pulse oximeter are compared with those from the iPPG acquisition system. Fig. 9 illustrates the generated heart-rate measurement curve.

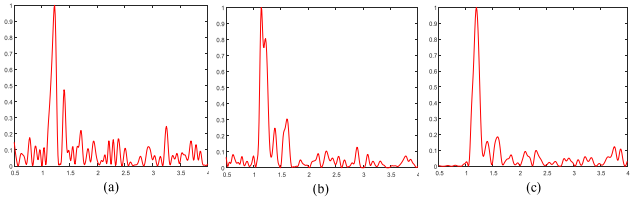


FIGURE 10. Normalized spectrograms of different ROI regions: (a) obtained from the whole face region; (b) obtained from the proposed ROI; (c) obtained after using RPCA.

None of the four methods can overcome the interference in the signal when the 5s sliding window is used (Fig. 9(a)). Significant differences are introduced by interference between the real and approximated heart rates. Fig. 9(d) shows that when the 20s sliding window is used, heart-rate values measured by the four methods follow the same downward trend as that by the values measured via pulse oximeter. These values also have larger errors than those obtained via pulse oximeter. An obvious delay is observed for the results acquired using the four methods when heart rates declined. The measurement accuracy of the iPPG system does not meet the requirements of clinical application when the sliding window is set to 20s.

Figs. 9(b) and (c) reveal that the heart-rate value measured via the contact system has a good linear relationship with the actual heart-rate value when the heart rate changed (before 25s). The measurement results obtained via the finger clip-type system are marginally lower than the noncontact measurement results obtained after 25s mainly because, after holding their breath, subjects begin to breathe freely and deeply after 25s of measurement. This action influences the blood perfusion of the skin. The estimated value recovers quickly because the human face is close to the heart than to the hand and is rich in capillaries and arterioles. Therefore, the noncontact measurement results are marginally higher than the measurements obtained via the finger clip-type system during the later stages of video recording.Fig. 9(b) shows that the proposed method has better real-time performance and interference suppression as well as less hysteresis curves than the method used to obtain the results in Fig. 9(c). In addition, the sliding window is set to 10s for performance analysis.

C. COMPARISON OF DIFFERENT METHODS

Fig. 10 presents the normalized spectrogram of sample 1 for Group 2. Fig. 10(a) presents the results after processing when the entire face was taken as the ROI, and Figs. 10(b) and 10(c) present the spectrograms with proposed ROI. The signal-to-noise ratio in Fig. 10(a) is lower than those acquired in Figs. 10(b) and 10(c), signifying that the different face regions influence the signal-to-noise ratio of the BVP signal.

For an intuitive comparison, the four methods are employed to extract BVP signals and analyze the time frequency for the videos of subjects in Groups 2 and 3. Figs. 11 and 12 show the time-frequency images of some samples.

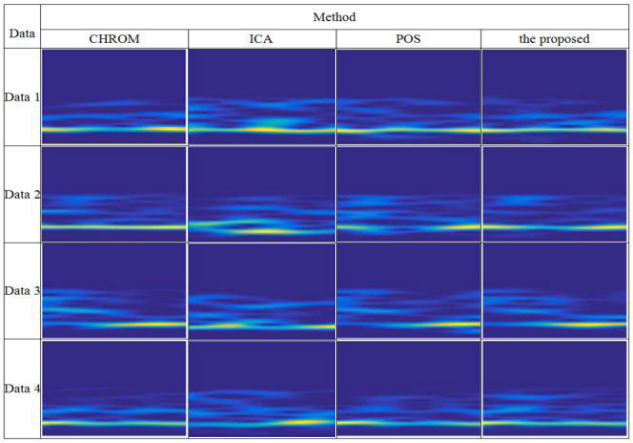


FIGURE 11. Time-frequency analysis of several subjects in Group 2.

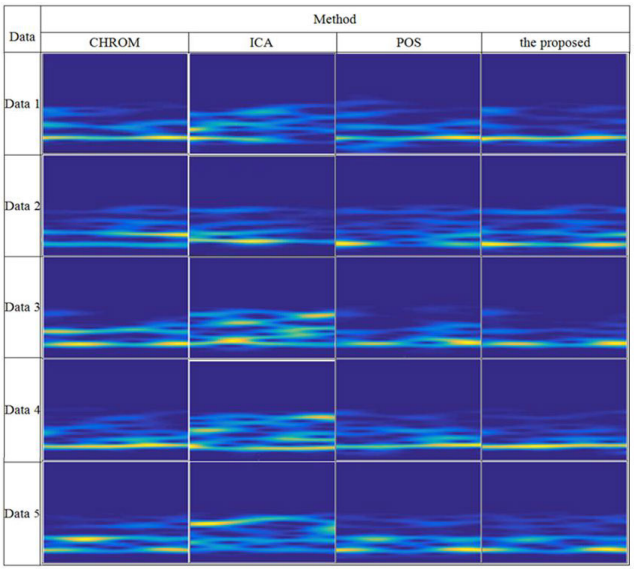


FIGURE 12. Time-frequency analysis of several subjects in Group 3.

Unpredictable interference may occur because of the uncontrolled movements of subjects and random variations in the test environment during long-term dynamic heart-rate monitoring based on face videos. Figs. 11 and 12 reveal the following: 1) ICA cannot effectively eliminate interference and obtain more accurate BVP signals. The poor precision of ICA is illustrated in the time-frequency images of samples 3 and 5 in Fig. 12. 2) CHROM, POS, and the proposed method obtained similar time-frequency images, as shown in the time-frequency images of Group 2. Meanwhile, the time-frequency images of Group 3 reveal that POS and the proposed method have advantages in eliminating interference when the sampled data have large interference and that the latter can obtain a clean time-frequency spectrum. As illustrated in the time-frequency images of sample 1 in Fig. 12, the signal extracted by the POS model has more noise than that extracted by the proposed method.

Fig. 13 illustrates the spectrogram (sparse part) obtained by the proposed approach. The spectrogram matches the

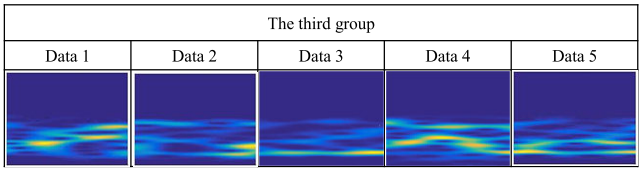


FIGURE 13. Spectrogram of the sparse component obtained by RPCA.

TABLE 1. Performance of heart rate detection.

Method	Proposed	POS	CHROM	ICA
Me	2.80	2.82	5.99	10.48
RMSE	4.91	5.04	15.34	19.12
STD	3.96	8.40	14.24	14.24

spectrogram (low-rank part) obtained by the proposed method in Fig. 12 to exhibit the denoising performance of RPCA and shows that the interference noise components of samples 1 and 4 are weakened. However, the samples present heart-rate signal is also weakened at the same time, as can be seen from samples 3 and 5. Future research will aim to provide a solution for this problem.

D. OBJECTIVE EVALUATION INDEX

A statistical method is adopted to evaluate the statistical characteristics of heart-rate values obtained by the aforementioned methods. Objective indices are utilized to evaluate our proposed heart-rate detection method quantitatively and systematically. Indices are computed to statistically analyze the difference between the reference heart-rate value obtained via pulse oximeter and the measured heart-rate values. The computation formulas used to assess the quality of the extracted pulse signal are as follows:

$$Me = \frac{1}{N}\sum_{i=1}^{N} d_i \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} d_i^2} \tag{12}$$

$$STD = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (d_i - Me)^2} \tag{13}$$

where d_i is the absolute value of the difference between reference value and measurement value of the heart rate; N is the number of heart-rate values; and *STD*, *Me*, and *RMSE* are the standard deviation, mean error, and root-mean-square error of the d_i set, respectively. Table 1 lists the statistical results for the four methods.

Table 1 presents the comparison of the performance of the discussed heart-rate detection techniques for all 85 subjects. The performance of the proposed method substantially is improved after sparse reduced-rank regression, and the statistical values of *RMSE*, average error, and standard variance are decreased.

V. CONCLUSION AND FUTURE DIRECTIONS

Operating rooms based on IoT technology utilize numerous sensors to gather physiological information from patients. These data can provide a medical reference for medical staff to improve medical services and promote the construction of smart cities. In the present study, noncontact heart-rate detection in an operating room scenario is achieved. The proposed method meets the requirements of the medical personnel and improves the comfort of the patients during operation.

Pixel variations in skin images can directly manifest variations in iPPG signals. Thus, this paper focuses on improving the signal-to-noise ratio of the iPPG signal, an approach that is beneficial for the noncontact measurement of physiological signals during surgery in surgical navigation. Blood pressure is another standard parameter that indicates vital signs, and it is monitored during the administration of surgical anesthesia. Therefore, future research should focus on the application of collected iPPG signals for monitoring blood pressure. Overall, the current research hotspot that is IoT technology is a vital aid to smart city construction. It will strengthen the interconnectedness of all facets of smart cities and continuously enhance the intelligence of urbanization.

ACKNOWLEDGMENT

(Hongwei Yue, Xiaorong Li, and Hongtao Wang contributed equally to this work.)

REFERENCES

[1] G. Piro, I. Cianci, L. A. Grieco, G. Boggia, and P. Camarda, "Information centric services in smart cities," *J. Syst. Softw.*, vol. 88, pp. 169–188, Feb. 2014.
[2] I. Azarmehr, K. Stokbro, R. B. Bell, and T. Thygesen, "Surgical navigation: A systematic review of indications, treatments, and outcomes in oral and maxillofacial surgery," *J. Oral Maxillofacial Surg.*, vol. 75, no. 9, pp. 1987–2005, Sep. 2017.
[3] W. B. Kannel, C. Kannel, R. S. Paffenbarger, Jr., and L. A. Cupples, "Heart rate and cardiovascular mortality: The framingham study," *Amer. Heart J.*, vol. 113, no. 6, pp. 1489–1494, Jun. 1987.
[4] A. Temko, "Accurate heart rate monitoring during physical exercises using PPG," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 9, pp. 2016–2024, Sep. 2017.
[5] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiolog. Meas.*, vol. 28, no. 3, p. R1, Feb. 2007.
[6] J. D'hooge, E. Konofagou, F. Jamal, A. Heimdal, L. Barrios, B. Bijnens, J. Thoen, F. van de Werf, G. Sutherland, and P. Suetens, "Two-dimensional ultrasonic strain rate measurement of the human heart *in vivo*," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 49, no. 2, pp. 281–286, Feb. 2002.
[7] A. Merla, P. A. Mattei, L. Di Donato, and G. L. Romani, "Thermal imaging of cutaneous temperature modifications in runners during graded exercise," *Ann. Biomed. Eng.*, vol. 38, no. 1, pp. 158–163, Jan. 2010.
[8] C. Takano and Y. Ohta, "Heart rate measurement based on a time-lapse image," *Med. Eng. Phys.*, vol. 29, no. 8, pp. 853–857, 2007.
[9] V. M. Banz, P. C. Müller, P. Tinguely, D. Inderbitzin, D. Ribes, M. Peterhans, D. Candinas, and S. Weber, "Intraoperative image-guided navigation system: Development and applicability in 65 patients undergoing liver surgery," *Langenbeck's Arch. Surg.*, vol. 401, no. 4, pp. 495–502, Jun. 2016.
[10] H.-H. Lin and L.-J. Lo, "Three-dimensional computer-assisted surgical simulation and intraoperative navigation in orthognathic surgery: A literature review," *J. Formosan Med. Assoc.*, vol. 114, no. 4, pp. 300–307, Apr. 2015.
[11] K. Cai, R. Yang, Q. Lin, and Z. Wang, "Tracking multiple surgical instruments in a near-infrared optical system," *Comput. Assist. Surg.*, vol. 21, no. 1, pp. 46–55, May 2016.

[12] B. Farahani, F. Firouzi, V. Chang, M. Badaroglu, N. Constant, and K. Mankodiya, "Towards fog-driven IoT eHealth: Promises and challenges of IoT in medicine and healthcare," *Future Generat. Comput. Syst.*, vol. 78, pp. 659–676, Jan. 2018.
[13] E. J. Parra, "Human pigmentation variation: Evolution, genetic basis, and implications for public health," *Amer. J. Phys. Anthropol.*, vol. 134, no. S45, pp. 85–105, Nov. 2007.
[14] M. W. Wukitsch, M. T. Petterson, D. R. Tobler, and J. A. Pologe, "Pulse oximetry: Analysis of theory, technology, and practice," *J. Clin. Monitor.*, vol. 4, no. 4, pp. 290–301, Oct. 1988.
[15] L. Kong, Y. Zhao, L. Dong, Y. Jian, X. Jin, B. Li, Y. Feng, M. Liu, X. Liu, and H. Wu, "Non-contact detection of oxygen saturation based on visible light imaging device using ambient light," *Opt. Express*, vol. 21, no. 15, pp. 17464–17471, 2013.
[16] Y. Sun, V. Azorin-Peris, R. Kalawsky, S. Hu, C. Papin, and S. E. Greenwald, "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, 2012, Art. no. 037005.
[17] Y. Sun, S. J. Hu, V. Azorin-Peris, R. Kalawsky, and S. Greenwaldc, "Noncontact imaging photoplethysmography to effectively access pulse rate variability," *J. Biomed. Opt.*, vol. 18, no. 6, 2013, Art. no. 061205.
[18] L.-M. Po, L. Feng, Y. Li, X. Xu, T. C.-H. Cheung, and K.-W. Cheung, "Block-based adaptive ROI for remote photoplethysmography," *Multimedia Tools Appl.*, vol. 77, no. 6, pp. 6503–6529, Mar. 2018.
[19] W. Wang, S. Stuijk, and G. de Haan, "Exploiting spatial redundancy of image sensor for motion robust rPPG," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 415–425, Feb. 2015.
[20] 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.
[21] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. A. 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, Mar. 2014.
[22] L. A. Aarts, V. Jeanne, J. P. Cleary, C. Lieber, J. S. Nelson, S. B. Oetomo, and W. Verkruysse, "Non-contact heart rate monitoring utilizing camera photoplethysmography in the neonatal intensive care unit—A pilot study," *Early Hum. Develop.*, vol. 89, no. 12, pp. 943–948, Dec. 2013.
[23] 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, 2011, Art. no. 077010.
[24] F. Peng, Z. Zhang, and X. Gou, H. Liu, and W. Wang, "Motion artifact removal from photoplethysmographic signals by combining temporally constrained independent component analysis and adaptive filter," *Biomed. Eng. Online*, vol. 13, no. 1, p. 50, Dec. 2014. doi: 10.1186/1475-925x-13-50.
[25] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Express*, vol. 18, no. 10, pp. 10762–10774, 2010.
[26] 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.
[27] T. Pursche, J. Krajewski, and R. Moeller, "Video-based heart rate measurement from human faces," in *Proc. IEEE Int. Conf. Consum. Electron.*, Jan. 2012, pp. 544–545.
[28] B. Chwyl, A. G. Chung, R. Amelard, J. Deglint, D. A. Clausi, and A. Wong, "SAPPHIRE: Stochastically acquired photoplethysmogram for heart rate inference in realistic environments," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 1230–1234.
[29] A. Lam and Y. Kuno, "Robust heart rate measurement from video using select random patches," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 3640–3648.
[30] G. De Haan and A. Van Leest, "Improved motion robustness of remote-PPG by using the blood volume pulse signature," *Physiolog. Meas.*, vol. 35, no. 9, pp. 1913–1926, Aug. 2014.
[31] 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.
[32] M. Khan, C. G. Pretty, A. C. Amies, R. Elliott, Y. S. Chiew, G. M. Shaw, and J. G. Chase, "Analysing the effects of cold, normal, and warm digits on transmittance pulse oximetry," *Biomed. Signal Process. Control*, vol. 26, pp. 34–41, Apr. 2016.
[33] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.
[34] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma, "Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, Dec. 2009, pp. 2080–2088.
[35] K. Cai, R. Yang, H. Chen, L. Li, J. Zhou, S. Ou, and F. Liu, "A framework combining window width-level adjustment and Gaussian filter-based multi-resolution for automatic whole heart segmentation," *Neurocomputing*, vol. 220, pp. 138–150, Jan. 2017.
[36] B. Xiao, K. Wang, X. Bi, W. Li, and J. Han, "2D-LBP: An enhanced local binary feature for texture image classification," *IEEE Trans. Circuits Syst. Video Technol.*, to be published. doi: 10.1109/tcsvt.2018.2869841.
[37] H. Tang, B. Xiao, W. Li, and G. Wang, "Pixel convolutional neural network for multi-focus image fusion," *Inf. Sci.*, vols. 433–434, pp. 125–141, Apr. 2018.
[38] J. Han, X. Ji, X. Hu, D. Zhu, K. Li, X. Jiang, G. Cui, L. Guo, and T. Liu, "Representing and retrieving video shots in human-centric brain imaging space," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2723–2736, Jul. 2013.
[39] B. Xiao, L. Li, Y. Li, W. Li, and G. Wang, "Image analysis by fractional-order orthogonal moments," *Inf. Sci.*, vols. 382–383, pp. 135–149, Mar. 2017.
[40] S. Javed, A. Mahmood, S. Al-Maadeed, T. Bouwmans, and S. K. Jung, "Moving object detection in complex scene using spatiotemporal structured-sparse RPCA," *IEEE Trans. Image Process.*, vol. 28, no. 2, pp. 1007–1022, Feb. 2019.
[41] S. E. Ebadi and E. Izquierdo, "Foreground segmentation with tree-structured sparse RPCA," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 9, pp. 2273–2280, Sep. 2018.

HONGWEI YUE received the Ph.D. degree in control theory and control engineering from the Guangdong University of Technology, China, in 2013. He is currently an Associate Professor with the Faculty of Intelligent Manufacturing, Wuyi University, China. His research interests include image processing, biomedical instrument, and information security.

XIAORONG LI received the master's degree in medical imaging and nuclear medicine from the Guangzhou University of Chinese Medicine, in 2012. He is currently the attending Doctor with the Department of Radiology, General Hospital of Southern Theater Command, PLA. His current research interests include the medical imaging diagnosis and computed-aided diagnosis.

HONGTAO WANG received the B.S. degree in automatic control from the Guangxi University of Technology, China, in 2004, the M.S. degree in control theory and control engineering from the Wuhan University of Technology, China, in 2007, and the Ph.D. degree in pattern recognition and intelligent systems from the South China University of Technology, China, in 2015. Since 2013, he has been with the Faculty of Intelligent Manufacturing, Wuyi University, as an Associate Professor. Since January 2017, he has been with the Centre for Life Sciences, Singapore Institute for Neurotechnology (SINAPSE), National University of Singapore, as a Visiting Research Fellow. His research interests include brain-like computation, pattern recognition, and hybrid intelligence.

HUAZHOU CHEN received the Ph.D. degree in applied mathematics from Shanghai University, China, in 2011. He is currently a Professor with the College of Science, Guilin University of Technology, China. His research interests include computational biology, bio- and chemo-informatics, and digitized-spectrum analytical technology.

XIAOJUN WANG received the master's degree from Nanchang University, in 1997. He is currently a Professor with Guangdong Polytechnic Normal University. His current research interests include intelligent control systems and industrial robots.

KEN CAI received the B.S. degree in measurement control technology and instruments and the M.S. degree in measuring and testing technologies and instruments from the Guangdong University of Technology, Guangzhou, China, in 2003 and 2006, respectively, and the Ph.D. degree in biomedical engineering from the South China University of Technology, Guangzhou, in 2011.

He is currently an Associate Professor with the College of Automation, Zhongkai University of Agriculture and Engineering, Guangzhou. His main research interests include machine vision and biomedical instrument.