Photoplethysmography-Based Heart Rate Monitoring Using Asymmetric Least Squares Spectrum Subtraction and Bayesian Decision Theory

Biao Sun and Zhilin Zhang, *Senior Member, IEEE*

***Abstract*— Motion artifacts (MAs) are strong interference sources in wearable photoplethysmography (PPG) signals, significantly affecting estimation of heart rate (HR) and other physiological parameters. In this paper, a novel method called SPECTRAP is proposed for accurate motion-tolerant estimation of HR using a PPG signal and a simultaneous acceleration signal. The method first calculates the spectra of the PPG signal and the acceleration signal, and then removes the MA spectral components from the PPG spectrum using a new spectrum subtraction algorithm. The new spectrum subtraction algorithm is based on asymmetric least square and overcomes drawbacks of the conventional spectrum subtraction algorithms. To find the spectral peak corresponding to HR on the resulting spectrum, SPECTRAP formulates the problem into a pattern classification problem and uses the Bayesian decision theory to solve it. Experimental results on the PPG database used in 2015 IEEE Signal Processing Cup showed that the proposed algorithm has excellent performance. The average absolute error on the twelve training sets was 1.50 beat per minute (BPM) (standard deviation: 1.95 BPM). The average absolute error on the ten testing sets was 2.13 BPM (standard deviation: 2.77 BPM).**

***Index Terms*— Photoplethysmography (PPG), heart rate, motion artifacts, asymmetric least squares, spectrum subtraction, Bayesian decision theory.**

I. Introduction

A. Background and Motivation

HEART rate (HR) monitoring during fitness is important for enabling exercisers to control their training load. Traditionally, people monitor their HR during exercise using Electrocardiogram (ECG). However, this technique is not convenient in practice, because it requires multiple sensors (such as the reference sensor) attached to the chest, and does not provide enough user comfortability. Recently, photoplethysmography (PPG) becomes a popular option for HR monitoring, because PPG sensors are simple, low-cost, easy-to-use, and do not require specific techniques to be attached to skin [1]. Thus, many modern wearable devices, such as smart watches and smart wristbands, use PPG to estimate HR in real time for fitness tracking and health monitoring.

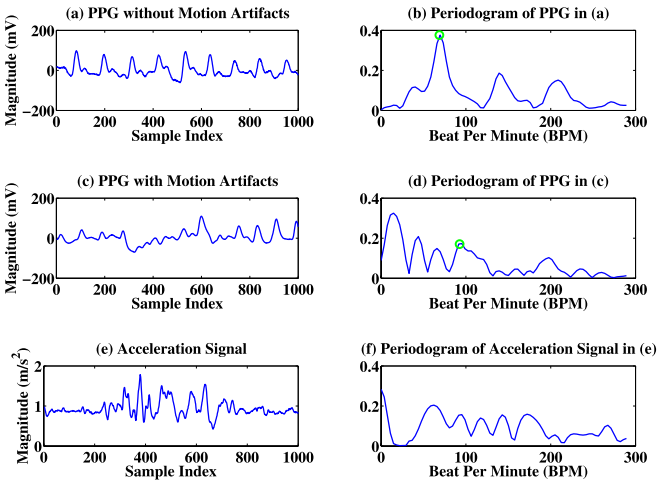

Fig. 1. An example of a segment of clean PPG and PPG with MA. Plot (a) and Plot (b) show the waveform and the spectrum (calculated by Periodogram) of a clean PPG signal, respectively. Plot (c) and Plot (d) show the waveform and the spectrum (calculated by Periodogram) of a PPG signal with strong MA, respectively. Plot (e) shows the simultaneous acceleration signal when recording the PPG signal in (c). Plot (f) shows the spectrum (calculated by Periodogram) of the acceleration signal. The green circles in Plot (b) and Plot (d) indicate the spectrum peaks corresponding to HR (calculated from simultaneously recorded ECG not shown here). Plot (d) shows that finding the spectrum peak corresponding to HR is difficult.

Despite the attractive attributes of PPG, a major disadvantage is that PPG is sensitive to motion artifacts (MA). Even a slight body movement of a wearer can introduce MA in the recorded PPG signal. Figure 1 gives an example of a clean PPG and an MA-contaminated PPG. However, removal of MA in raw PPG signals is always challenging. Therefore, there is a high demand for developing effective MA removal methods.

B. State of the Art

To date, researchers have developed numerous approaches to removing MA in PPG, such as wavelet-based methods [2]–[4], non-linear methods [5], time-frequency analysis [6], Kalman filtering [7], and empirical mode decomposition [8], [9].

Manuscript received April 20, 2015; revised July 27, 2015; accepted August 17, 2015. Date of publication August 26, 2015; date of current version October 15, 2015. This work was supported in part by the Ph.D. Programs Foundation through the Ministry of Education of China under Grant 20120032110068, in part by the Tianjin Key Technology Research and Development Program under Grant 14ZCZDS F00025, and in part by the National Natural Science Foundation of China under Grant 61271321, Grant 61401303, and Grant 61473207. The associate editor coordinating the review of this paper and approving it for publication was Prof. M. R. Yuce. *(Corresponding author: Zhilin Zhang.)*

B. Sun is with the School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China (e-mail: sunbiao@tju.edu.cn).

Z. Zhang is with Samsung Research America–Dallas, Richardson, TX 75082 USA (e-mail: zhilinzhang@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSEN.2015.2473697

Adaptive filtering is a popular approach to remove MA, provided that a reference signal is available. The reference signal can be acquired with extra hardware such as accelerometers [10]–[15], photoelectric devices [16], or another PPG sensor [17]–[19].

Spectrum subtraction (SS) [20] is another widely adopted approach to remove the spectrum of acceleration data from the spectrum of a PPG, thus suppressing the effects of MA [9], [21]. This technique has low computational complexity and can be easy to implement, but may not be robust and effective in some practical scenarios [20].

Recently, Zhang *et al.* [22] and Zhang [23] proposed a novel method named TROIKA, which consists of three stages, namely, signal decomposition, sparse signal recovery-based spectrum estimation, and spectral peak tracking with verification. The signal decomposition stage is used to partially remove MA in raw PPG signals. The sparse signal recovery-based spectrum estimation is used to overcome some drawbacks of conventional power spectrum estimation in PPG spectrum estimation, facilitating the finding of spectral peaks corresponding to heartbeat in the third stage. Later, the method was improved by using an advanced sparse signal recovery model and SS [24]. The main drawback of these methods is that the PPG spectrum calculation often requires heavy computational load.

C. Contributions and Paper Organization

This study aims to improve HR estimation performance while reducing computational load. A novel algorithm which consists of SPECtrum subtraction, peak TRAcking, and Post-processing is proposed, denoted by SPECTRAP. Its first part is based on the SS framework. A tri-axis simultaneous acceleration signal is used as a noise reference signal, and an asymmetric least squares spectrum subtraction (ALS-SS) algorithm is developed to remove MA from PPG, which effectively makes the PPG power spectrum cleaner and smoother. In the second stage, a spectral peak selection approach is proposed to find the heartbeat frequency, which is based on the Bayesian decision theory. In this approach, both the statistical properties of PPG signals and a priori knowledge about the characteristic of spectral peaks of interest are taken into account. Finally, to prevent incorrect estimations of the HR, a post-processing stage is performed.

The main contributions of the work are as follows.

1) Instead of using conventional SS methods, an asymmetric least square method is proposed to remove MA spectral components from spectra of raw PPG signals, overcoming drawbacks of conventional SS methods. Using the method, false spectral peaks, which are often generated by conventional SS methods, are largely reduced.
2) The spectral peak selection problem is formulated into a pattern classification problem, and the Bayesian decision theory is used to find the spectral peaks corresponding to HR. To the best of our knowledge, most existing algorithms [9], [21], [22], [24] used heuristic methods to find the spectral peaks corresponding to HR, which involve many user-tuning parameters. This is the first time that the spectral peak selection problem is solved by using a pattern classification method.

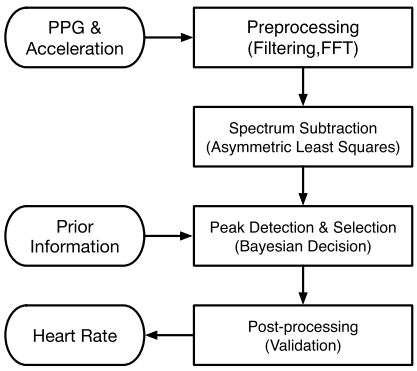

Fig. 2. Block diagram of the framework for HR monitoring. A PPG signal, a simultaneous acceleration signal and prior information about the true HR peak are the input. Estimated HR is the output. Spectrum subtraction and peak detection/selection are the two key stages. The framework also contains pre- and post-processing stages.

The rest of this paper is organized as follows. Section II introduces the proposed method. The experimental set-up and validation of the algorithm are presented in Section III. Conclusion is drawn in the last section.

II. Proposed Method

The proposed method estimates HR by using a raw PPG signal and a simultaneous acceleration signal. The method contains three main stage, i.e., an SS stage to remove MA, a spectral peak selection stage to estimate HR, and a post-processing stage to further refine the estimated HR. The method is demonstrated in Fig 2.

A. Preprocessing

A time window is used to slide the simultaneous PPG and acceleration signals, and HR is estimated in the time window. The time window is T seconds long with an incremental step of S seconds. In our experiments, T and S were set to 8 seconds and 2 seconds, respectively.

Generally, healthy adults' HR varies from 0.5 to 3.0 Hz. Therefore, in the pre-processing step, the PPG and acceleration signals are processed by a bandpass filter with a low cut-off frequency and a high cut-off frequency of 0.5 Hz and 3.0 Hz, respectively. After the filtering procedure, most noise and MA outside the frequency band are removed.

The power spectra of the PPG and acceleration signals in the current window are then calculated by Periodogram [25]. Each power spectrum is normalized by dividing by its maximum value. In our experiments, the acceleration signal is a tri-axis signal, and thus the power spectrum of acceleration along each axis was calculated separately.

B. MA Reduction by Spectrum Subtraction Based on Asymmetric Least Squares

A motion not only changes the acceleration signal but also affects the simultaneous PPG signal. Thus, the spectral

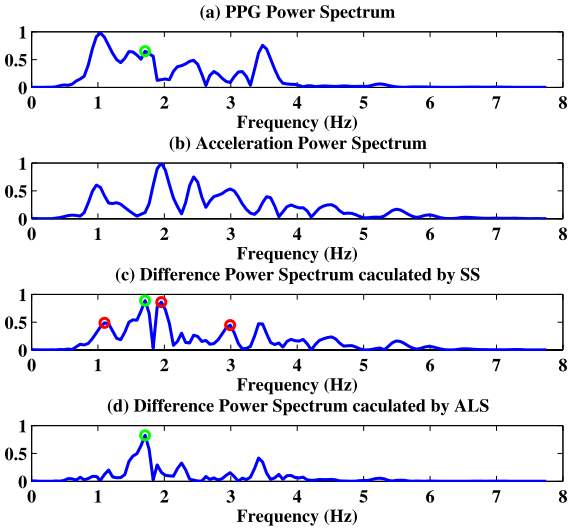

Fig. 3. Performance of SS on MA removal. (a) Normalized power spectrum of a raw PPG signal contaminated by strong MA. The true spectral peak corresponding to HR is indicated by a green circle. (b) Normalized power spectrum of a simultaneous acceleration signal. (c) The resulting difference spectrum after direct SS. The red circles indicates generated false peaks. (d) The difference spectrum calculated by ALS. Most false peaks are suppressed, and the true peak corresponding to HR can be easily detected.

components of MA in the PPG signal are correlated to the spectral components of the acceleration signal. Therefore, SS can be used to remove the MA. However, subtracting the acceleration spectrum from the PPG spectrum directly may generate many negative residuals and false spectral peaks, which can degrade performance of the following peak selection procedure. Figure 3 gives an example showing the drawbacks of the basic SS [20] in MA removal.

To solve the issue caused by conventional SS, we propose a new SS method based on asymmetric least squares (ALS) [26], [27]. ALS is a global optimization method independent of reference peaks, and can largely reduce the generation of false peaks in spectra. However, ALS generally suffers from bias in the presence of intense peaks (in relation to the noise level). To alleviate this problem, an adaptive iterative re-weighting is used to better reject false spectral peaks.

Denote the PPG spectrum by $\mathbf{y} = (y_1, y_2, \ldots, y_N)^T$ and the acceleration spectrum by $\mathbf{x} = (x_1, x_2, \ldots, x_N)^T$, where N is the spectrum length or the spectral grid number. The basic SS equation is expressed by:

$$\mathbf{z} = \mathbf{y} - \mathbf{a} \odot \mathbf{x},$$

where $\mathbf{z}$ is the difference spectrum, $\mathbf{a}$ is the scaling vector, i.e., each element of $\mathbf{x}$ is multiplied by a different scaling factor. $\odot$ denotes the operation of point-wise multiplication. The goal is to estimate the scaling vector to make negative peaks in $\mathbf{z}$ as weak as possible. To solve this problem, we propose the following ALS cost function

$$S = \sum_i \mathbf{w}_i \mathbf{z}_i^2 + \lambda \sum_i \left(\Delta^2 \mathbf{z}_i\right)^2 \tag{1}$$

where $\mathbf{z}_i = \mathbf{y}_i - \mathbf{a}_i \mathbf{x}_i$ and $\Delta^2 \mathbf{z}_i = \mathbf{z}_i - 2\mathbf{z}_{i-1} + \mathbf{z}_{i-2}$. $\mathbf{w}_i$ is the weight to $\mathbf{z}_i$. The first term in S accounts for the *fidelity* from $\mathbf{x}$ to $\mathbf{y}$, while the second term imposes *smoothness* to $\mathbf{z}$. The balance between fidelity and smoothness is controlled by the parameter λ, usually chosen between $10^2 \leq \lambda \leq 10^9$.

Note that in the cost function, the weights $\mathbf{w}_i$, if properly chosen, are able to increase the penalization produced by regions where the acceleration spectrum is above the PPG spectrum (i.e., false peak regions). We propose the following method to calculate $\mathbf{w}_i$,

$$\mathbf{w}_i = \begin{cases} p, & \text{if } \mathbf{z}_i > 0 \\ (1-p) \cdot \dfrac{2}{1 + e^{\left(\frac{|\mathbf{z}_i|}{k}\right)}}, & \text{otherwise} \end{cases} \tag{2}$$

where p is usually chosen as $10^{-3} \leq p \leq 10^{-1}$. The regions where the PPG spectrum is above the acceleration spectrum will contribute much less to the penalty, while the false peak regions are penalized much. Thus, minimization of S will effectively reduce the generation of false peaks in the difference spectrum $\mathbf{z}$. Note that in (2) $1-p$ is multiplied by a sigmoid-like function $\frac{2}{1+e^{\left(\frac{|\mathbf{z}_i|}{k}\right)}}$. Its goal is to overcome the drawback of using a fixed weight to the penalty of false peak regions, and thus, with the adaptive weighting, the estimation is more immune to the relative amplitudes of false (negative) peaks.[1] k is an additional parameter that controls the exponential decay of the weights. An easy way to estimate k is by setting it to the peak amplitude we want to start rejecting.

Rewriting (1) in a compact matrix form, we have:

$$S = \mathbf{z}^T \mathbf{W} \mathbf{z} + \lambda \|\mathbf{D}\mathbf{z}\|^2 \tag{3}$$

Here $\mathbf{D}$ is a differential matrix, i.e., $\mathbf{D}\mathbf{z} = \Delta^2 \mathbf{z}$. For example, when $\mathbf{z}$ is a 5 by 1 vector, the matrix $\mathbf{D}$ is given by:

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}$$

$\mathbf{W}$ is a diagonal matrix with the ith principal diagonal element being $\mathbf{w}_i$, i.e,

$$\mathbf{W} = \text{diag}\left([\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_N]\right).$$

Minimization of (3) can be solved by iteration. The gradient descent algorithm can be used starting from $\mathbf{w}_i = 1(\forall i)$ and simultaneously updating $\mathbf{W}$ and $\mathbf{a}$. Experiments showed that less than 10 iterations is enough for a proper estimation of $\mathbf{a}$.

Remark 1: Because in our experiments tri-axis acceleration data were used, the acceleration data along three axes were processed separately. Thus, we had three difference spectra, each corresponding to one axis.

C. Spectral Peak Tracking Method Based on Bayesian Decision Theory

After the spectrum subtraction stage, MA is partially removed in raw PPG signals. The next step is tracking the HR-related spectral peak.

We propose a peak tracking algorithm based on Bayesian decision theory. In this approach, both statistical properties of

[1]The relative amplitude of false spectral peaks could be of one or more orders.

the PPG signal and a priori knowledge about the characteristics of the peaks of interest are taken into account. The problem of peak tracking can be divided into two parts: peak detection and peak selection. Peak detection is used in finding possible peaks in the difference spectrum. In peak selection, only the peak that corresponds to the HR is chosen.

Peak detection is performed using amplitude thresholding with the threshold η automatically set to:

$$\eta = \delta \cdot \max\{|\mathbf{z}|\},$$

where $\mathbf{z}$ is the difference spectrum for each channel, δ is a parameter controlling the amplitude threshold, and $\max\{\cdot\}$ denotes the maximum value of the difference spectrum. The frequencies of spectral peaks with amplitudes larger than the threshold η are gathered as the candidate peak set.

In peak selection, the goal is to select the best possible spectral peak from the candidate peak set. This can be considered as a classification problem. Suppose there are m candidate peaks in the candidate peak set in the current time window, and the true spectral peak corresponding to the HR[2] is one of them. The goal is equivalent to classifying which peak is the true spectral peak.

Many features of the true spectral peak can be used in the classification. To find effective features, statistical properties of true peaks in available PPG data are investigated. Results showed that:

1) About 75% of spectral peaks that have good amplitude are true peaks. Peaks with good amplitude are defined as peaks with the highest amplitudes in their corresponding time windows.
2) About 84% of peaks that have good positions are true peaks. A peak with a good position refers to the one with the shortest distance from its previous true peak.
3) About 96% of true peaks have good amplitude and good position.

Based on these statistics, *peak amplitude* and *peak-to-peak separation* are used as features for classification. The peak-to-peak separation is defined as the frequency separation between the previous true peak and the current candidate peak. The two features are treated as two variables with specific distributions.

In particular, based on the investigation on extensive PPG data, we assume that the true peak's amplitude has the following distribution:

$$f(amp) = \frac{amp_{\max}}{1 + e^{-amp}}, \tag{4}$$

where amp is the peak amplitude and $amp_{\max}$ is the maximum peak amplitude of the spectrum in the current time window. Also, we assume that the peak-to-peak separation feature, denoted by sep, has the Laplacian distribution, which is defined as:

$$f_L(sep; \mu, b) = \frac{1}{2b} e^{-\frac{|sep-\mu|}{b}}, \tag{5}$$

where μ is the location parameter, and $b \geq 0$ is the scale parameter. Fig. 4 shows that the Laplacian distribution can well model the peak-to-peak separation of true peaks. Note that peak amplitude and peak-to-peak separation are assumed to be mutually independent.

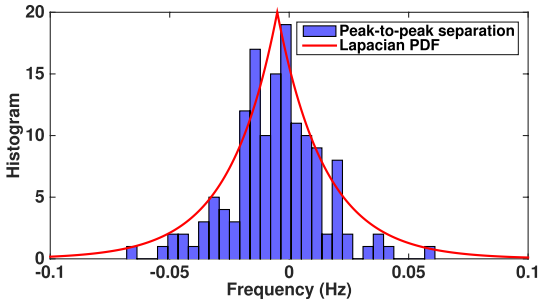

Fig. 4. Histogram of the peak-to-peak separation in a PPG dataset. The Laplacian PDF with parameters $\mu = 0$ and $\sigma = 0.02$ is depicted by a solid line.

Suppose that m candidate peaks have been detected in the current spectrum, without any other information about these peaks, every peak is equally likely to be the true peak, i.e., the prior probability of the i^{th} candidate peak to be TRUE will be:

$$P(\omega_i = \mathcal{T}) = \frac{1}{m}.$$

where $\omega_i = \mathcal{T}$ denotes the event that the ith peak is the true spectral peak corresponding to HR.

Suppose the ith peak has the amplitude amp_i and peak-to-peak separation sep_i. The feature vector can be constructed as $\mathbf{v} = [amp_i, sep_i]^T$, where $\mathbf{v}$ is in a 2-dimensional Euclidean space $\mathbf{R}^2$, called the feature space. Let $p(\mathbf{v}|\omega_i)$[3] be the state-conditional probability density function for $\mathbf{v}$, i.e., the probability density function for $\mathbf{v}$ conditioned on ω_i. The posterior probability $P(\omega_i = \mathcal{T}|\mathbf{v})$ can then be computed from $p(\mathbf{v}|\omega_i = \mathcal{T})$ from Bayes' rule,

$$\begin{aligned} P(\omega_i = \mathcal{T}|\mathbf{v}) &= \frac{p(\mathbf{v}|\omega_i = \mathcal{T})P(\omega_i = \mathcal{T})}{p(\mathbf{v})} \\ &\sim p(\mathbf{v}|\omega_i = \mathcal{T})P(\omega_i = \mathcal{T}), \end{aligned} \tag{6}$$

where we omit the evidence factor $p(\mathbf{v})$.

Because the peak amplitude and peak-to-peak separation are assumed to be independent, $p(\mathbf{v}|\omega_i = \mathcal{T})$ can then be calculated as:

$$p(\mathbf{v}|\omega_i = \mathcal{T}) = p(amp_i|\omega_i = \mathcal{T})p(sep_i|\omega_i = \mathcal{T}).$$

Therefore, the posterior probability (6) can be calculated as:

$$P(\omega_i = \mathcal{T}|\mathbf{v}) \sim p(amp_i|\omega_i = \mathcal{T})p(sep_i|\omega_i = \mathcal{T})P(\omega_i = \mathcal{T}),$$

where $p(amp_i|\omega_i = \mathcal{T}) = f(amp_i)$ is defined in (4) and $p(sep_i|\omega_i = \mathcal{T}) = f_L(sep_i; \mu, b)$ is defined in (5).

After calculating the posterior probability, the decision procedure is developed based on the *minimum-error-rate* classification method [28]. Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the finite set of m possible actions, and the action α_j corresponds to deciding that the jth peak is the true peak. Define the loss function $L(\alpha_j|\omega_i = \mathcal{T})$ which describes the loss incurred for taking

[2]For convenience, in the following context such a spectral peak is briefly called the true peak.

[3]We generally use an upper-case $P(\cdot)$ to denote a probability mass function and a lower-case $p(\cdot)$ to denote a probability density function.

action α_j when $\omega_i = \mathcal{T}$. The expected loss associated with taking action α_j is given by:

$$R(\alpha_j|\mathbf{v}) = \sum_{i=1}^{m} L(\alpha_j|\omega_i = \mathcal{T}) P(\omega_i = \mathcal{T}|\mathbf{v}) \tag{7}$$

There are many ways to define the loss function $L(\alpha_j|\omega_i = \mathcal{T})$. For simplicity, we define the loss function as the zero-one loss function

$$L(\alpha_j|\omega_i = \mathcal{T}) = \begin{cases} 0, & j = i, \\ 1, & j \neq i. \end{cases}$$

With the loss function, (7) can be further expressed as

$$\begin{aligned} R(\alpha_j|\mathbf{v}) &= \sum_{i=1}^{m} L(\alpha_j|\omega_i = \mathcal{T}) P(\omega_i = \mathcal{T}|\mathbf{v}) \\ &= \sum_{i \neq j} P(\omega_i = \mathcal{T}|\mathbf{v}) \\ &= 1 - P(\omega_j = \mathcal{T}|\mathbf{v}). \end{aligned}$$

The result states that to minimize the expected loss when taking action α_j, one should select the jth peak whose posterior $P(\omega_j = \mathcal{T}|\mathbf{v})$ is the largest. Thus, we have the following decision rule:

$$\begin{aligned} &\text{DECIDE } \omega_j = \mathcal{T}, \\ &\text{IF } P(\omega_j = \mathcal{T}|\mathbf{v}) > P(\omega_i = \mathcal{T}|\mathbf{v}), \quad \forall i \neq j. \end{aligned}$$

D. Post-Processing

After the peak tracking stage, an incorrect HR estimate may be obtained due to MA and other inference. For example, in the presence of extremely strong MA, a raw PPG signal may not contain any heartbeat-related component [22], and thus the peak tracking algorithm may find a wrong spectral peak. Also, in the presence of rhythmic hand movements such as during running and walking, MA's frequencies may overlap with heartbeat frequency. In this situation, the true peak corresponding to HR may be removed by SS, and the peak tracking algorithm may select a spectral peak far from the true peak. Thus, a post-processing stage is necessary to correct the wrong estimates.

To achieve this, similar as in [21], a moving average smoothing filter is developed as follows. First, a smoothing window of 20 seconds (containing 10 HR estimates) is constructed. For the ith HR estimate f_i, the average HR frequency $\overline{f_i}$ and the standard deviation σ_i are computed as:

$$\begin{aligned} \overline{f_i} &= \frac{1}{10} \sum_{j=i-5}^{i+4} f_j, \\ \sigma_i &= \sqrt{\frac{1}{10} \sum_{j=i-5}^{i+4} (f_j - \overline{f_i})^2}. \end{aligned}$$

Once the difference between f_i and $\overline{f_i}$ is greater than σ_i, this HR estimate f_i will be determined to be a wrong estimate, and its value is set to $\overline{f_i}$, i.e.,

$$f_i = \begin{cases} f_i, & |f_i - \overline{f_i}| < \sigma_i, \\ \overline{f_i}, & \text{otherwise.} \end{cases}$$

Although the post-processing procedure may not be used in real-time HR monitoring, it can be used for long-term HR monitoring for heart disease diagnosis and other scenarios in which real-time monitoring is not required. Besides, the post-processing method in [21] can be used if complete real-time monitoring is needed.

III. Experimental Results

A. Data Description

The PPG database was used for the 2015 IEEE Signal Processing Cup [29]. It includes 12 datasets for training, and 10 datasets for testing. Each dataset consists of two channels of PPG signals, three channels of simultaneous acceleration signals, and one channel of simultaneous ECG signal. The PPG signals were recorded from subjects' wrist (dorsal locations) using PPG sensors built in a wristband. The PPG sensors used green LEDs working at 515 nm. The acceleration signals were recorded using a tri-axis accelerometer also built in the wristband. The ECG signals were recorded using wet ECG sensors locating at the chest of subjects. The ground-truth of heart rate was calculated from the ECG signal, which was used to evaluate algorithms' performance. All signals were sampled at 125 Hz and wireless transmitted via blue-tooth to a nearby computer.

During the recording of the 12 training datasets, 12 male subjects with yellow skin and ages ranging from 18 to 35 participated. The participants walked or ran with the following speeds in order: the speed of 1-2 km/hour for 0.5 minute, the speed of 6-8 km/hour for 1 minute, the speed of 12-15 km/hour for 1 minute, the speed of 6-8 km/hour for 1 minutes, the speed of 12-15 km/hour for 1 minute, and the speed of 1-2 km/hour for 0.5 minute. During the recording, the participants intentionally performed other actions such as wiping sweat on forehead, pushing buttons on the treadmill, and freely swinging hands. During the recording of the 10 testing datasets, 8 subjects (7 male and 1 female) with yellow skin and ages ranging from 19 to 58 participated. The subject performed many intensive activities including various forearm and upper arm exercise (e.g. shake hands, stretch, push, and so on, which are common in arm rehabilitation exercise), running, jump, and push-up. Thus the testing datasets include much stronger MA than the training datasets.[4] Although there are two PPG signals in each dataset, only the first PPG signal was used in the proposed algorithm.

The training datasets were used to train the algorithm and select appropriate parameter values. The relevant parameters in the experiments were adjusted as follows: the spectrum grid parameter in DFT was set to $N = 2048$, the smoothing parameter was set to $\lambda = 10^2$, the weighting parameter in (3) to $p = 10^{-3}$, and the parameter in (6) to $\delta = 0.7$. In the peak selection part, the parameters of the Laplacian distribution function were set to $\mu = 0$ and $b = 0.02$. After that all training datasets and test datasets are used to evaluate the performance of the SPECTRAP.

[4]Details on the database can be found at: http://www.signalprocessingsociety.org/spcup2015/.

TABLE I
COMPARISION OF SPECTRAP AND TROIKA IN TERMS OF AVERAGE ABSOLUTE ERRORS (ERROR1) ON THE 12 TRAINING DATASETS. THE UNIT IS BPM, SD INDICATES THE STANDARD DEVIATION

	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10	No.11	No.12	**Average**
SPECTRAP	1.18	2.42	0.86	1.38	0.92	1.37	1.53	0.64	0.60	3.65	0.92	1.25	**1.50 (SD=1.95)**
TROIKA	2.87	2.75	1.91	2.25	1.69	3.16	1.72	1.83	1.58	4.00	1.96	3.33	**2.42 (SD=2.47)**

TABLE II
COMPARISON OF SPECTRAP AND TROIKA IN TERMS OF AVERAGE ABSOLUTE ERROR PERCENTAGE (ERROR2) ON THE 12 TRAINING DATASETS. SD INDICATES THE STANDARD DEVIATION

	No.1	No.2	No.3	No.4	No.5	No.6	No.7	No.8	No.9	No.10	No.11	No.12	**Average**
SPECTRAP	1.04%	2.33%	0.66%	1.31%	0.74%	1.14%	1.36%	0.55%	0.52%	2.27%	0.65%	1.02%	**1.12% (SD=1.47%)**
TROIKA	2.18%	2.37%	1.50%	2.00%	1.22%	2.51%	1.27%	1.47%	1.28%	2.49%	1.29%	2.30%	**1.82% (SD=2.07%)**

TABLE III
AVERAGE ABSOLUTE ERRORS (ERROR1) AND AVERAGE ABSOLUTE ERROR PERCENTAGE (ERROR2) OF SPECTRAP ON THE 10 TEST DATASETS. SD INDICATES THE STANDARD DEVIATION

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10	**Average**
Error1	4.89	1.58	1.83	3.05	1.62	1.24	2.04	2.49	1.16	0.66	**2.13 (SD=2.77)**
Error2	6.29%	1.98%	1.49%	2.0%	1.36%	0.92%	2.23%	1.81%	0.92%	0.79%	**2.04% (SD=3.01%)**

B. Data Analysis and Statistics

Post-processing and analysis of the recordings were performed using custom software written in MATLAB (The MathWorks, Inc., Natick, MA, USA). Ground-truth of HR was calculated from simultaneously recorded ECG signal as follows. In each time window, the number of cardiac cycles H and the duration D (in seconds) were counted, and the HR was then calculated as $60H/D$ (in BPM).[5]

To evaluate the proposed algorithm, multiple measurement indexes were used. The first measurement index was the average absolute error (in BPM), defined as

$$\text{Error1} = \frac{1}{\text{L}} \sum_{\text{i}=1}^{\text{L}} |\text{BPM}_{\text{est}}(\text{i}) - \text{BPM}_{\text{ECG}}(\text{i})|,$$

where L is the total number of time windows, $\text{BPM}_{\text{ECG}}(i)$ is the ground-truth of HR in the ith time window calculated from the simultaneously recorded ECG signal, and $\text{BPM}_{\text{est}}(i)$ is the estimated HR.

The second was the average absolute error percentage, defined as

$$\text{Error2} = \frac{1}{\text{L}} \sum_{\text{i}=1}^{\text{L}} \frac{|\text{BPM}_{\text{est}}(\text{i}) - \text{BPM}_{\text{ECG}}(\text{i})|}{\text{BPM}_{\text{ECG}}(\text{i})}.$$

As the third measurement index, Bland–Altman plots [30] were used for combined graphical and statistical interpretation of the two measurement techniques. The differences between heart rate measurements from the wrist and ECG were expressed as percentages of the averages in both techniques and plotted against the averages. The 95% Limit of Agreement (LOA) was calculated in this analysis, and defined as the average absolute error (μ_E) ± 1.96 standard deviation (σ_E) of the absolute error, i.e., $([\mu_E - 1.96\sigma_E, \mu_E + 1.96\sigma_E])$, reflecting the range where 95% of all errors between measurements were expected to lie.

[5]Note that the ground-truth of HR is available in the PPG database, since we used the same sliding window as required in the 2015 IEEE Signal Processing Cup.

C. Results

1) Performance of SPECTRAP: To evaluate SPECTRAP, it was compared with the recently proposed TROIKA algorithm [22]. The average absolute error (Error1) and the average absolute error percentage (Error2) on the 12 training datasets are given in Table I and Table II, respectively. It can be seen that SPECTRAP had better performance than TROIKA. In addition, averaged across the 12 training datasets, the absolute estimation error (Error1) of SPECTRAP was 1.50 ± 1.95 BPM (mean $\pm$ standard deviation), and the error percentage (Error2) was $1.12\% \pm 1.47\%$. In contrast, TROIKA had the performance of Error1 $= 2.42 \pm 2.47$ BPM and Error2 $= 1.82\% \pm 2.07\%$. (These results were directly adopted from [22].)

SPECTRAP was further evaluated on the 10 testing datasets. The average absolute error (Error1) and the average absolute error percentage (Error2) are given in Table III. Averaged across the 10 testing datasets, the Error1 of SPECTRAP was 2.13 ± 2.77 BPM, and the Error2 was $2.04\% \pm 3.01\%$.

To deeply look into the performance of SPECTRAP, Figure 5 gives the estimated HR trace on the training dataset 9 (randomly chosen). The HR trace estimated by SPECTRAP without post-processing is also plotted. The results show that the HR trace estimated by complete SPECTRAP was almost the same as the ground-truth of HR trace. If the

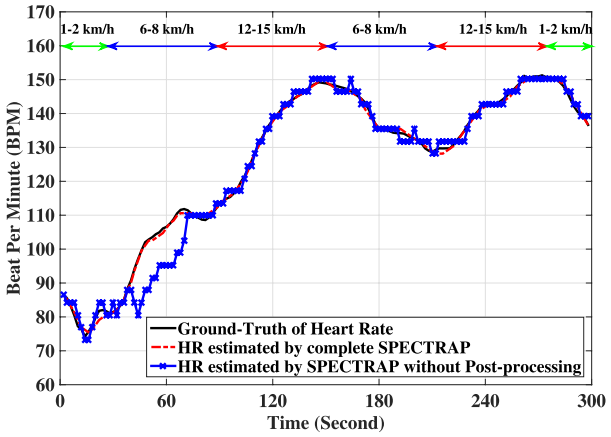

Fig. 5. Estimation results on training dataset 9. The HR trace estimated by complete SPECTRAP was almost the same as the ground-truth of HR trace. Missing the post-processing step resulted in degraded performance, but still with acceptable HR monitoring ability. The walking/running speed at every duration is indicated on the top of the figure.

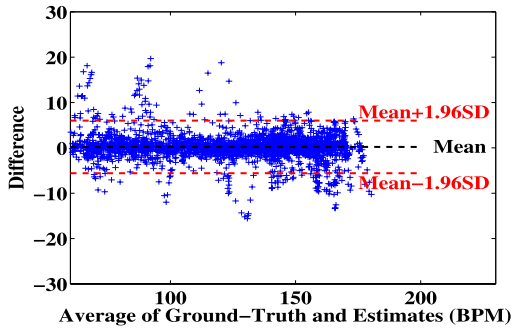

Fig. 6. The Bland–Altman plot of the estimation results over all 22 datasets. The LOA is [−5.59, 6.01] BPM.

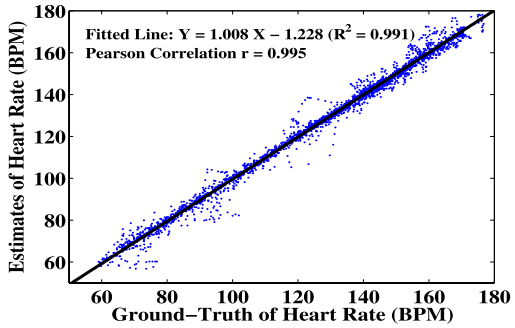

Fig. 7. Scatter plot between the ground-truth heart rate values and the associated estimates over all 22 datasets. The fitted line was $Y = 1.008X - 1.228$, where X indicates the ground-truth heart rate value, and Y indicates the associated estimate. The R^2 value, a measure of goodness of fit, was 0.991. The Pearson correlation was 0.995.

post-processing step was omitted, the performance was degraded, but still acceptable.

The Bland–Altman plot over all 22 datasets is depicted in Figure 6, where the LOA is [−5.59, 6.01] BPM. The Scatter Plot between the ground-truth HR values and the associated estimates is given in Figure 7, which shows the fitted line is $Y = 1.008X - 1.228$, where X indicates the ground-truth heart rate value, and Y indicates the associated estimate.

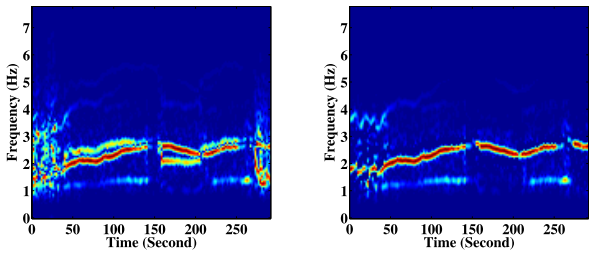

Fig. 8. Difference spectrum calculated using traditional SS and ALS-SS on training dataset 3. (a) Normal SS. (b) ALS-SS.

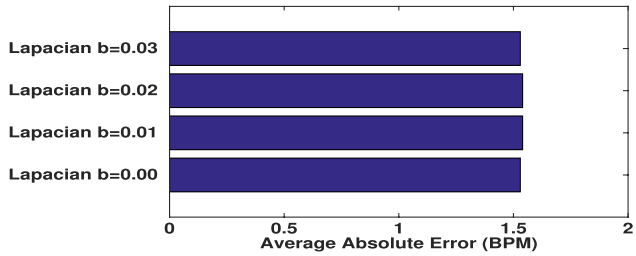

Fig. 9. Robustness of SPECTRAP to different parameters evaluated using the Test Dataset 2. The parameters of the Laplacian distribution function were set to $\mu = 0$, and b changed from 0 to 0.03.

The Pearson coefficient was 0.995. The goodness of fit characterized by R^2 value was 0.991.

2) Effect of Motion Artifact Removal: To see the effect of MA removal of the proposed ALS-based SS algorithm (denoted by ALS-SS), the difference spectra calculated by traditional SS [20] and ALS-SS on training dataset 3 (randomly chosen) are shown in Figure 8(a) and (b), respectively. The difference spectra were computed using Periodogram with a window size of 2048. The false peaks and high-power noise components which led to inaccurate and unreliable peak tracking can be clearly seen in the difference spectrum in Figure 8(a), indicating that the HR trace was not extracted precisely. For the ALS-SS algorithm, the effect of removing false peaks and high-power noise components can be seen in Figure 8(b), where most false peaks were suppressed and some true peaks were enhanced and more easy to detect. A very clear HR trace can be seen in this figure.

3) Robustness to Peak-to-Peak Separation Distribution: To evaluate robustness to peak-to-peak separation distribution, the sensitivity of SPECTRAP to different distribution parameters is tested using the test dataset 2 (randomly chosen). In previous experiments, the parameters of the Laplacian distribution function were set to $\mu = 0$ and $b = 0.02$. In this experiment, SPECTRAP was performed using the Laplacian distribution function with parameter b changed from 0 to 0.03. The results are given in Figure 9, showing the algorithm is robust to various parameter values.

4) Computational Complexity of SPECTRAP: To evaluate computational complexity of SPECTRAP, Table IV gives running time of SPECTRAP and TROIKA on the training dataset 7 (randomly chosen). The experiment was carried

TABLE IV
RUNNING TIME (ON DATA IN EACH TIME WINDOW OF 8 SECONDS) OF SPECTRAP AND TROIKA ON TRAINING DATASET 7

	Average (s)	Standard Deviation (s)
SPECTRAP	0.0162	0.0015
TROIKA	0.9419	0.1187

out on a computer with Intel Core 2 Duo P7550 @2.26GHz and 4 GB RAM. Results show that the computational cost of SPECTRAP was significantly lower than that of TROIKA.

IV. CONCLUSIONS

In this study, a method named SPECTRAP for heart rate estimation using a wrist-type PPG during physical exercise is presented. The method contains three parts: asymmetric least squares based spectrum subtraction, Bayesian decision theory based spectral peak selection, and a post-processing procedure. Experimental results proved the efficacy of SPECTRAP for reliable and accurate estimation of heart rate.

REFERENCES

[1] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, no. 3, pp. R1–R39, 2007.

[2] C. M. Lee and Y. T. Zhang, "Reduction of motion artifacts from photoplethysmographic recordings using a wavelet denoising approach," in *Proc. IEEE EMBS Asian-Pacific Conf. Biomed. Eng.*, Oct. 2003, pp. 194–195.

[3] J. Y. A. Foo, "Comparison of wavelet transformation and adaptive filtering in restoring artefact-induced time-related measurement," *Biomed. Signal Process. Control*, vol. 1, no. 1, pp. 93–98, 2006.

[4] M. Raghuram, K. V. Madhav, E. H. Krishna, and K. A. Reddy, "Evaluation of wavelets for reduction of motion artifacts in photoplethysmographic signals," in *Proc. 10th Int. Conf. Inf. Sci. Signal Process. Appl. (ISSPA)*, May 2010, pp. 460–463.

[5] M. J. Hayes and P. R. Smith, "Artifact reduction in photoplethysmography," *Appl. Opt.*, vol. 37, no. 31, pp. 7437–7446, 1998.

[6] Y.-S. Yan, C. C. Poon, and Y.-T. Zhang, "Reduction of motion artifact in pulse oximetry by smoothed pseudo wigner-ville distribution," *J. Neuroeng. Rehabil.*, vol. 2, no. 3, pp. 1–9, 2005.

[7] S. Seyedtabaii and L. Seyedtabaii, "Kalman filter based adaptive reduction of motion artifact from photoplethysmographic signal," *World Acad. Sci., Eng. Technol.*, vol. 37, no. 2, pp. 173–176, 2008.

[8] X. Sun, P. Yang, Y. Li, Z. Gao, and Y.-T. Zhang, "Robust heart beat detection from photoplethysmography interlaced with motion artifacts based on empirical mode decomposition," in *Proc. IEEE-EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Jan. 2012, pp. 775–778.

[9] Y. Zhang, B. Liu, and Z. Zhang, "Combining ensemble empirical mode decomposition with spectrum subtraction technique for heart rate monitoring using wrist-type photoplethysmography," *Biomed. Signal Process. Control*, vol. 21, pp. 119–125, Aug. 2015.

[10] L. B. Wood and H. H. Asada, "Low variance adaptive filter for cancelling motion artifact in wearable photoplethysmogram sensor signals," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, Aug. 2007, pp. 652–655.

[11] S. H. Kim, D. W. Ryoo, and C. Bae, "Adaptive noise cancellation using accelerometers for the PPG signal from forehead," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBS)*, Aug. 2007, pp. 2564–2567.

[12] H. H. Asada, H.-H. Jiang, and P. Gibbs, "Active noise cancellation using MEMS accelerometers for motion-tolerant wearable bio-sensors," in *Proc. 26th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (IEMBS)*, vol. 1. Sep. 2004, pp. 2157–2160.

[13] H. Han and J. Kim, "Artifacts in wearable photoplethysmographs during daily life motions and their reduction with least mean square based active noise cancellation method," *Comput. Biol. Med.*, vol. 42, no. 4, pp. 387–393, 2012.

[14] B. Lee, J. Han, H. J. Baek, J. H. Shin, K. S. Park, and W. J. Yi, "Improved elimination of motion artifacts from a photoplethysmographic signal using a Kalman smoother with simultaneous accelerometry," *Physiol. Meas.*, vol. 31, no. 12, p. 1585, 2010.

[15] M.-Z. Poh, N. C. Swenson, and R. W. Picard, "Motion-tolerant magnetic earring sensor and wireless earpiece for wearable photoplethysmography," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 786–794, May 2010.

[16] H. H. Asada, P. Shaltis, A. Reisner, S. Rhee, and R. C. Hutchinson, "Mobile monitoring with wearable photoplethysmographic biosensors," *IEEE Eng. Med. Biol. Mag.*, vol. 22, no. 3, pp. 28–40, May/Jun. 2003.

[17] M. R. Ram, K. V. Madhav, E. H. Krishna, N. R. Komalla, and K. A. Reddy, "A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 5, pp. 1445–1457, May 2012.

[18] J. M. Goldman, M. T. Petterson, R. J. Kopotic, and S. J. Barker, "Masimo signal extraction pulse oximetry," *J. Clin. Monitor. Comput.*, vol. 16, no. 7, pp. 475–483, 2000.

[19] R. Yousefi, M. Nourani, S. Ostadabbas, and I. Panahi, "A motion-tolerant adaptive algorithm for wearable photoplethysmographic biosensors," *IEEE J. Biomed. Health Inform.*, vol. 18, no. 2, pp. 670–681, Mar. 2014.

[20] P. C. Loizou, *Speech Enhancement: Theory and Practice*. Boca Raton, FL, USA: CRC Press, 2013.

[21] H. Fukushima, H. Kawanaka, M. S. Bhuiyan, and K. Oguri, "Estimating heart rate using wrist-type photoplethysmography and acceleration sensor while running," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug./Sep. 2012, pp. 2901–2904.

[22] Z. Zhang, Z. Pi, and B. Liu, "TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 522–531, Feb. 2015.

[23] Z. Zhang, "Heart rate monitoring from wrist-type photoplethysmographic (PPG) signals during intensive physical exercise," in *Proc. 2nd IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Dec. 2014, pp. 698–702.

[24] Z. Zhang, "Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 1902–1910, Aug. 2015.

[25] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ, USA: Prentice-Hall, 2005.

[26] W. K. Newey and J. L. Powell, "Asymmetric least squares estimation and testing," *Econometrica, J. Econ. Soc.*, vol. 55, no. 4, pp. 819–847, 1987.

[27] P. H. C. Eilers, "A perfect smoother," *Anal. Chem.*, vol. 75, no. 14, pp. 3631–3636, 2003.

[28] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York, NY, USA: Wiley, 2012.

[29] Z. Zhang, "Undergraduate students compete in the IEEE signal Processing cup: Part 3," *IEEE Signal Process. Mag.*, vol. 32, no. 6, 2015.

[30] J. M. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, vol. 327, no. 8476, pp. 307–310, 1986.

Biao Sun, photograph and biography not available at the time of publication.

Zhilin Zhang, photograph and biography not available at the time of publication.