IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 3, 2017, accepted May 19, 2017, date of publication May 24, 2017, date of current version June 27, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2707472

A Machine Learning-Empowered System for Long-Term Motion-Tolerant Wearable Monitoring of Blood Pressure and Heart Rate With Ear-ECG/PPG

QINGXUE ZHANG[1], (Member, IEEE), XUAN ZENG[2], (Member, IEEE), WENCHUANG HU[1], (Senior Member, IEEE), AND DIAN ZHOU[1,2], (Senior Member, IEEE)

[1]Department of Electrical Engineering, The University of Texas at Dallas, Richardson, TX 75080 USA
[2]Department of Microelectronics, Fudan University, Shanghai 200433, China

Corresponding author: Dian Zhou (qingxue.zhg@gmail.com)

This work was supported by the Recruitment of Global Experts (the Thousand Talents Plan), oversea collaborative scholar funding and National Natural Science Foundation of China under Grant 61574044, Grant 61376040, and Grant 61574046.

ABSTRACT In this paper, we propose a fully ear-worn long-term blood pressure (BP) and heart rate (HR) monitor to achieve a higher wearability. Moreover, to enable practical application scenarios, we present a machine learning framework to deal with severe motion artifacts induced by head movements. We suggest situating all electrocardiogram (ECG) and photoplethysmography (PPG) sensors behind two ears to achieve a super wearability, and successfully acquire weak ear-ECG/PPG signals using a semi-customized platform. After introducing head motions toward real-world application scenarios, we apply a support vector machine classifier to learn and identify raw heartbeats from motion artifacts-impacted signals. Furthermore, we propose an unsupervised learning algorithm to automatically filter out residual distorted/faking heartbeats, for ECG-to-PPG pulse transit time (PTT) and HR estimation. Specifically, we introduce a dynamic time warping-based learning approach to quantify distortion conditions of raw heartbeats referring to a high-quality heartbeat pattern, which are then compared with a threshold to perform purification. The heartbeat pattern and the distortion threshold are learned by a K-medoids clustering approach and a histogram triangle method, respectively. Afterward, we perform a comparative analysis on ten PTT or PTT&HR-based BP learning models. Based on an acquired data set, the BP and HR estimation using the proposed algorithm has an error of -1.4 ± 5.2 mmHg and 0.8 ± 2.7 beats/min, respectively, both much lower than the state-of-the-art approaches. These results demonstrate the capability of the proposed machine learning-empowered system in ear-ECG/PPG acquisition and motion-tolerant BP/HR estimation. This proof-of-concept system is expected to illustrate the feasibility of ear-ECG/PPG-based motion-tolerant BP/HR monitoring.

INDEX TERMS Wearable computers, blood pressure, heart rate, photoplethysmogram, electrocardiography, pulse transit time, fitness, signal processing, machine learning.

I. INTRODUCTION

High blood pressure (BP), also called hypertension, is a common but dangerous condition, impacting over 35 percent of people, relating to many cardiovascular, circulatory and cerebrovascular diseases, and causing 12.8 percent of total deaths worldwide [1]. Long-term BP monitoring is a key factor in hypertension control for several reasons. Firstly, hypertension is often associated with few or no symptoms which makes timely early diagnosis and treatment highly challenging without daily BP tracking. Moreover, a frequent BP measurement is significant to check up how well the treatment of hypertension is going on. Besides, BP always fluctuates over time, making the measurement at specified times and circumstances insufficient for effective BP monitoring and analysis. However, the traditional BP measurement methods usually cannot well fit the needs of long-term application scenarios, such as the invasive catheterization method sacrificing the comfortableness and the noninvasive cuff-based oscillometry approach lacking of a good wearability [2].

Wearable computers are paving a promising way for pervasive smart health wearables for people around the world, especially for developing worlds where major problems include lack of health infrastructure and limited health coverage. Wearable computers can provide health management

in a more affordable manner than traditional health services, especially when long-term continuous health data collection is needed for effective diagnosis/treatment of chronic diseases like hypertension. Many investigations in wearable BP monitoring have been reported and summarized in several recently published surveys [2], [3]. The most popular BP estimation theories are based on the fact that BP is often reversely correlated with the pulse transit time (PTT), i.e., the blood wave propagation time between two arterial sites [2]. In the arterial vessel, a higher BP usually generates a higher velocity of propagation, which results in a smaller time (i.e. PTT) for the wave to travel along the vessel, and vice versa. To measure the PTT start and end time, the electrocardiography (ECG) and photoplethysmogram (PPG) signals are the most widely used ones. The ECG heartbeat peak corresponds to the pressure wave occurrence time on the proximal site, i.e., the thoracic aorta, and thus can represent the PTT start time. The PPG heartbeat foot corresponds to when the pulse arrives the distal site, i.e., the location where the PPG sensor is placed, and thus can reflect the PTT end time. In these works, the most frequently applied ECG/PPG sensors placement methods are two-wrists/finger, chest/finger, and chest/chest. However, these placement approaches may impact the wearability and comfortableness, considering the former two require extra connection overhead or wearing more than one devices, and the last one may need a chest strap to fix the sensors and suffer from sweating. Some works [4], [5] proposed an in-ear PPG signal monitor to measure HR and other information, but they did not acquire ECG signal and measure BP. Another work [6] proposed placing the PPG sensor behind the left ear and placing two ECG electrodes behind the left ear and neck, respectively. However, the signal quality may be impacted if the collar coat touches the electrode on the neck in long-term daily applications. Moreover, this work did not evaluate the BP estimation performance after obtaining the PTT measurements, and did not consider daily movements-induced motion artifacts.

Another significant concern lacking of enough attention and study is whether BP estimation systems can tolerate to large amounts of motion artifacts [7], [8], since the body movements in long-term daily applications inevitably induce time varying skin-sensor contact variations which usually impact or even corrupt the ECG and PPG signals acquired. The accelerometers can be applied to track the motion information, which can be used as a reference for motion artifacts removal, such as discarding signal periods or adaptive filter-based motion artifact cancellation [9]. However, the acceleration information and sensor-electrode contact condition may not be well correlated, considering that the sensor-electrode contact is highly complicated and lack of effective modeling methods [10]. Moreover, there are diverse BP modeling theories and strategies being studied, to deal with the underlying complicated blood pressure wave generation and propagation mechanisms, nevertheless, the comparative analysis of major BP models is rather limited [2], [3]. One thing worth noting is that, we have previously reported a single-arm-worn ECG&PPG-based blood pressure monitor [11] which can provide a super wearability, but in this study we focus on another novel easy-wearing blood pressure monitor with novel sensor placement methods. Besides, body movements during blood pressure estimation were not considered in our previous work, however, in this study, we have made lots of efforts to deal with motion artifacts towards all-day application scenarios.

In this paper, focusing on above-mentioned challenges, we propose a novel fully ear-worn ECG&PPG-based BP and HR monitor to provide a much higher wearability, and we further present a machine learning framework to deal with many severe and random motion artifacts induced in daily applications. Firstly, to meet critical requirements on the wearability and comfortableness in long-term daily applications, a highly convenient sensors placement method is suggested which allows situating the ECG and PPG sensors all behind two ears and the possibility to integrate them into glasses or ear headsets. To the best of our knowledge, it is the first time to place ECG sensors behind two ears, which is a highly challenging non-standard single lead configuration for a super wearability purpose. Using our semi-customized hardware prototype, the weak ECG and PPG signals are successfully acquired with good morphologies.

Secondly, a machine learning-based framework is proposed for heartbeat identification from weak ear signals with large amounts of motion artifacts induced by breathing, blood vessels movements and especially, head movements (the participants were asked to perform head movements towards practical application scenarios). In the proposed machine learning-based framework, we firstly try to identify the raw heartbeats using our previously reported support vector machine (SVM)-based approach [12]. Afterwards, to filter out many residual distorted or faking raw heartbeats, we further propose an unsupervised learning approach to automatically label the heartbeat quality levels and purify the raw heartbeats. Specifically, we introduce a dynamic time warping (DTW)-based learning approach [13] which can perform non-linear mapping of two time-varying sequences, to measure the dissimilarity between each raw heartbeat and a high quality heartbeat pattern. The measured DTW dissimilarity values are used to quantify the degree of distortion of raw heartbeats and are compared with a distortion threshold, to generate the heartbeat-specific signal quality indices and perform ECG/PPG heartbeats purification. The high quality heartbeat pattern is learned using a K-medoids clustering method [14], and the distortion threshold is learned by a histogram triangle-based method [15].

The purified heartbeats are used for HR and PTT estimation. Afterwards, ten PTT-BP and PTT&HR-BP learning models [2], [3], [16]–[20] are taken into account to thoroughly compare their difference and determine an effective one for the ear application scenario. In this study, we take special interest in systolic BP (SBP) estimation. Our proof-of-concept system is validated on an acquired ear signal dataset. The proposed machine learning framework is also

compared to state-of-the-art approaches and shows a superior motion-tolerant ability, which is expected to contribute to pervasive long-term hypertension, heart health and fitness management.

II. MATERIAL AND METHOD

A. SYSTEM OVERVIEW

The proposed machine learning-enabled system is illustrated in Fig. 1. The top part (Fig. 1a) gives the customized hardware prototype and the sensors placement method for ear-ECG and PPG signals acquistion. The bottom part (Fig. 1b) shows the flow of the signal processing and HB/SBP estimation algortihm including three stages, i.e., stage I – supervised learning of heartbeat (HB) identification, stage II – unsupervised learning of signal quality labelling and signal purification, and stage III – HR estimation and supervised learning of SBP estimation. The proposed algorithms can be run on the mobile devices such as cellphones. Therefore, no computation resources will be introduced to the wearables and the form factor will not be increased. One thing worth noting is that the ML-based algorithms are expected to deal with large amounts of motions artifacts in daily applications, which is very challenging but necessary to enable 24-hour continuous blood pressure/heart rate tracking.

B. CUSTOMIZED HARDWARE PLATFORM AND SENSORS PLACEMENT

The customized hardware platform [11] shown in Fig. 1a includes two parts, i.e., the ECG [12] and PPG subsystems. In the former one, the ECG signal is acquired by an ADS1299 24-bit analog-to-digital (ADC) with a sampling rate of 500 Hz on a TI ADS1299EEG-FE evaluation board (green one) [21] and is then sent via the SPI port to a TI TivaTM C series LaunchPad (red one) [22], which is equipped with an ARM Cortex M4 microcontroller to configure the ADC and relay the signal to a PC via the USB port. In the latter one, the PPG signal is acquired by a 22-bit ADC with a sampling rate of 128 Hz on a TI AFE4490SPO2 evaluation board [23], which also owns an MSP430F5529IPN microcontroller to configure the ADC and relay the data to the PC. A higher sampling rate for ECG is based on the consideration that it is used for both HR and PTT estimation. This prototype can be conveniently used in long-term applications after removing evaluation-specific components and adding a wireless module.

The sensors placement method proposed is illustrated in the top right part of Fig. 1a, where the R/B/S correspond to the reference/bias/signal electrodes for single-lead ECG acquisition, and P represents the sensor for PPG measurement. All the sensors can be integrated into glasses or ear headsets to achieve a much higher wearability and comfortableness, compared with the chest or wrists placement.

C. DATASET RECORDING

The customized platform was applied to acquire an ear ECG/PPG dataset from fourteen subjects, to evaluate the effectiveness of the proposed proof-of-concept system. The data collection was approved by the university IRB office and the informed consent was obtained from all participants. The data collection comprises a thirty-minute training session and a thirty-minute testing session for each subject. Each session can be further split to fifteen two-minute trials. During the first seven trials, the subject stayed still to get low SBP conditions, and during each of the other eight trails, the subject rode a recumbent exercise bike in the first minute and stayed still in the second minute, to perturb the SBP to a larger range similar to the methods used in many studies [2], [3]. The reference SBP, denoted as SBPcuff, was measured on the left arm in the second minute of each trial, using an ambulatory BP monitor CONTEC ABPM50 [24]. Correspondingly, the ear signals in the second minute of each trial are used for HR and SBP estimation. The chest-ECG signal was also collected to obtain the ground truth of heartbeat occurrence time.

One thing worth noting is that deleting time periods with distortions may over-discard signal periods which are still of an acceptable signal quality. It means that there may still be a portion of good heartbeats during a signal period with distortions. So, it may be helpful to provide a high temporal resolution of BP estimates (we aim to report minute-level BP), if the good heartbeats can be extracted from all signal periods based on beat-specific quantitative distortion values.

We notice that the ear signals are frequently impacted by motion artifacts, even the users take a sitting or standing position, there are still lots of motion artifacts, such as continuous background motion artifacts due to uncontrolled neck muscle and blood vessels movements, and motion artifacts induced by little head movements such as slightly looking around or up and down. Taking practical scenarios into account, we further introduced much severer motion artifacts by performing head movements including shaking the head and nodding for one third of each signal period. Specifically, in the second minute of each trial, the subject was asked to shake the head during the first ten seconds and nod during the fourth ten seconds. These head movements corrupt a large portion of signals and make heartbeat identification much more challenging. Therefore, it is necessary to utilize lots of signal periods even distorted by motion artifacts to guarantee the high-resolution BP tracking.

D. SUPERVISED LEARNING OF HB IDENTIFICATION

The stage I of the proposed algorithm in Fig. 1b performs raw heartbeat identification from both pre-processed ear-ECG and PPG signals. Considering the ECG signal is of relatively richer signal characteristics (especially the QRS complex) than the PPG signal (detailed visualization will be given in the results section), we firstly introduce an advanced supervised machine learning approach for raw ECG heartbeats identification, based on which the raw PPG heartbeat pairs are then determined by a minima searching method [2].

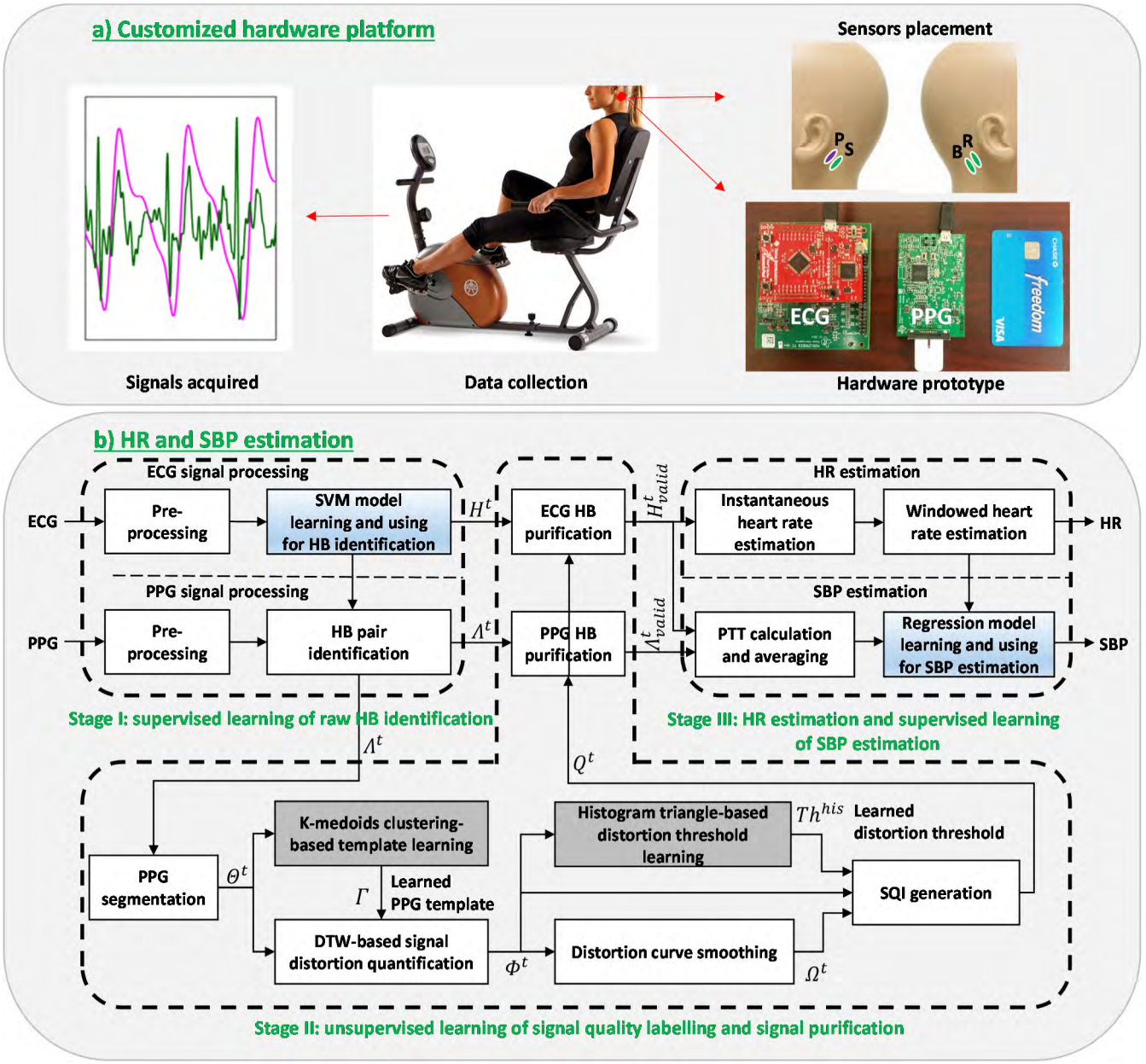

FIGURE 1. The proposed machine learning-enabled system for wearable cuff-less SBP and HR monitoring from motion artifacts-sensitive ear-ECG/PPG signals. **Algorithm block coloring in b):** white colored blocks are active in both training and testing phases; each blue colored block include both a supervised model learning process and a model using process, which are active in the training phase and the testing phase, respectively; each gray colored block includes an unsupervised model learning process and is active only in the training phase. **Abbreviations:** R/B/S represent the reference/bias/signal electrodes used for single-lead ECG signal measurement, respectively; P corresponds to the reflective PPG sensor; SVM, support vector machine; HB, heartbeat; DTW, dynamic time warping; SQI, signal quality index; PTT, pulse transit time; HR, heart rate; SBP, systolic blood pressure. **Definitions of the Greek letters:** refer to section II material and method => unsupervised learning of signal quality labelling and purification => algorithm 1.

1) SIGNAL PRE-PROCESSING

The raw ear-ECG and PPG signals are both processed by a six-order Butterworth bandpass filter (2-30 Hz and 0.5-8 Hz, respectively). Then PPG is resampled to 500 Hz to obtain a same time resolution as ECG. An example of acquired ear-ECG and PPG signals with only background motion artifacts is given in Fig. 2, where clear signal morphologies can be observed. More analysis about the signal quality with deliberately introduced severe motion artifacts will be given later.

2) ECG-BASED AND PPG-BASED HEARTBEAT IDENTIFICATION

To identify raw heartbeats from weak ear-ECG signal impacted or corrupted by large amounts of background and deliberately introduced motion artifacts, our previously reported SVM-based approach is applied [12]. Specifically, after segmenting the ECG stream to heartbeat candidates by an adaptive threshold-based auto-segmentation approach, ten critical multi-domain features are extracted from each

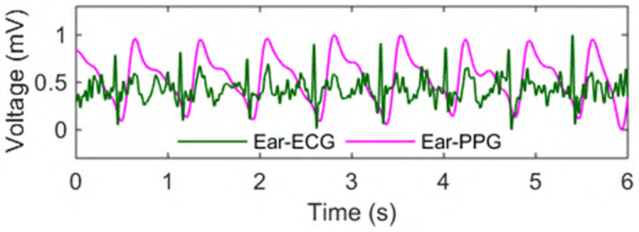

FIGURE 2. Ear-ECG/PPG Signals, with the amplitude both scaled to be between 0 and 1 for good readability (further analysis on the signal quality such as the signal strength and morphology will be given in the results section).

candidate and fed into the SVM classifier for raw heartbeat identification. The SVM model is learned based on the ear-ECG training data and tested on the ear-ECG testing data for each subject. The heartbeat locations from the chest-ECG stream collected at the same time are used as reference to label whether the ear-ECG heartbeat candidates are real heartbeats or motion artifacts-induced interferential spikes, to enable a supervised learning process.

The SVM decision function is presented by (1), where x is the ten-dimensional input vector, x_i is the $i-th$ support vector in m support vectors learned and y_i is its class label, α_i is the weight learned, b is the bias learned, and k is the kernel which is chosen as a linear one to lower computation load in wearable applications. Based on the raw ECG heartbeats, the raw PPG heartbeats (PPG feet) can be determined by searching the minimum points between each two adjacent R peaks [2].

$$f(x) = sign(\sum_{i=1}^{m} \alpha_i y_i \cdot k(x, x_i) + b) \tag{1}$$

E. UNSUPERVISED LEARNING OF SIGNAL QUALITY LABELLING AND SIGNAL PURIFICATION (SQLSP)

Considering that large amounts of back ground and head movements-induced motion artifacts usually severely corrupt a large portion of weak ear signals, many highly distorted or faking heartbeats in some signal segments need to be suppressed. Therefore, an unsupervised learning approach is proposed to automatically purify the raw heartbeats, as shown in stage II of the proposed algorithm in Fig. 1b. It can self-learn behaviors of raw PPG heartbeats, automatically generate heartbeat-specific signal quality indices and then purify the raw ECG/PPG heartbeat pairs identified by the stage I of the algorithm. Two major considerations made here include: 1) choosing an unsupervised learning strategy not a supervised one, and 2) further learning motion artifacts-sensitive behaviors of raw PPG heartbeats not ECG heartbeats.

The former one is based on the finding that it is hard to generate ground truth signal quality labels for raw heartbeats (e.g., labelled as a good or poor quality level when using a binary labelling method) which are necessary for supervised signal quality learning. Firstly, the background motion artifacts induced by uncontrolled neck muscle and blood vessels movements, especially with exercise stress, usually occur randomly making it difficult to manually label the signal quality for raw heartbeats. Moreover, the head movements deliberately introduced (shaking and nodding) to generate more critical motion artifacts cannot be strictly controlled, because: 1) it is usually difficult to precisely control the start and end time of head movements according to the data recording protocol due to the subject-dependent command response delay [25]; 2) the non-linear distortion behaviors of the raw heartbeats cannot be easily quantified, resulting from both high inter-subject and intra-subject variabilities in the swing and speed of head movements. Therefore, we propose an unsupervised learning approach to automatically label the signal quality after self-learning the diverse behaviors of raw heartbeats corrupted by motion artifacts.

Further learning motion artifacts-impacted behaviors of raw PPG heartbeats is based on the observation that the PPG signal is of less signal characteristics and thus more sensitive to motion artifacts than the ECG signal, which will be visualized in the results section. Therefore, the difference between high quality and low quality raw PPG heartbeats is more learnable for an unsupervised learner.

The detailed learning process is given in Algorithm 1. Firstly, the PPG stream is split to raw PPG heartbeat segments (s1-segmentation), which are then fed to a K-medoids clustering-based learner to determine a high quality heartbeat template (s2-template learning). Afterwards, the learned template is used to screen the raw PPG heartbeats to quantify the degree of distortion based on the DTW approach (s3-HB distortion), which can effectively measure the dissimilarity between raw heartbeats owning time varying length/morphology and the learned template using a dynamic programming strategy. The distortion values measured are used to learn by a histogram triangle-based method a distortion threshold (s4-threshold learning), which can be applied to generate binary heartbeat-specific signal quality indices for heartbeat purification purpose (s5-s8).

One thing worth noting is that the step 2) 'PPG template learning' and the step 4) 'distortion threshold learning' are only active in the training phase, and all the other steps are active in both training and testing phases. Moreover, in the testing phase, the algorithm can be performed in real time, since two 'for loops' in Algorithm 1 can be merged such that the algorithm can be executed trail by trail.

1) PPG SEGMENTATION

The PPG stream is segmented based on the raw PPG heartbeat locations identified in stage I. In the second minute of each 2-minute trial, the foot-to-foot PPG segments Θ^t are obtained as shown in step 1) 'segmentation' in Algorithm 1.

2) K-MEDOIDS CLUSTERING-BASED TEMPLATE LEARNING

As mentioned above, a PPG heartbeat template with a good morphology is needed by the DTW algorithm, to screen the raw PPG heartbeats to calculate their distortion values used in signal quality labelling. However, there is no pre-labeled signal quality information to directly perform PPG

Algorithm 1 Unsupervised Learning of SQLSP

Input:
raw PPG HB locations in T trials $\Lambda^t, \forall t \in [1, T]$
raw ECG HB locations in T trials $\mathrm{H}^t, \forall t \in [1, T]$
number of centroids in K-medoids clustering K
number of points used in distortion smoothing S
shrinkage factor used in distortion threshold learning δ
lower limit of percentage of HBs to be protected ρ
step size in percentage to adaptively adjust the threshold η
Output:
heartbeat signal quality indices in T trials $Q^t, \forall t \in [1, T]$
valid PPG HB locations $\Lambda^t_{valid}, \forall t \in [1, T]$
valid ECG HB locations $H^t_{valid}, \forall t \in [1, T]$
quality indicator for the $t-th$ BP estimate $BPQI^t, \forall t \in [1, T]$
Procedure:
for $t = 1$ to T **do** ▷ sweep trials
  $\Theta^t \leftarrow$ PPG segments split by Λ^t ▷ s1-segmentation
  if $t = 1$ **then** ▷ s2-template learning
    $L \leftarrow$ averaged length of segments in Θ^t
    $\hat{\Theta}^t \leftarrow$all segments in Θ^t resampled to a length of L
    $\Gamma \leftarrow K\_medoids\_clustering(\hat{\Theta}^t, K)$
  end if
  $\Phi^t \leftarrow \{\phi^t_j | \phi^t_j = DTW(\Gamma, \theta^t_j), \forall \theta^t_j \in \Theta^t\}$ ▷ s3-HB distortion
end for
$\Psi \leftarrow \Phi^1 \cup \Phi^2 \cup \ldots \cup \Phi^T$
$Th^{his} \leftarrow HIS\_TRI(\Psi, \delta)$ ▷ s4-threshold learning
for $t = 1$ to T **do** ▷ sweep trials
  $\Omega^t \leftarrow smooth(\Phi^t, S)$ ▷ s5-distortion smoothing
  $Q^t, BPQI^t \leftarrow SQI(\Phi^t, \Omega^t, Th^{his}, \rho, \eta)$ ▷ s6-SQI generation
  $\Lambda^t_{valid} \leftarrow$ PPG HBs validated by Q^t ▷ s7-HB purification
  $\mathrm{H}^t_{valid} \leftarrow$ ECG HBs validated by Q^t ▷ s7-HB purification
end for

template selection, otherwise, it is unnecessary to perform the unsupervised learning of signal labelling in the proposed system.

It is known that when there are more motion artifacts, there is also more morphological randomness in the raw heartbeats [12], which results in a decreasing consistency among them. If we partition the raw heartbeats into different groups according to beat-to-beat consistency, the high quality heartbeats are more likely to be clustered together benefitting a better inter-beat consistency, and the low quality heartbeats tend to be partitioned into multiple clusters due to much more diverse motion artifacts-induced morphologies. Based on this consideration, a K-medoids clustering approach is introduced to learn a good PPG template from the raw heartbeats. K-medoids clustering is a classical unsupervised machine learning algorithm which breaks the objects (raw heartbeats) up into clusters and attempt to minimize the distance (consistency) between objects belonging to a cluster and the representative object designated as the center (medoid) of that cluster [14]. Therefore, the medoid that represents a highest number of objects is selected as the high quality PPG heartbeat template.

Since the K-medoids clustering method usually applies an iterative strategy to search the final K representative medoids, to lower the computation load the PPG template is only learned from Θ^1, i.e., the raw PPG heartbeat segments in the second minute of the first trial in the training session. Moreover, also to lower the computation load, the Euclidean distance is chosen to measure inter-object distance. Since the time-varying raw heartbeats are usually of different lengths, they are all resampled to own a length of L, which is the averaged length of all raw heartbeats in Θ^1, to enable the Euclidean distance calculation. The template learning process is given in step 2) 'template learning' in Algorithm 1.

A K-means++ algorithm [26] is applied for choosing initial cluster medoid seeds, as shown in (2-3), where c_j is the $j-th$ seed to be selected from all resampled heartbeats $\hat{\theta}^1_j \epsilon \hat{\Theta}^1$ with probability $w_{\hat{\theta}^1_j}$, $d^2\left(\hat{\theta}^1_j, c_p\right)$ is the Euclidean distance between $\hat{\theta}^1_j$ and its closest pre-selected medoid $c_p, p < j$, D_p is the set of all objects closest to medoid c_p, and $\hat{\theta}^1_h \epsilon D_p$. That is, each subsequent medoid seed is selected with a probability proportional to its distance to the closest pre-selected closest seed. The number of medoid seeds K, is set as 20, considering that a smaller one may not be able to effectively separate diverse signal morphologies of raw heartbeats, and a larger one may over partition the raw heartbeats which makes multiple medoids respresent similar numbers of objects and lower the robustness of the selected template.

To solve the K-medoids problem, a partitioning around medoids (PAM) method is chosen [27], which uses a greedy search to iteratively evaluate whether the swapping of each medoid c_j and each non-medoid $\hat{\theta}^1_j$ can decrease the total object-to-medoid dissimilarity ξ as (4). If yes, then update the medoid until no dissimilarity reduction can be achieved.

$$c_j = Select(\hat{\theta}^1_j | w_{\hat{\theta}^1_j}, \forall \hat{\theta}^1_j \epsilon \hat{\Theta}^1) \tag{2}$$

$$w_{\hat{\theta}^1_j} = \frac{d^2\left(\hat{\theta}^1_j, c_p\right)}{\sum_{\{h|\hat{\theta}^1_h \epsilon D_p\}} d^2\left(\hat{\theta}^1_h, c_p\right)}, \quad \forall \hat{\theta}^1_j \epsilon \hat{\Theta}^1 \tag{3}$$

$$\xi = \sum_{\{c_p|p=1,\ldots,K\}} \sum_{\{h|\hat{\theta}^1_h \epsilon D_p\}} d^2\left(\hat{\theta}^1_h, c_p\right) \tag{4}$$

3) DTW-BASED SIGNAL DISTORTION EVALUATION

The DTW approach is introduced to screen all raw PPG heartbeats by the self-learned high quality PPG template to calculate the raw heartbeat-specific distortion value. DTW is a popular pattern learning technique firstly used in speech recognition applications and is seen as a branch of machine learning techniques. Compared with the Euclidean

distance-based method, DTW can more effectively measure the dissimilarity between two time series with time-varying morphologies and lengths [13], which is always the case when processing dynamic signals such as PPG heartbeats. DTW has an ability to warp the testing sequence in a non-linear approach to measure its distance from a pre-defined template sequence. In our study, we use this DTW distance to quantify the degree of distortion for raw PPG heartbeats, as shown in step 3) 'HB distortion' in Algorithm 1.

To measure the dissimilarity between a PPG heartbeat template $\Gamma = \{y_l | l = 0, \ldots, L-1\}$ and a raw PPG heartbeat $\theta_j^t = \{x_i^{t,j} | i = 0, \ldots, N_j^t - 1\}$, the DTW performs two steps of operation, i.e., the local distance matrix generation, and the path distance matrix generation, respectively. In the former step as (5), a $N_j^t - by - L$ local distance table is created and its each element $d_{i,l}^{t,j}$ (local distance) equals to the Euclidean distance between sample $x_i^{t,j}$ in θ_j^t and sample y_l in Γ. This table gives all possible sample-to-sample dissimilarities between θ_j^t and Γ. In the latter step as (6), a path distance table is created by a dynamic programming strategy, with the element $D_{i,l}^{t,j}$ (accumulated path distance) equaling to the current local distance $d_{i,l}^{t,j}$ plus the minima of three preceding accumulated path distances. In such a way, many warping paths are generated which all try to minimize the accumulated path distance, i.e., to find an optimal path to match the testing sequence to the template. Therefore, the last element $D_{N_j^t,L}^{t,j}$ can effectively represent the minimum accumulated sequence-to-sequence distance, i.e., the sequence level dissimilarity value, which is used as the raw heartbeat-specific distortion level as (7). Consequently, we can get quantified distortion values for raw heartbeats in each trail, denoted as Φ^t, and for all trials, denoted as Ψ, as shown in step 3) 'HB distortion' in Algorithm 1.

$$d_{i,l}^{t,j} = \left| x_i^{t,j} - y_l \right|, \quad \forall i, \ \forall l \tag{5}$$

$$D_{i,l}^{t,j} = \begin{cases} d_{i,l}^{t,j} + \min \begin{cases} D_{i-1,l}^{t,j} \\ D_{i-1,l-1}^{t,j} \\ D_{i,l-1}^{t,j} \end{cases} & \forall i > 0 \& \forall l > 0 \\ d_{i,l}^{t,j} & i = 0 \& l = 0 \\ inf & otherwise \end{cases} \tag{6}$$

$$\phi_j^t = D_{N_j^t,L}^{t,j} \tag{7}$$

4) HISTOGRAM TRIANGLE-BASED DISTORTION THRESHOLD LEARNING

Based on quantified distortion evaluation of the raw PPG heartbeats, the next is to learn an appropriate distortion threshold to differentiate heartbeats with a good or a poor quality. The same consideration used in K-medoids clustering-based template learning is applied here, i.e., low quality raw heartbeats owning much more diverse distorted morphologies due to random motion artifacts. Therefore, statistically, in a distortion histogram, the raw heartbeats with a relatively good quality should concentrate in the low distortion area (the left side of the x-axis), while the ones with gradually worse signal quality conditions usually spread over the higher distortion area (the middle and right side of the x-axis), due to poor consistency induced by random motion artifacts. Leveraging this interesting *left-skewed* histogram, we use an unsupervised learning approach called histogram triangle method [15] to learn the distortion threshold.

Based on a normalized density histogram of the distortion values, the maximum point $(b_{max}, his(b_{max}))$ is firstly determined, which corresponds to the distortion bin most frequently hit by good quality PPG heartbeats in the low distortion area, where b_{max} is the histogram bin and $his(\cdot)$ is a function returning the density value for a given bin. Then a histogram hypotenuse is constructed by connecting the maximum point $(b_{max}, his(b_{max}))$ and the right boundary of the histogram envelope $(b_{right}, his(b_{right}))$, and denote this hypotenuse as $b_{max} \sim b_{right}$. A threshold Th_{01}^{his} is learned as the bin corresponding to a maximum perpendicular distance from this bin to the hypotenuse as (8). After de-normalization based on the maxima and minima of Ψ (a set of distortion values in all trials in the training session), we get the threshold in the original scale Th_{ori}^{his} as (9), and after shrinking it by a factor of δ (50%) to enhance the robustness, we obtain the final threshold Th^{his} (10).

$$Th_{01}^{his} = \underset{b_{max} \le b \le b_{right}}{\text{argmax}} \ Dis\left\{((b, his(b)), b_{max} \sim b_{right}\right\} \tag{8}$$

$$Th_{ori}^{his} = Th_{01}^{his} \left(\max(\Psi) - \min(\Psi)\right) + \min(\Psi) \tag{9}$$

$$Th^{his} = \delta Th_{ori}^{his} \tag{10}$$

5) DISTORTION CURVE SMOOTHING

To further enhance the robustness before separating the raw PPG heartbeats to binary groups with a good or poor signal quality levels, the raw PPG heartbeats distortion values are smoothed by a 10th order moving average method ($S = 10$ in Algorithm 1). This is based on the consideration that when some raw PPG heartbeats own high distortion values, they are either real heartbeats highly corrupted by severe motion artifacts, or motion artifacts-induced interferential spikes. Therefore, their neighboring raw PPG heartbeats with lower distortion values may also have a high possibility to be impacted by motion artifacts. The smoothing operation can elevate the low distortion values for these neighboring heartbeats, and help cluster raw heartbeats in suspicious time periods to the low signal quality group more strictly.

6) SQI GENERATION

Based on the learned distortion threshold and the smoothed distortion curve, the raw PPG heartbeats can now be clustered to binary groups with a good or poor quality level. The proposed SQI generation algorithm is shown in Algorithm 2, where two prudential considerations are made to further enhance the robustness. Firstly, although the smoothed distortion curve Ω^t in the $t-th$ trial can help elevate low distortion values when they are close to high distortion values

(i.e., suspicious time periods), the smoothing operation usually lowers the high distortion values at the same time. It means that the unsmoothed distortion curve Φ^t can still contribute to highlight the heartbeats with high distortion values. Therefore, we compare not only the smoothed distortion curve but also the unsmoothed one to the learned distortion threshold Th^{his} for the SQI set generation (step 2 in Algorithm 2).

Algorithm 2 SQI Generation

$Q^t, BPQI^t \leftarrow SQI(\Phi^t, \Omega^t, Th^{his}, \rho, \eta)$
Input:
unsmoothed distortion values in the $t-th$ trial $\Phi^t = \{\phi_j^t | \forall j\}$
smoothed distortion values in the $t-th$ trial $\Omega^t = \{\pi_j^t | \forall j\}$
learned distortion threshold Th^{his}
lower limit of percentage of HBs to be protected ρ
step size in percentage to adaptively adjust the distortion threshold η
Output:
heartbeat signal quality indices in the $t-th$ trial $Q^t = \{q_j^t | \forall j\}$
quality indicator for the $t-th$ BP estimate $BPQI^t$
Procedure:
Step 1 – initialize the parameters
$P \leftarrow \rho \| \Phi^t \|$ ▷ # raw heartbeats protected
$\gamma^t \leftarrow Th^{his}$ ▷ initialize the adaptive threshold
$\Delta \leftarrow \eta \max\left(\max_{\forall j} \phi_j^t, \max_{\forall j} \pi_j^t\right)$ ▷ step size for threshold adjusting
Step 2 – generate the SQI set
$c \leftarrow 0$
for $j = 1$to $\|\Phi^t\|$ **do** ▷ sweep raw heartbeats
  if $\phi_j^t \geq \gamma^t$ **and** $\pi_j^t \geq \gamma^t$ **then** ▷ check two conditions
    $q_j^t \leftarrow 1$ ▷ a good quality heartbeat
    $c \leftarrow c + 1$
  else
    $q_j^t \leftarrow 0$ ▷ a poor quality heartbeat
  end if
end for
$BPQI^t \leftarrow c / \|\Phi^t\|$ ▷ BP quality indicator
Step 3 – adaptively adjust the threshold to protect the best P heartbeats if necessary
if $c < P$ **then** ▷ need to protect the best P heartbeats
  $\gamma^t = \gamma^t + \Delta$ ▷ increase the threshold
  go to Step 2 ▷ re-generate the SQI set
else
  stop
end if

Secondly, the motion artifacts due to time-varying electrode-skin contact variations are so random that it is impractical to cover all motion artifacts scenarios in the training session. If there happen to be some severe motion artifacts resulting in very high distortion values in the testing session, they may over-elevate many low distortion values in the corresponding suspicious time periods. Consequently, this strict SQI generation procedure may filter out too many raw heartbeats in some trials. However, based on our observation, even after aggressively introducing twenty-second head movements-induced motion artifacts in each trial (the subjects are usually asked to stay during estimation, but we aggressively asked them to perform movements for one third of each trail time) , the heartbeats corrupted are still lower than fifty percent. Leveraging this observation, we introduce a heartbeats protection strategy to protect the best ρ (20%) heartbeats in each trail, by adaptively increasing the learned threshold Th^{his} with a step size of η (1%) until at least ρ heartbeats are labeled with a good quality level (step 3 in Algorithm 2). To guarantee the consistency of the proposed SQI generation algorithm, this protection operation is also applied to the training session.

The BP estimate quality indicator is also reported to reflect the percentage of raw heartbeats left after step 2 but before heartbeat protection. The indicator can be used to select out high confident BP estimates according to specific application requirements.

After strictly labelling low quality raw PPG heartbeats and performing necessary heartbeats protection operations, the generated SQI set based on raw PPG heartbeats can now be used for heartbeats purification.

7) PPG AND ECG HEARTBEATS PURIFICATION

Considering there are still many residual highly corrupted and faking heartbeats, both raw ECG and PPG heartbeats are purified according to the SQI information, i.e., filtering raw heartbeats with an SQI of 0 and keeping those with an SQI of 1. The purified heartbeats are then sent to the stage III of the proposed HR and SBP estimation algorithm.

F. HR ESTIMATION AND SUPERVISED LEARNING OF SBP ESTIMATION

Based on the purified ECG heartbeats, the HR estimates can now be achieved, and together with purified PPG heartbeats, the PTT can also be measured. Afterwards, the SBP model can be firstly calibrated in the training session by a supervised learning process referring to the left arm cuff-based ground truth SBP, and then used for SBP estimation on the unseen data in the testing session.

1) HEART RATE ESTIMATION

As mentioned above, the ECG signal is of a relatively better motion artifacts-tolerant ability than the PPG signal, therefore, the purified ECG heartbeats are used for instantaneous heart rate estimation. Then the windowed heart rate (denoted as HR, with a unit of beats-per-minute, denoted as BPM) estimates can be achieved, where the window corresponds to the second minute in each two-minute trial during which the SBPcuff is measured. The performance of the estimated HR will be evaluated in terms of mean error ± standard deviation (ME ± STD), mean absolute error (MAE) and root mean absolute error (RMSE).

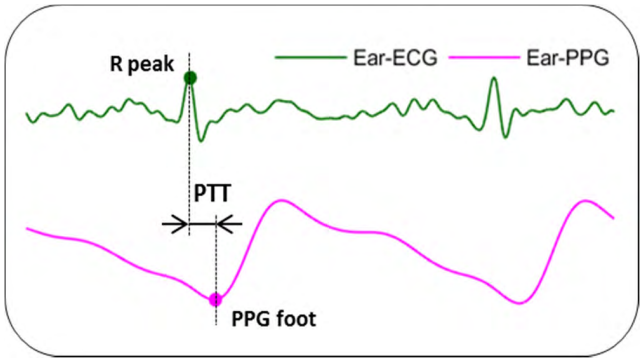

FIGURE 3. Pulse transit time (PTT) measured with ECG and PPG signals (This illustration of PTT is based on ear-ECG/PPG signals).

2) PULSE TRANSIT TIME CALCULATION

Pulse transit time is the time consumed by the pressure pulse to flow from the proximal (PTT start time) to the distal (PTT end time) arterials sites. As shown in Fig. 3, the ECG R peak represents when the pulse leaves the proximal site, i.e., the thoracic aorta, and the PPG foot corresponds to when the pulse arrives the distal site, i.e., the skin surface where the PPG sensor is placed on. The measured instantaneous PTT_i, i.e., the PTT value for the $i-th$ ECG/PPG heartbeat pair is calculated as (11), where PPG_i^{foot} is the PPG foot occurrence time and ECG_i^{Rpeak} corresponds to the ECG R peak occurrence time. Similar to windowed HR, the instantaneous PTT measured in the second minute of each trial is also averaged to obtain the windowed PTT estimates.

$$PTT_i = PPG_i^{foot} - ECG_i^{Rpeak} \tag{11}$$

3) BLOOD PRESSURE MODEL LEARNING AND TESTING

Due to the complicated underlying blood pressure wave generation and propagation mechanisms, many SBP learning models have been reported based on diverse assumptions and strategies. To thoroughly compare them and determine an appropriate one for ear application scenarios, ten popular SBP learning models are taken into account in our study as shown in Table 1, including seven PTT-SBP models, and three PTT&HR-SBP models with HR information enhanced.

Among PTT-SBP models 1 to 7, various styles of equations are applied, such as linear, quadratic, exponential ones and so on, based on different deduction processes. For example, the model 2 reflects the reverse correlation between PTT and SBP shown by large amounts of studies, based on the fact that a high SBP will reduce the time consumed by the pressure pulse to propagate from the proximal to the distal sites, and vice versa [2]. The model 7 is based on the combined action of the pulse wave and the energy of wave. Among PTT&HR-SBP models 8-10, the HR information is introduced to model establishment. They are based on the consideration that when HR increases, the cardiac output flow usually increased at the same time which causes a higher SBP, and vice versa. One thing worth noting is that, for simplicity and convenience purpose, the PPT measurement method introduced above actually includes another extra item, i.e., the pre-ejection period (PEP). PEP corresponds to the aortic valve opening time and usually significantly increases the PTT measured. PEP can be measured by adding extra hardware components, such as the phonocardiogram (PCG) sensor or the impedance cardiography (ICG) sensor [16]. Here, the PEP term is ignored for simplification purpose which is a common strategy used in many previous works [2], [3]. The HR information has been used to enhance the SBP model [18], therefore, we also consider PTT&HR-SBP models, for comparison purpose. The HR information is already carried by the ECG signal and no extra hardware components are needed. In future, new sensors can be added to measure PEP for further model enhancement.

TABLE 1. Ten blood pressure models for comparative analysis.

No.	Equation	HR information
1	$SBP = a\,ln(PTT) + b$	w/o
2	$SBP = a\,PTT^{-1} + b$	w/o
3	$SBP = a\,PTT + b$	w/o
4	$SBP = a\,PTT^2 + b\,PTT + c$	w/o
5	$SBP = a\,PTT^2 + b$	w/o
6	$SBP = a\,e^{b\,PTT}$	w/o
7	$SBP = a\,PTT^{-2} + b$	w/o
8	$SBP = a\,PTT^{-2} + b\,HR^{-2} + c$	w/
9	$SBP = a\,ln(PTT) + b\,ln(PTT) + c$	w/
10	$SBP = a\,PTT + b\,HR + c$	w/

The SBP models are learned on the training data and tested on the unseen testing data to show the generalization ability. The left-arm cuff-based SBP is used as reference to enable a supervised learning process. The performance is reported in terms of ME $\pm$ STD, MAE and RMSE.

G. PERFORMANCE COMPARISON

The proposed motion-tolerant machine learning framework is also compared with two state-of-the-art approaches, Kalman filtering (KLMF) [28] and discrete wavelet transformation (DWT) [29]. For KLMF, the raw ECG heartbeats are identified by a well-known Pan&Tompkins algorithm, and then purified by an impulse rejection filter which is designed to check the beat-to-beat interval changing trend and remove outliers due to motion artifact. The purified ECG heartbeats are used to calculate instantaneous HR (IHR) and identify PPG heartbeats for PPT calculation. Each IHR is estimated based on a prediction process and a measurement process, where the noise covariance matrix Q is set as 0.1, and the measurement noise covariance matrix is defined as $M = M_0 \cdot exp\left(1/w^2 - 1\right)$. w is to weight the current measurement ($w = 1$ for non-outliers and $w = 10^{-5}$ for outliers) and M_0 is learned to minimize the HR error on the training data. For DWT, the ECG stream is decomposed

to 8 levels based on the Daubechies6 wavelet. To remove motion artifacts, only the $3^{rd}/4^{th}/5^{th}$ detail coefficients are used to reconstruct the new signal. Then a threshold learned also by minimizing the HR error is used to detect the ECG heartbeats, which are then used in next PPG heartbeat identification, as well as HR and SBP estimation.

III. RESULTS AND DISCUSSION

In this section, both the proposed hardware prototype and the HR/SBP estimation algorithms are evaluated in detail, according to the signal processing flow shown in Fig. 1.

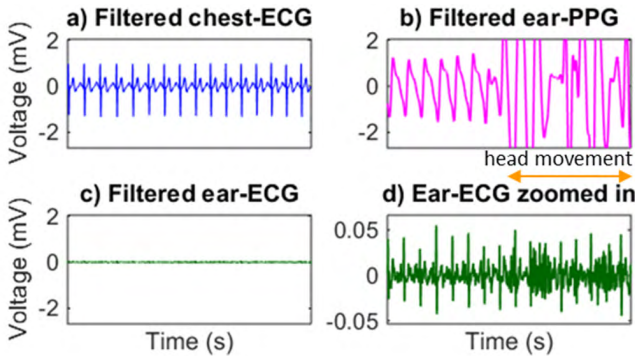

FIGURE 4. An example of the signal segments acquired (chest-ECG, ear-PPG, and ear-ECG), showing that the weak ear-ECG has a peak-to-peak voltage only around 5% of that of the chest-ECG.

A. SIGNALS ACQUIRED

After situating the ECG electrodes behind two ears and the PPG sensor behind the left ear, our semi-customized biopotential acquisition platform successfully collected the weak ear-ECG (Fig. 4c-d) and ear-PPG (Fig. 4b) signals, where the chest-ECG signal (Fig. 4a) is also given for comparison purpose. The acquired ear-ECG signal is only around 5% of the chest-ECG signal in terms of peak-to-peak voltage, resulting from a much smaller potential difference between the back locations of two ears. Although the ear-ECG is highly weak, it can still show distinguishable heartbeats, especially clear QRS complex morphologies, even with continuous background motion artifacts due to uncontrolled neck muscle and blood vessels movements, indicating the effectiveness of the proposed non-standard highly convenient single lead ECG configuration. Meanwhile, the acquired PPG signal also owns a clear heartbeat morphology leveraging many blood vessels around the back location of the ear.

When performing head movements, many motion artifacts are induced to both ear-ECG and ear-PPG signals (Fig. 4d and 4b) which make heartbeats identification highly challenging. Therefore, advanced signal processing and machine learning algorithms for robust heartbeat recognition are proposed to enable this highly wearable ear signal acquisition solution.

B. HEARTBEAT IDENTIFICATION

To identify raw heartbeats from motion artifacts-impacted or even corrupted ear signals, our previously proposed SVM-based approach is applied on the weak ear-ECG signal. Two examples of heartbeat identification results in the *testing session* of subject 1 are given in Fig. 5, where the raw ECG heartbeats are firstly identified and then the PPG heartbeats are determined by a simple minima searching method. There are several interesting observations as follows to support why we firstly identify the raw heartbeats from the ECG signal, and why we need to learn signal quality labelling based on the PPG signal in an unsupervised manner and purify both ECG/PPG heartbeats.

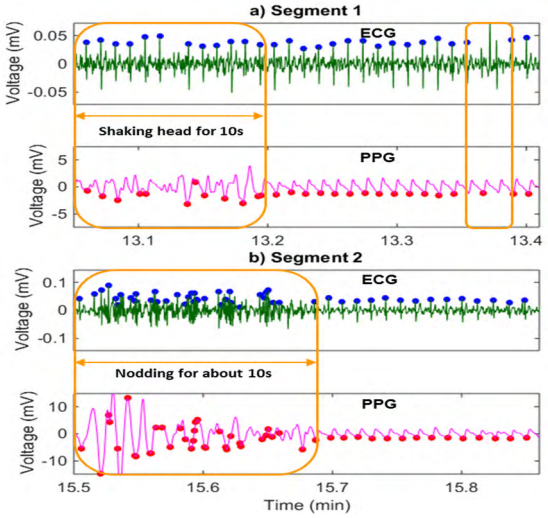

FIGURE 5. Two examples of heartbeat identification results in the testing session of the subject 1. Blue dots: identified raw ECG heartbeat locations; red dots: identified raw PPG heartbeat locations; wide orange rectangles: signal periods with deliberately introduced motion artifacts due to head movements; narrow orange rectangle: signal period with missing or fake heartbeats due to severe background motion artifacts; all the weak ECG signal is continuously impacted by background motion artifacts.

Firstly, the ECG signal acquired with the non-standard signal lead configuration is so weak that it is continuously impacted by the background motion artifacts due to uncontrolled neck muscle and blood vessels movements, as shown in Fig. 5. Specially, in the signal periods not covered by the wide orange rectangles, the PPG signal owns a better signal morphology compared with the ECG signal, which is even highly corrupted by the background motion artifacts such as the segment in the narrow orange rectangle in Fig. 5a). However, as shown in the wide rectangle (corresponding to deliberately introduced head movements) in Fig. 5a, the ECG signal actually owns a relatively better motion artifacts tolerant ability than the PPG signal levering sharply changing QRS complexes, although the ECG signal is weaker than the PPG signal. Therefore, we firstly perform raw heartbeat identification from the ECG signal.

Secondly, the PPG signal is highly distorted in the wide orange rectangles due to deliberately introduced motion artifacts by head movements, as shown in wide rectangles in Fig. 5a-b. Since PPG heartbeat morphological characteristics (PPG feet) during these signal periods cannot reflect the

heartbeat occurrence time for PTT calculation, we need to filter out these signal segments according to the signal quality of these raw PPG heartbeats.

However, learning how to label the signal quality of the raw PPG heartbeats is not trivial since it is hard to generate ground truth signal quality labels as reference in the *training session* to enable a supervised learning. This is because (using the testing data in Fig. 5 as an example here): 1) the background motion artifacts are highly random; 2) the start and end time of head movements cannot be precisely controlled due to subject-dependent command response delay as shown in the wide rectangle in segment 2 (Fig. 5b); 3) the non-linear distortion behaviors of raw PPG heartbeats is so random that it is difficult to manually generate signal quality labels. Based on these considerations, we choose an unsupervised learning approach to learn how to generate the SQI information for each raw PPG heartbeat. One thing worth noting is that the SQI information is also used to purify the ECG heartbeats which may also be corrupted at the same time when the PPG heartbeats are severely corrupted as shown in the wide rectangle in Fig. 5b.

C. PPG SEGMENTATION AND TEMPLATE LEARNING

To perform unsupervised learning of signal quality labelling, we apply a DTW method to quantify the degree of distortion for each raw PPG heartbeat. The PPG template for the DTW method is learned by the K-medoids clustering approach on the segmented raw PPG heartbeats. The clustering results are given in Fig. 6, where raw heartbeats with a relatively good quality concentrate in the M1 cluster, and raw heartbeats with a poor quality are grouped into many other clusters due to the high randomness induced by motion artifacts. Consequently, the medoid in cluster 12 which represents a highest number of instances (#=65%, i.e., 65% of raw heartbeats in the second minute in the first trial are grouped into cluster 12) is selected as the high quality PPG heartbeat template. One interesting observation is that some slightly distorted heartbeats are also grouped into this cluster since the other clusters correspond to raw heartbeats so randomly corrupted by severe motion artifacts due to head movements. The high quality PPG template can still be well leaned, because the K-medoids clustering algorithm makes each medoid represent the majority of instances in each cluster, i.e., minimizing the object-to-medoid dissimilarity ξ as (4).

D. PPG DISTORTION EVALUATION AND THRESHOLD LEARNING

After quantifying the degree of distortion for all raw PPG heartbeats using the DTW method, the histogram triangle-based approach is proposed for PPG distortion threshold learning which will be used to generate the SQI information. An example is given in Fig. 7, where a skewed intensity histogram of the DTW distances is constructed. The relatively good quality heartbeats concentrate in the low distortion area (the left side of the x-axis) and poor quality heartbeats spread over a larger range resulting from high and diverse distortion

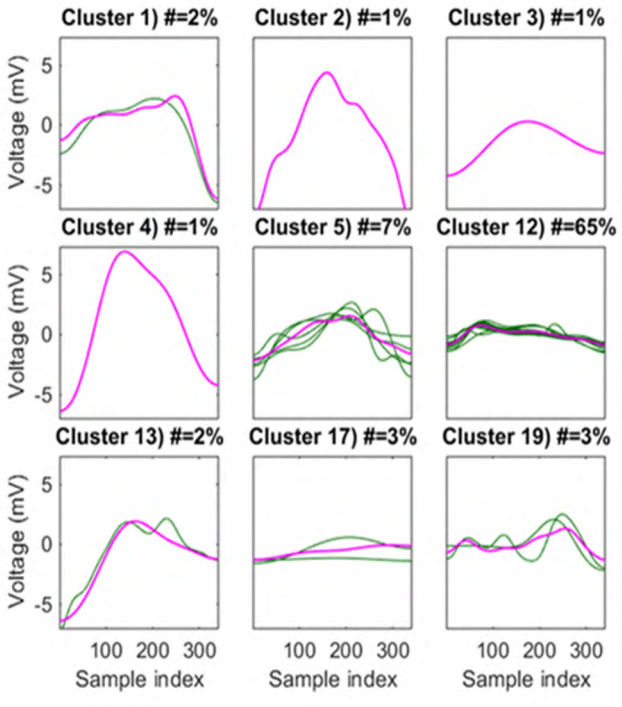

FIGURE 6. K-medoids clustering-based PPG template learning in the first trial in the training session of subject 1, with K equals to 20 (only top 9 clusters are visualized here). Pink curve: medoid of each cluster; green curves: instances represented by the medoids; #=: percentage of instances in current cluster.

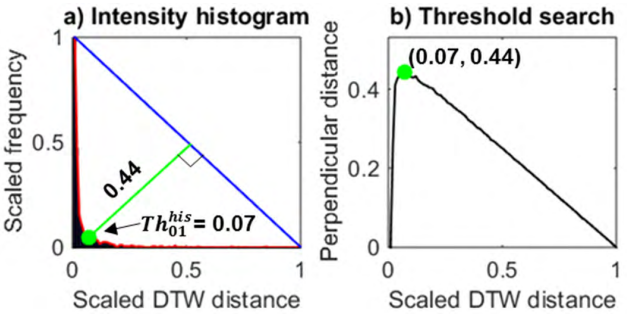

FIGURE 7. PPG distortion threshold learning in the first trial in the training session of subject 1. Blue line: the histogram hypotenuse; red curve: the histogram envelope; green line: the maximum perpendicular distance; green dot: the learned normalized (0-1 range) threshold.

values due to random and severe motion artifacts. The global searching process effectively captures the transition point of the histogram and determines the normalized threshold Th_{01}^{his} in this example as 0.07, which is then de-normalized and multiplied by a shrinkage factor to get the final distortion threshold Th^{his} which is 11.9 in this example.

E. SQI GENERATION AND HEARTBEATS PURIFICATION

Based on the DTW distance-based distortion values and the learned distortion threshold, we now can generate the SQI information for all PPG raw heartbeats in the training or the testing session. An example of the whole SQI generation

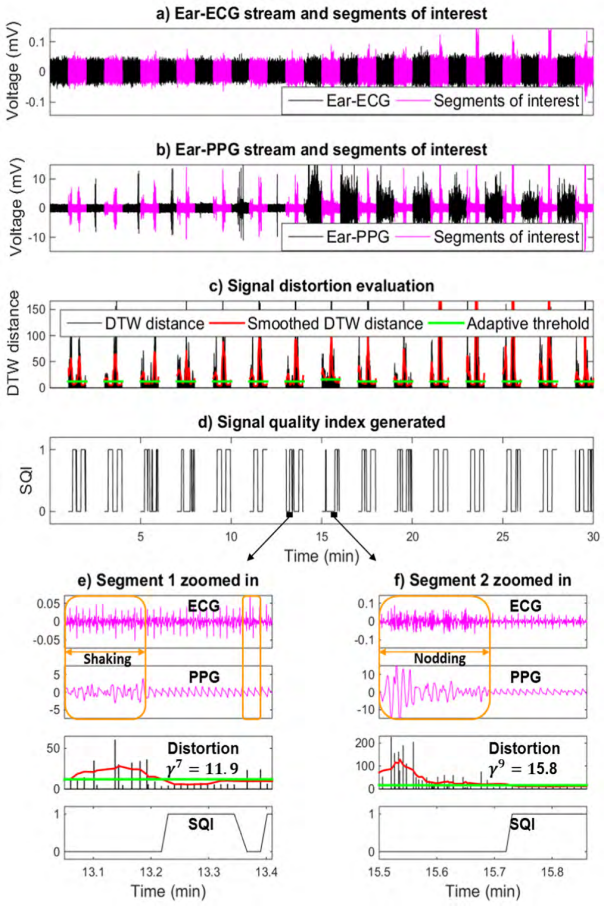

FIGURE 8. The whole SQI generation process in the testing session of subject 1, including the signals acquired, quantified degree of distortion for raw PPG heartbeats, the adaptive distortion threshold and the SQI sequence generated.

process in the testing session of subject 1 is given in Fig. 8. The acquired thirty-minute ear-ECG and ear-PPG streams are shown in Fig. 8a and 8b, where fifteen pink segments in each stream corresponding to the second minute in each of fifteen trials. In each pink segment, there is head shaking movement during the first ten seconds and nodding movement during the fourth ten seconds, resulting in many severe motion artifacts which increase the peak-to-peak voltage. In the last eight black segments, there are exercise-induced signal variations (riding the bike), especially in the ear-PPG stream. In the first seven black segments, there are also some signal variations due to normal head movements. The calculated unsmoothed DTW distance and the smoothed one are given in Fig. 8c, which shows diverse degree of distortion caused by both background and deliberately introduced motion artifacts. Based on SQI algorithm proposed, the distortion threshold is adaptively elevated if the best ρ heartbeats need to be protected. Finally, the raw PPG heartbeat-specific SQI information is generated as Fig. 8d.

Two same segments as those in Fig. 5 are used here to further illustrate the details of the SQI generation process,

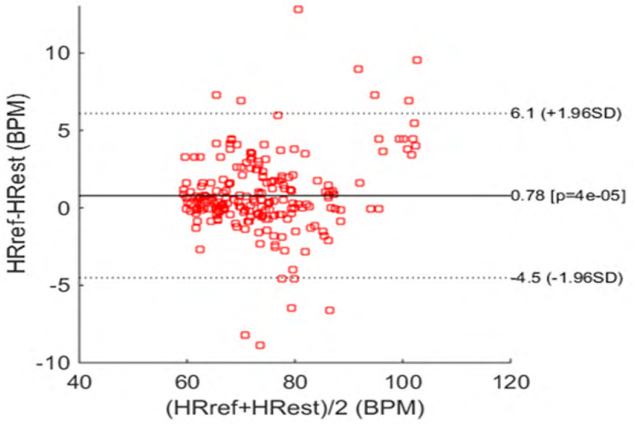

FIGURE 9. Bland-Altman plot for estimated HR (HRest) and reference HR (HRref).

TABLE 2. HR estimation performance comparison in the testing session.

	ME	STD	MAE	RMSE
This Study	-1.7	6.5	5.2	6.7
KLMF	-26.3	12.6	26.4	29.1
DWT	-42.6	19.3	42.6	46.7

Notes. Unit: BPM.

as shown in Fig. 8e and 8f. During deliberately introduced head movements (wide orange rectangles), the DTW distance-based distortion values (black bars) are much higher than those in other time periods. The smoothing operation of the distortion values can make the low distortion values above the distortion threshold and thus pose a more critical distortion evaluation during these highly suspicious periods (wide orange rectangles). On the other hand, the unsmoothed distortion values can still highlight the heartbeats with a high distortion condition (around minute 13.37). Besides, the motion artifacts are so random that it is difficult to include all motion artifacts scenarios in the training session. If there happen to be some highly random motion artifacts shown as in Fig. 8f, the corrupted PPG morphologies may induce dramatically high distortion values which generate a large range of suspicious period. This shows the necessity to introduce a protection mechanism to adaptively elevate the distortion threshold for some trials to protect the best ρ (20%) heartbeats (e.g., threshold $\gamma^9 = 15.8$ in trial 9 is adaptively adjusted to be higher than threshold $\gamma^7 = Th^{his} = 11.9$ in trial 7). Based on the proposed SQI generation algorithm, the raw heartbeats are labelled as accepted (SQI = 1) or rejected (SQI = 0), which helps filter out many signal segments highly corrupted by severe motion artifacts such that the remaining purified heartbeats can be used in HR and PTT estimation later.

F. HEART RATE ESTIMATION

The windowed heart rate estimates are achieved based on the ECG heartbeats. A Bland-Altman plot for estimated and

TABLE 3. SBP estimation performance comparison in the testing session.

SBP Model No.	This study				KLMF				DWT			
	ME	STD	MAE	RMSE	ME	STD	MAE	RMSE	ME	STD	MAE	RMSE
1	-1.7	6.5	5.2	6.7	-3.0	7.7	6.2	8.3	3.9	7.4	6.2	8.4
2	-1.6	6.6	5.3	6.8	-3.0	7.2	6.0	7.8	4.5	8.0	6.7	9.2
3	-1.8	6.1	5.0	6.4	-3.7	7.3	6.3	8.2	3.0	6.6	5.6	7.3
4	-2.1	6.6	5.4	6.9	-3.3	10.1	7.4	10.6	4.0	12.6	9.8	13.2
5	-1.9	6.4	5.2	6.7	-4.2	7.8	6.6	8.9	2.5	6.8	5.3	7.2
6	-1.8	6.1	5.0	6.3	-3.5	7.1	6.2	7.9	3.3	6.9	5.9	7.6
7	-2.0	6.3	5.2	6.5	-2.8	6.7	5.7	7.2	5.4	8.5	7.7	10.1
8	**-1.4**	**5.2**	**4.2**	**5.4**	1.7	6.9	5.2	7.1	6.7	10.8	8.2	12.7
9	**-1.2**	**5.0**	**4.1**	**5.2**	2.4	7.5	5.7	7.8	10.4	14.9	11.5	18.1
10	**-1.1**	**5.1**	**4.0**	**5.1**	2.8	7.9	6.0	8.3	12.9	18.6	14.0	22.6

reference HR is given in Fig. 9 to illustrate the HR estimation performance. It shows that most of the HR estimates concentrate in the low error area, indicating the potential of using ear-ECG for robust long-term HR monitorting applications. Averaged over the acquired ear signal dataset, the ME±STD, MAE and RMSE of HR estimation are 0.8±2.7, 1.8 and 2.8 BPM, respectively.

In table 2, the HR estimation performance using our approach greatly outperforms KLMF and DWT. For KLMF, a smoothing operation is applied to suppress the abnormal measurements based on the outlier indicators generated by the impulse rejection filter. However, the robustness of the outlier indicators is still very low due to the fact that the beat-to-beat checking rules used in this filter cannot effectively cover highly random cases due to large amounts of motion artifacts. For DWT, the low performance mainly suffers from the fact that the motion artifacts usually own a frequency spectrum highly overlapping that of the ECG signal. Therefore, there are still many residual motion artifacts in the reconstructed signal which lower the HR estimation performance.

G. BLOOD PRESSURE ESTIMATION

Based on HR and PTT estimates, ten diverse SBP modles including seven PTT-SBP models (1-7) and three PTT&HR-SBP models (8-10) are thoroughly compared to explore their abilities in SBP estimation. According to the *Advancement of Medical Instrumentation (AAMI) standard* [30], the BP estimation error should be less than 5.0 ± 8.0 mmHg in terms of mean error (ME) $\pm$ standard deviation (SD). To thoroughly evaluate the SBP estimation performance, we consider four different criteria including ME, STD, MAE and RMSE. Moreover, although many wearable BP estimation studies only reported the performance on the training data [2], [3], we test the proposed algorithm on the unseen testing data to emphasize the generalization ability of the learned SBP models.

The performance comparison of ten SBP models and three signal processign approaches on the unseen testing data is summarized in Table 3. Using the proposed framework, model 8-10 effectively outperform model 1-7 leveraing the additionally introduced robust HR information, with the ME±STD, MAE and RMSE no more than -1.4 ± 5.2, 4.2 and 5.4 mmHg, respectively. Compared with our framework, MLMF and DWT both show a worse performance due to the reason mentioned above, i.e., many residual motion artifacts. One thing worth noting is that model 8-10 for KLMF and DTW may even own worse performance than model 1-7, due to the introduction of low robust HR information into the BP models.

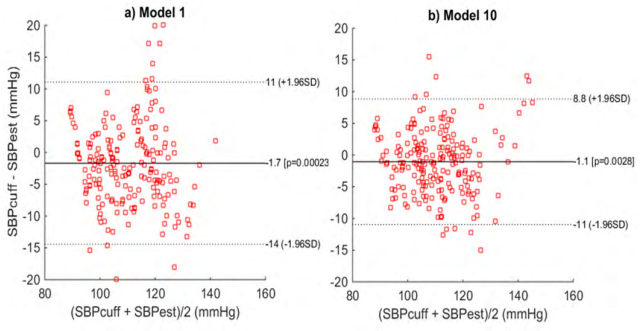

FIGURE 10. Bland-Altman plots for SBP model 1 and 10 based on our framework.

To further illustrate the performance difference between the PTT-SBP models and the PTT&HR-SBP models, the Bland-Altman plots for model 1 and 10 based on the proposed machine learning framework are given in Fig. 10. The latter one owns a smaller mean error and a more concentrated distribution (a smaller standard deviation) compared to the former one, indicating that the PTT-SBP model can be enhanced by the HR information, yielding a more robust PTT&HR-SBP model. On thing worth noting is that we have introduced

exercise to perturb the BP to make the trained SBP model be able to cover a large range of BP (similar to many studies [2], [3]), and we did found strong correlations between heart rate and SBP (similar to the previous study [18]), but more data acquisition protocols are also necessary considering that the relation between heart rate and SBP may need to be further explored.

IV. CONCLUSION

In this paper, we propose a fully ear-worn system to provide a high wearability for long-term cuff-less blood pressure and heart rate monitoring, and further present a machine learning-based signal processing framework which owns a high motion-tolerant ability to enable daily applications. Firstly, we proposed a sensor placement approach to meet the critical requirements on the wearability and comfortableness, which allows for situating all sensors behind two ears and the possibility to integrate them into glasses or ear headsets. We successfully acquired the weak ear-ECG and ear-PPG signals with our semi-customized bio-potential acquisition platform, and explored diverse background and head movements-induced motion artifacts towards practical application scenarios. Secondly, to deal with many severe motion artifacts, the raw heartbeats are identified using the SVM classifier, and then distorted or faking raw heartbeats are automatically purified by an unsupervised learning algorithm. The DTW approach is introduced to quantify the degree of distortion for raw heartbeats, referring to a high quality heartbeat template, which are then compared with a distortion threshold to generate SQI information for heartbeat purification. The high quality heartbeat template and the distortion threshold are learned using the K-medoids clustering approach and the histogram triangle method, respectively. The purified ECG and PPG heartbeats are used for HR and PTT estimation, which are then applied for robust SBP estimation.

Applying the proposed machine learning-based signal processing framework to an acquired ear signal dataset, the ME$\pm$STD, MAE and RMSE for HR estimation are 0.8$\pm$2.7, 1.8 and 2.8 BPM, respectively, and for PTT&HR-SBP models, they are no more than -1.4 ± 5.2, 4.2 and 5.4 mmHg, respectively. The performance greatly outperforms state-of-the-art works. This study is expected to demonstrate the feasibility of the proof-of-concept system in wearable ear-ECG/PPG acquisition and motion-tolerant BP/HR estimation, to enable pervasive hypertension, heart health and fitness management. In future, we will acquire data from more subjects, and also further introduce motion artifacts from more scenarios, such as walking, running, sleeping, eating, etc. Another interesting work is to enhance the power efficiency of this easy-wearing blood pressure system for long-term wearable application scenarios [31]–[33].

REFERENCES

[1] WHO. *Raised Blood Pressure*. [Online]. Available: http://www.who.int/gho/ncd/risk_factors/en/
[2] R. Mukkamala *et al.*, "Toward ubiquitous blood pressure monitoring via pulse transit time: Theory and practice," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 1879–1901, Aug. 2015.
[3] D. Buxi, J.-M. Redouté, and M. R. Yuce, "A survey on signals and systems in ambulatory blood pressure monitoring using pulse transit time," *Physiol. Meas.*, vol. 36, pp. R1–R26, Mar. 2015.
[4] S. Vogel, M. Hulsbusch, T. Hennig, V. Blazek, and S. Leonhardt, "In-ear vital signs monitoring using a novel microoptic reflective sensor," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 882–889, Nov. 2009.
[5] B. Venema, J. Schiefer, V. Blazek, N. Blanik, and S. Leonhardt, "Evaluating innovative in-ear pulse oximetry for unobtrusive cardiovascular and pulmonary monitoring during sleep," *IEEE J. Transl. Eng. Health Med.*, vol. 1, 2013, Art. no. 2700208.
[6] D. Da He, E. S. Winokur, and C. G. Sodini, "An ear-worn vital signs monitor," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 11, pp. 2547–2552, Nov. 2015.
[7] Y.-T. Zhang, C. C. Y. Poon, C.-H. Chan, M. W. W. Tsang, and K.-F. Wu, "A health-shirt using e-textile materials for the continuous and cuffless monitoring of arterial blood pressure," in *Proc. 3rd IEEE/EMBS Int. Summer School Med. Devices Biosensors*, Sep. 2006, pp. 86–89.
[8] A. Jadooei, O. Zaderykhin, and V. Shulgin, "Adaptive algorithm for continuous monitoring of blood pressure using a pulse transit time," in *Proc. IEEE 33rd Int. Sci. Conf. Electron. Nanotechnol. (ELNANO)*, Apr. 2013, pp. 297–301.
[9] S.-H. Liu, "Motion artifact reduction in electrocardiogram using adaptive filter," *J. Med. Biol. Eng.*, vol. 31, no. 1, pp. 67–72, 2011.
[10] Y. M. Chi, T.-P. Jung, and G. Cauwenberghs, "Dry-contact and noncontact biopotential electrodes: Methodological review," *IEEE Rev. Biomed. Eng.*, vol. 3, pp. 106–119, Oct. 2010.
[11] Q. Zhang, D. Zhou, and X. Zeng, "Highly wearable cuff-less blood pressure and heart rate monitoring with single-arm electrocardiogram and photoplethysmogram signals," *Biomed. Eng. OnLine*, vol. 16, no. 1, p. 23, 2017.
[12] Q. Zhang, D. Zhou, and X. Zeng, "A novel machine learning-enabled framework for instantaneous heart rate monitoring from motion-artifact-corrupted electrocardiogram signals," *Physiol. Meas.*, vol. 37, no. 11, pp. 1945–1967, 2016.
[13] X. Long, P. Fonseca, J. Foussier, R. Haakma, and R. M. Aarts, "Sleep and wake classification with actigraphy and respiratory effort using dynamic warping," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 4, pp. 1272–1284, 2014.
[14] T. Velmurugan and T. Santhanam, "Computational complexity between K-means and K-medoids clustering algorithms for normal and uniform distributions of data points," *J. Comput. Sci.*, vol. 6, no. 3, pp. 363–368, 2010.
[15] G. W. Zack, W. E. Rogers, and S. A. Latt, "Automatic measurement of sister chromatid exchange frequency," *J. Histochem. Cytochem.*, vol. 25, no. 7, pp. 741–753, 1977.
[16] S. Mazaheri and E. Zahedi, "A comparative review of blood pressure measurement methods using pulse wave velocity," in *Proc. IEEE Int. Conf. Smart Instrum., Meas. Appl. (ICSIMA)*, Nov. 2014, pp. 1–5.
[17] S. Mottaghi, M. Moradi, and L. Roohisefat, "Cuffless blood pressure estimation during exercise stress test," *Int. J. Biosci., Biochem. Bioinform.*, vol. 2, no. 6, p. 394, 2012.
[18] F. S. Cattivelli and H. Garudadri, "Noninvasive cuffless estimation of blood pressure from pulse arrival time and heart rate with adaptive calibration," in *Proc. 6th Int. Workshop Wearable Implant. Body Sensor Netw.*, 2009, pp. 114–119.
[19] R. Payne, C. Symeonides, D. Webb, and S. Maxwell, "Pulse transit time measured from the ECG: An unreliable marker of beat-to-beat blood pressure," *J. Appl. Physiol.*, vol. 100, no. 1, pp. 136–141, 2006.
[20] T. Pereira, R. Sanches, P. Reis, J. Pego, and R. Simoes, "Correlation study between blood pressure and pulse transit time," in *Proc. IEEE 4th Portuguese Meeting Bioeng. (ENBENG)*, Feb. 2015, pp. 1–5.
[21] Texas Instrument. (2012). *ADS1299EEG-FE Evaluation Board*. [Online]. Available: http://www.ti.com/tool/ads1299eegfe-pdk
[22] Texas Instrument. (2013). *Tiva C Series LaunchPad*. [Online]. Available: http://www.ti.com/tool/ek-tm4c123gxl
[23] Texas Instrument. (2014). *AFE4490SPO2 Evaluation Board*. [Online]. Available: http://www.ti.com/tool/afe4490spo2evm
[24] CONTEC. *ABPM50*. [Online]. Available: http://www.contecmed.com/
[25] A. Pentland and A. Liu, "Modeling and prediction of human behavior," *Neural Comput.*, vol. 11, no. 1, pp. 229–242, 1999.

[26] D. Arthur and S. Vassilvitskii, "k-means++: The advantages of careful seeding," in *Proc. 18th Annu. ACM-SIAM Symp. Discrete Algorithms*, 2007, pp. 1027–1035.
[27] L. Kaufman and P. J. Rousseeuw, "Partitioning around medoids (program PAM)," in *Finding Groups in Data: An Introduction to Cluster Analysis*. New York, NY, USA: Wiley, 1990, pp. 68–125. [Online]. Available: http://www.redlightgreen.com/title/finding-groups-in-data-an-introduction-to-cluster-analysis/oclc/800813602?referer=br&ht=edition
[28] A. V. Gribok, X. Chen, and J. Reifman, "A robust method to estimate instantaneous heart rate from noisy electrocardiogram waveforms," *Ann. Biomed. Eng.*, vol. 39, pp. 824–834, Feb. 2011.
[29] S. Mukhopadhyay, S. Biswas, A. B. Roy, and N. Dey, "Wavelet based QRS complex detection of ECG signal," *Int. J. Eng. Res. Appl.*, vol. 2, pp. 2361–2365, Sep. 2012.
[30] AAMI. (2007). *ANSI/AAMI SP10-1992 Standard*. [Online]. Available: http://www.aami.org/
[31] J. Birjandtalab, Q. Zhang, and R. Jafari, "A case study on minimum energy operation for dynamic time warping signal processing in wearable computers," in *Proc. IEEE Int. Conf. Pervas. Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2015, pp. 415–420.
[32] J. Eriksson *et al.*, "PerMoby'15: The fourth IEEE international workshop on the impact of human mobility in pervasive systems and applications, 2015—Program," in *Proc. IEEE Int. Conf. Perv. Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2015, pp. 1–2.
[33] Q. Zhang, C. Zahed, V. Nathan, D. A. Hall, and R. Jafari, "An ECG dataset representing real-world signal characteristics for wearable computers," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Oct. 2015, pp. 1–4.

QINGXUE ZHANG (M'14) received the B.S. and M.S. degrees in science and technologies of electronics from Xi'an Jiaotong University, Xi'an, China, in 2004 and 2007, respectively, and the Ph.D. degree in electrical engineering from The University of Texas at Dallas, Richardson, TX, USA, in 2017. From 2007 to 2013, he was a System Architect and a Team Lead with Hisilicon Co., Ltd., in Huawei. His current research interests include smart health and wearable computers, with emphasis on wearable prototyping, signal processing/machine learning algorithms, and emerging applications.

XUAN ZENG (M'97) received the B.S. and Ph.D. degrees in electrical engineering from Fudan University, Shanghai, China, in 1991 and 1997, respectively. She was a Visiting Professor with the Department of Electrical Engineering, Texas A&M University, College Station, TX, USA, and the Department of Microelectronics, Technische Universiteit Delft, Delft, The Netherlands, in 2002 and 2003, respectively. She is currently a Full Professor with the Department of Microelectronics, Fudan University, where she was the Director of the State Key Laboratory of ASIC & System, from 2008 to 2012. Her current research interests include smart city, smart health, machine learning, big data, design for manufacturability, high-speed interconnect analysis and optimization, analog behavioral modeling, circuit simulation, and ASIC design.

Dr. Zeng was a recipient of the Chinese National Science Funds for Distinguished Young Scientists in 2011 and the First-Class of Natural Science Prize of Shanghai in 2012. She is the Changjiang Distinguished Professor with the Ministry of Education Department of China in 2014. She is currently the Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II.

WENCHUANG HU received the B.S. degree from Peking University, China, in 1999, and the Ph.D. degree from the University of Notre Dame, IN, in 2004. He was a Post-Doctoral Research Fellow with the Department of Electrical Engineering, University of Michigan, Ann Arbor, MI. In 2005, he joined The University of Texas at Dallas and currently an Associate Professor of the Department of Electrical Engineering.

His research has been focused on biosensors, bioelectronics, nanolithography, nanofabrication, and applications in medical, semiconductors, and renewable energy. He is a Senior Member of the IEEE, a member of Sigma Xi, AVS, MRS, ACS, and SPIE.

DIAN ZHOU (M'89–SM'07) received the B.S. and M.S. degrees in physics from Fudan University, Shanghai, China, in 1982 and 1985, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 1990. He is currently a Professor with The University of Texas at Dallas, Richardson, TX, USA. His current research interests include smart city, smart health, machine learning, big data, internet of things, VLSI design, circuit and systems, CAD tools, and algorithms.

Prof. Zhou was a recipient of the IEEE Circuits and Systems Society Darlington Award in 1993, the National Science Foundation (NSF) Young Investigator Award in 1994, the Outstanding Overseas Young Investigator Award from NSF of China in 2000, the Changjiang Honor Professor from the Ministry of Education Department of China in 2002, and the Grand Award HDTV Terrestrial Broadcasting ASIC Chip: Chinese Television Number One from the Electronics Society of China in 2005, and selected as the Thousand People Plan Professor from China in 2010.