

# A Robust Heart Rate Monitoring Scheme Using Photoplethysmographic Signals Corrupted by Intense Motion Artifacts

Emroz Khan, Forsad Al Hossain, Shiekh Zia Uddin, S. Kaisar Alam,  
and Md. Kamrul Hasan\*

**Abstract— Goal:** Although photoplethysmographic (PPG) signals can monitor heart rate (HR) quite conveniently in hospital environments, trying to incorporate them during fitness programs poses a great challenge, since in these cases, the signals are heavily corrupted by motion artifacts. **Methods:** In this paper, we present a novel signal processing framework which utilizes two channel PPG signals and estimates HR in two stages. The first stage eliminates any chances of a runaway error by resorting to an absolute criterion condition based on ensemble empirical mode decomposition. This stage enables the algorithm to depend very little on the previously estimated HR values and to discard the need of an initial resting phase. The second stage, on the other hand, increases the algorithm’s robustness against offtrack errors by using recursive least squares filters complemented with an additional novel technique, namely time-domain extraction. **Results:** Using this framework, an average absolute error of 1.02 beat per minute (BPM) and standard deviation of 1.79 BPM are recorded for 12 subjects performing a run with peak velocities reaching as high as 15 km/h. **Conclusion:** The performance of this algorithm is found to be better than the other recently reported algorithms in this field such as TROIKA and JOSS. **Significance:** This method is expected to greatly facilitate the presently available wearable gadgets in HR computation during various physical activities.

**Index Terms**—Ensemble empirical mode decomposition (EEMD), photoplethysmography (PPG), recursive least squares (RLS) filter, wearable biomedical computing.

## I. INTRODUCTION

**P**HOTOPLETHYSMOGRAPHIC (PPG) signals [1], [2] are becoming a popular means of monitoring heart rate (HR) because of their wearable implementation compared to the conventional electrocardiography (ECG) technology. These signals are obtained through pulse oximeters which are embedded in a small wearable device to be put on at earlobes or fingertips (transmission type) or at wrists (reflection type). The pulse oximeters cast light on the wearer’s skin through light-emitting diodes (LED) and obtain the transmitted or reflected light whose intensity depends on the amount of blood present in the arteries

Manuscript received March 4, 2015; revised July 8, 2015; accepted August 1, 2015. Date of publication; date of current version. Asterisk indicates corresponding author.

E. Khan, F. A. Hossain, and S. Z. Uddin are with the Bangladesh University of Engineering and Technology.

S. K. Alam is with Improlabs Pte, Ltd.

\*M. K. Hasan is with the Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka 1000, Bangladesh (e-mail: khasan@eee.buet.ac.bd).

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/TBME.2015.2466075

under the skin. Since the amount of blood varies with the cardiac cycle, the light intensity also varies with the cardiac rhythm and, thus, can be used to extract HR information.

Initially after its emergence, PPG signals were used to monitor patients’ HRs 24 h in hospital beds, mainly using the fingertip-type devices. Even in these situations, slight movements of the patients’ fingers would cause significant distortion in the signals and made the HR estimation difficult. Several attempts have been made to address this problem [3]–[8]. Ram *et al.* [6] used adaptive noise cancellation technique to remove the motion artifact (MA), though it suffered from high sensitivity to the choice of the reference signal. Sun *et al.* [7] performed artifact reduction using empirical mode decomposition (EMD), but considered all the modes with soft thresholding, instead of extracting only the correct PPG modes. Peng *et al.* [8] considered constrained-independent component analysis together with least mean square (LMS) adaptive filters to recover HR information. Though this algorithm and the rest others perform well for limited motion tolerance (e.g., horizontal and vertical movements of the finger, finger bending etc., with the subject sitting still), trying to apply them for significant MA (e.g., when the subject is in motion) does not yield satisfactory results.

Apart from clinical scenarios, many exercisers, sportsman, and elderly people require real-time mobile surveillance on their HR, so as to restrain them from excessive physical exertion and thereby, to prevent any chances of heart seizures. Following this line, many smart wearable gadgets have already been merchandised; however, these still require further robustness in their performance. Motivated by this future aspect of PPG signals, the IEEE Signal Processing Society has also arranged a research contest to invite attention to this problem [9].

To aid the signal processing task when PPG signals are being used during physical exercise, often three-axis accelerometry data are recorded in an attempt to model the MA [10]. Yousefi *et al.* [11] first attempted to bring the tissue effects under consideration and suggested a two-stage normalized-LMS adaptive noise canceller to remove MA. Though the method gives appreciable improvements over other previously reported algorithms, it requires pulse oximeters with different wavelengths. López *et al.* [12] on the other hand first pointed out the importance of heuristic approach to this problem and made interesting observations in the trend of PPG signals under intense running exercise. Using these observations, an algorithm was devised which yielded promising results emphasizing that previously estimated HR values can play a significant role in HR

estimation. But the algorithm does not work well in other types of exercises (e.g., boxing); in fact, even during rest, the algorithm faces difficulty in estimating HR.

Very recently, Zhang *et al.* [13] proposed a three-stage general framework and obtained benchmark results for substantial MA. But it still suffers from robustness to initial estimates. As a result, the subjects are required to stay still initially so that the initial estimates are correct and the tracker can perform accurately. As a sequel to this study, Zhang proposed another algorithm in [14] which though manages to further improve the previous results by a more vigorous signal sparsification technique, that is still heavily dependent on the tracking mechanism. In fact in his new algorithm, the HR monitoring process is not initialized at all if the MA is significant at the beginning. This was to make sure that HR estimates are on the right track from the start since the algorithm is not immune to initial wrong estimates.

In this paper, we consider the aforementioned shortcomings of the previous methods and propose a novel technique which works essentially in two stages. In the first stage, it is ensured that the algorithm is not suffering from a *runaway* error; that is, if for some reason the algorithm loses track of the true HR, it can still manage to get back on track no matter how far the ground truth is from the previous estimates. To accomplish this, we applied a noise-assisted ensemble empirical mode decomposition (EEMD) as a signal denoising technique together with an *absolute criterion* (AC) which dictates the algorithm to overthrow the dependence on previous estimates. Consequently, this stage also makes no need of an initial resting boot phase. The second stage works when the estimates are assumed to be already on the right track and deals with close proximity of MA and PPG peaks and spectral shadowing, so that the algorithm does not choose wrong estimates and get *offtrack*. To do so, we use a recursive least squares (RLS) adaptive filter complemented with an additional novel technique, namely *time-domain extraction*. Using this two-stage method, HR computation is performed on both single-channel and double-channel PPG signals on a number of sample datasets, and its performance is compared with the previously reported algorithms in the literature such as [13], [14].

This paper is organized as follows: Section II discusses the various cases where HR estimation is quite difficult, and Section III suggests ways to come around them. Section IV puts all the ideas together and establishes a well-structured algorithm. Section V illustrates the experimental results and compares with other benchmark algorithms in this field. Finally, Section VI concludes this paper.

## II. PROBLEM INSIGHTS

HR monitoring from motion-corrupted PPG signals is a challenging problem for a number of reasons. Some of them are listed below:

- 1) During intense hand movement (running, for example), the space between the wrist and the oximeter changes, and as a result, the measured intensity of the PPG signal varies. This variation is often correlated with the frequency of the hand motions. To illustrate this, Fig. 1(a) shows a segment

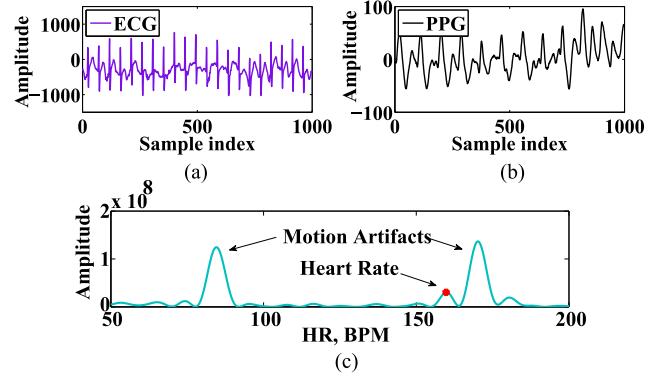


Fig. 1. Segment of the (a) ECG signal and the (b) corresponding PPG signal in time domain. (c) Periodogram of the PPG signal shows that there are large spurious MA peaks near the HR peak.

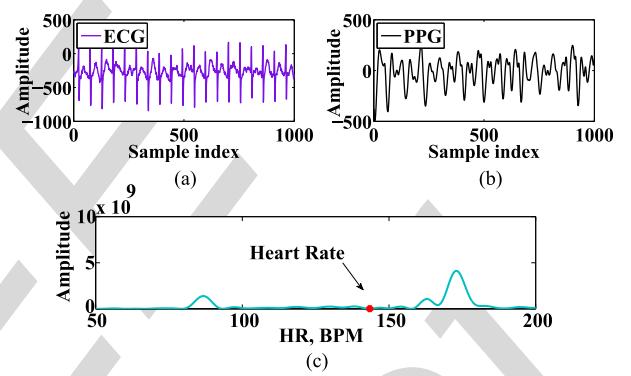


Fig. 2. Segment of the (a) ECG signal and the (b) corresponding PPG signal in time domain. (c) Periodogram of the PPG signal shows that there is no peak corresponding to the HR.

of the ECG signal, whereas (b) shows the corresponding PPG signal in time domain. As given by the periodogram of the PPG signal [see Fig. 1(c)], we can see that there are large spurious MA peaks near the HR peak.

- 2) Sometimes the sensors might be so far apart from the skin due to exhaustive motion that the original peak corresponding to the HR might be totally absent in the spectrum. Fig. 2(a) and (b) shows such ECG and PPG segments, respectively, together with the periodogram of the PPG in (c). This situation might persist for several seconds and poses a great difficulty in HR estimation.
- 3) Sometimes the peak corresponding to MA might be so close to the HR peak that they may get indistinguishable for a given resolution. Fig. 3(a) and (b) gives an example of such ECG–PPG pair along with the periodogram of PPG in (c). Increasing the data length will not circumvent the problem because that would require the HR process to be stationary for an extended amount of period, which seldom is the case. To make things worse, often the side-lobe of a MA peak is so strong that it masks the original HR peak.
- 4) As pointed out in [11], intense motion like boxing, jumping, etc., may cause so much acceleration in the blood at the arteries that even correct intensity might not correspond to actual HR.

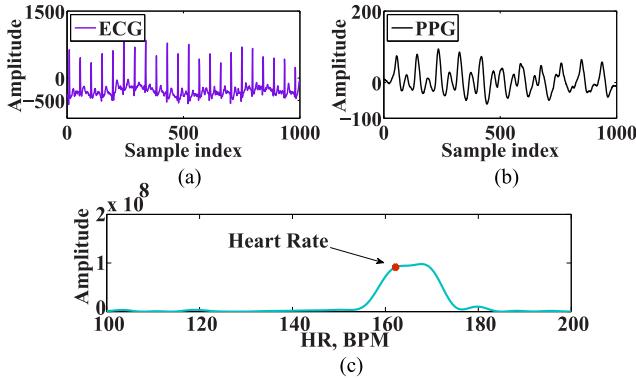


Fig. 3. Segment of the (a) ECG signal and the (b) corresponding PPG signal in time domain. (c) Periodogram of the PPG signal shows that the MA peak and the HR peak are almost indistinguishable from each other.

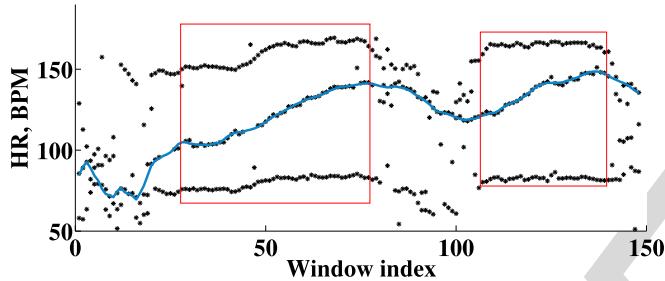


Fig. 4. Plot (in blue line) of the ground truth HRs as a subject starts running with an initial rest. After a while, the subject pauses to take a brief rest and starts running again at a different speed. The overlay dots (in black) are the strongest PPG peaks found in each time window. Clearly, the trend mentioned in [12] is observed, while the subject runs steadily (inside the red rectangles), but the pattern is broken elsewhere.

- 5) HR is a periodic process (if the time window is not too long) and, therefore, can have higher order harmonics. But the presence of harmonics is not sufficient enough to differentiate between HR and MA peaks (as was done in [13]), since MA can also be aperiodic or periodic. In rest, the random movements of the hands cause the MA to be aperiodic, whereas the free swinging of the hands in a running exercise renders the MA periodic.
- 6) The observation made in [12] shows that when a subject performs a run, then for each time window the HR is usually among the top three dominant peaks with the other two coming from MA (similar to the example given in Fig. 1). In fact, if we plot the overlays of the top three dominant peak locations over all the time windows (as shown in Fig. 4), then a certain pattern is observed. In this pattern, the highest and the lowest frequency peak locations remain almost constant and correspond to MA, whereas the middle frequency peak location shows a rising or falling trend corresponding to the actual HR. Though this holds to a good approximation when the subject runs steadily (inside the red rectangles in Fig. 4), the pattern is destroyed at rest or at change in running speed (similar to the cases illustrated in Figs. 2 and 3). In fact, estimating the HR during running is easier sometimes than the rest

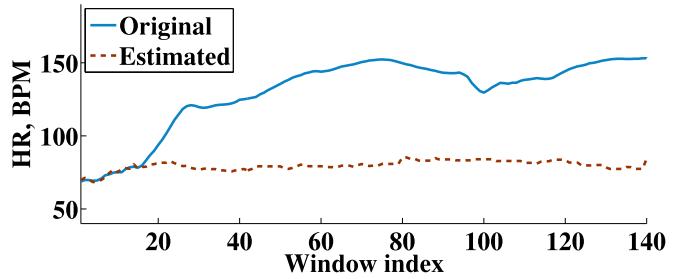


Fig. 5. Example of a runaway situation which arises from too much dependence upon the tracking mechanism. When a MA peak is close to the HR peak and is mistaken to be a HR estimate, the tracking mechanism would keep following the MA peaks in successive windows, even if the original HR peak gets stronger later in time.

position, since in the latter case, the MA peaks are large in number and are scattered all over the spectrum.

- 7) Since actual HRs do not show a sudden discontinuity, this provides with an important *a priori* information which allows for a provision of tracking mechanism. This gives good results most of the time, but for intense MA, this can, in fact, mislead the HR estimates go astray. For example, sometimes when a MA peak is close to the HR peak and is strong in magnitude and, thus, is mistaken to be a HR estimate, the tracking mechanism would follow the MA peaks in successive time windows, and would not allow for a come around to the correct track, even if the original HR peak gets stronger later in time. Thus, stability of the algorithm is sacrificed as too much confidence is placed upon the tracking mechanism resulting in a large runaway error as shown in Fig. 5.

Though the difficulties jotted above can certainly be reduced by applying various signal processing techniques, there are always chances for them to arise in certain time windows. In the next section, we discuss ways to minimize their occurrences and countermeasures to overcome them should they arise anyway.

### III. MATERIALS AND METHODS

In this section, we illustrate the key features to be used in our algorithm and expound on their effectiveness and functionality.

#### A. Data Acquisition and Modeling

For benchmarking purpose, we run our algorithm on the same datasets used in [13] and [14] where 12 healthy male subjects (age ranging from 18 to 35) ran on a treadmill with speeds reaching up to 15 km/h. Among the 12 datasets, we find datasets 1, 2, 4, and 10 the most corrupted ones and take them as our training data. The remaining datasets were used as testing purpose once the algorithm along with its parameters was derived from the training phase. Moreover, we run the same algorithm on a new dataset where the subject, instead of running, performed intensive forearm and upper arm movements (e.g., boxing). We also run the algorithm for a subject (age 58, female) with abnormal heart rhythm and blood pressure. For all datasets, the PPG signals were recorded from the wrist by two pulse oximeters with green LEDs (wavelength: 515 nm). Their distance

(from center to center) was 2 cm. The acceleration signal was also recorded from the wrist by a triaxis accelerometer. Both the pulse oximeters and the accelerometer were embedded in a wristband, which was comfortably worn. In order to calculate the ground truth HR for evaluating the algorithm's performance, ECG signal was recorded simultaneously from the chest using wet ECG sensors.

The proposed algorithm can estimate HR from either single-channel or two-channel PPG signals. Let us describe the algorithm for the two channel case and let each channel come with successive time windows of length  $M$  with some overlap  $L$ . Then, we can model the motion-corrupted raw PPG segments  $y_{i,\text{raw}}(n)$  ( $i = 1, 2$ ) by

$$y_{i,\text{raw}}(n) = d_i(n) + v_i(n), \quad n = 0, 1, \dots, M - 1$$

where  $d_i(n)$  is the desired clean PPG signal and  $v_i(n)$  is the noise introduced by motion. Let the three-axis raw acceleration data for the same time window be denoted by  $a_{x,\text{raw}}(n)$ ,  $a_{y,\text{raw}}(n)$ , and  $a_{z,\text{raw}}(n)$ . Since human HRs usually reside between 40 and 200 beat per minute (BPM), we bandpass filter the PPG and the acceleration segments in this frequency range to obtain  $y_i(n)$ ,  $a_x(n)$ ,  $a_y(n)$ , and  $a_z(n)$ , respectively. Also, let us assume that the sampling frequency of all the signals is  $F_s$ . Like Zhang *et al.* [13], [14], we also consider 8-s-long data windows with 75% overlap (this implies  $M = 1000$  and  $L = 750$  for  $F_s = 125$  Hz).

### B. Ensemble Empirical Mode Decomposition

In this paper, we use EEMD as a tool for signal decomposition to separate the clean PPG signal  $d_i(n)$  from the MA  $v_i(n)$ . It is reported that EEMD works very well for cleansing biomedical signals for postprocessing [15], [16]. EMD is a nonlinear signal processing technique that decomposes a given signal into a number of intrinsic mode functions (IMF). The IMFs satisfy the following two conditions: 1) the number of extrema and the number of zero crossings over the whole signal length must either be the same or differ at most by 1, and 2) the mean value of the envelope defined by the maxima and the envelope defined by the minima must be zero at all points in the signal.

For a given bandpass-filtered PPG signal  $y(n)$  (let us drop the channel index  $i$  in this section), its EMD is implemented using a sifting process that first finds all the local maxima and minima, and then connects them using cubic splines to create two corresponding envelopes: One upper envelope for the maxima and one lower envelope for the minima. The average of the two envelopes  $m(n)$  is then subtracted from the original signal  $y(n)$  to obtain a new signal

$$h(n) = y(n) - m(n).$$

This new signal is then put in place of  $y(n)$  and the above process is repeated until  $h(n)$  satisfies the two conditions of an IMF. When this occurs,  $h(n)$  becomes the first IMF  $c_1(n)$ .

Another round of sifting process is repeated on the residual signal  $y_1(n) = y(n) - c_1(n)$  to sift the next IMF. The whole process is halted when the residual signal  $y_K(n)$  after the extraction of the  $K$ th IMF becomes a monotonic function. The

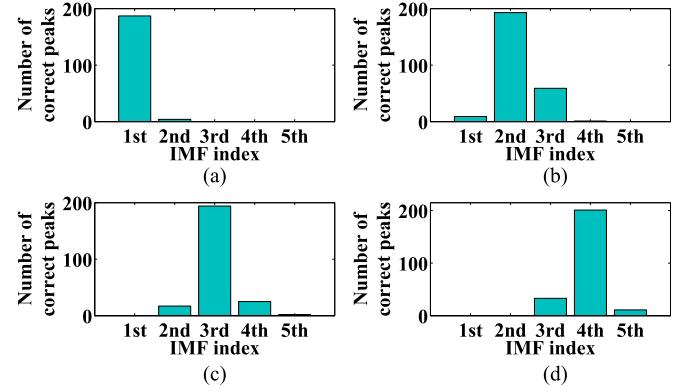


Fig. 6. Bar plot of the number of time windows of dataset 4 for which the maximum peak of the periodogram of an IMF gives the correct HR at sampling frequencies (a)  $F_s = 25$  Hz, (b) 125 Hz, (c) 250 Hz, and (d) 500 Hz. From this plot, it is evident that the dominant IMF for the clean PPG mode changes with  $F_s$ .

IMFs have the property that they, when added together with the final residual signal, can reconstruct the original signal

$$y(n) = \sum_{j=1}^K c_j(n) + y_K(n).$$

The EMD algorithm is, however, quite sensitive to mode mixing where an IMF includes uncontrolled oscillations or transient spectral content. To address mode mixing, an extension of the EMD algorithm, known as EEMD is proposed in [16]. Here, an ensemble of  $N_E$  signals is created from the given signal  $y(n)$  by adding white Gaussian noise  $w_p(n)$  of the same variance ( $p = 1, \dots, N_E$ ). That is, the ensemble contains

$$\tilde{y}_p(n) = y(n) + w_p(n), \quad p = 1, \dots, N_E.$$

Next, the EMD algorithm is applied on each of them to extract the IMFs  $c_{pj}(n)$  ( $j = 1, \dots, K$ ) so that

$$\tilde{y}_p(n) = \sum_{j=1}^K c_{pj}(n) + y_{pK}(n), \quad p = 1, \dots, N_E$$

$y_{pK}(n)$  being the residual function for  $\tilde{y}_p(n)$ . Finally, the optimum choice of the IMF is taken as the ensemble average

$$\bar{c}_j(n) = \frac{1}{N_E} \sum_{p=1}^{N_E} c_{pj}(n), \quad j = 1, \dots, K.$$

In the context of PPG, we find that if the PPG signal contains HR information, then it is usually in a specific IMF for a particular sampling frequency. Fig. 6 illustrates this point where we bar plot the number of time windows of dataset 4 for which the maximum peak of the periodogram of a certain IMF coincides with the ground truth HR. From this figure, we see that the dominant PPG mode is the first IMF for sampling frequency  $F_s = 25$  Hz [see Fig. 6(a)], the second IMF for  $F_s = 125$  Hz [see Fig. 6(b)], the third IMF for  $F_s = 250$  Hz [see Fig. 6(c)], and the fourth IMF for  $F_s = 500$  Hz [see Fig. 6(d)].

Fig. 7 illustrates the use of EEMD to extract HR information from ECG and PPG signals (here,  $F_s = 125$  Hz is used). Fig. 7(a) shows an ECG segment along with its periodogram in Fig. 7(b).

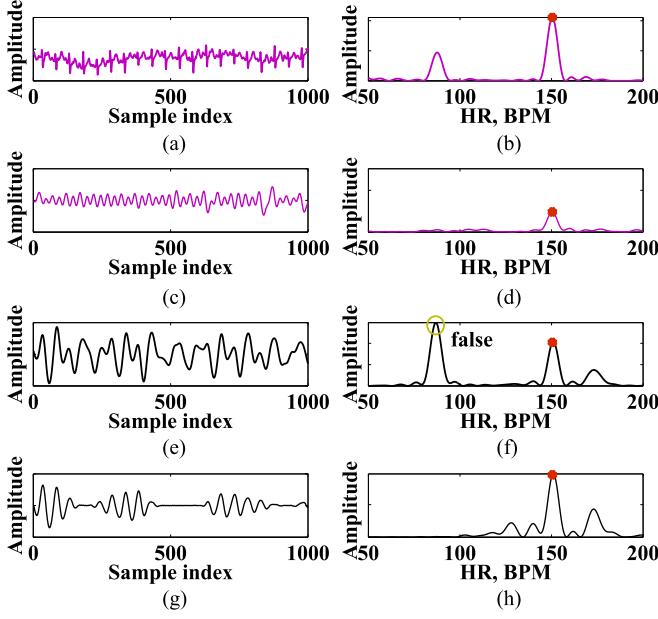


Fig. 7. Illustration of the effectiveness of using EEMD to extract the HR information from the noisy PPG signal. Here, both the ECG and the noise-corrupted PPG signal segments contain the HR information in their second IMFs (sampling frequency  $F_s = 125$  Hz). (a) Original ECG. (b) Periodogram of original ECG. (c) Second IMF of ECG. (d) Periodogram of second IMF of ECG. (e) Original PPG. (f) Periodogram of original PPG. (g) Second IMF of PPG. (h) Periodogram of second IMF of PPG.

If EEMD is performed on this segment, then we get the second IMF as shown in Fig. 7(c). Clearly, its maximum peak coincides with the ground truth HR shown as a red dot in the figure. Now, if we look at the PPG signal for the same time window, we see that it is corrupted by MA in both time [see Fig. 7(e)] and frequency domain [see Fig. 7(f)]. Its maximum peak (yellow circled in the figure) does not give the correct HR since it comes from MA. But if we perform EEMD on it to extract its second IMF [see Fig. 7(g)], we find that its maximum peak now coincides with the HR.

### C. Absolute Criterion

As discussed in Section II, stability and tracking mechanism work in opposite directions. To make sure that the proposed algorithm is stable, it should have a definite criterion which can allow to overthrow the previous track of HR estimates and jump to a distant estimate. This way, if for some reason, the HR estimates get derailed, it can still make it back to the original track and, thus, avoid any chance of runaway error. Also, no initialization would be required in this case, since no matter where it started from, it can still have the opportunity to catch up with the track should the criterion is fulfilled. We propose this AC to be as follows:

*If the PPG signal, after being suitably cleansed, shows a reliable peak, then this peak is assumed to correspond to the original HR.*

The existence of such an AC is vouched by the fact that a given signal no matter how much contaminated by MA, after suitably

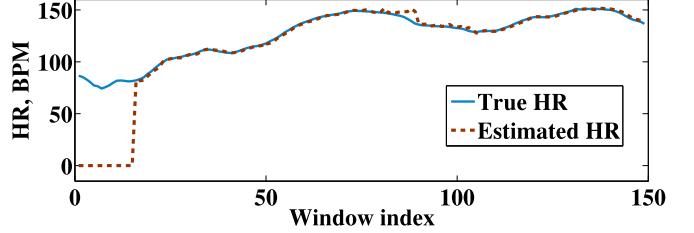


Fig. 8. Illustration of the effectiveness of using an AC which enables the algorithm in catching up with the original HR, and therefore, discards the requirement of an initial resting boot phase.

cleansed, is very unlikely to not give a strong reliable peak ever in any of its numerous time windows. In fact, our observation has shown that, even in statistical sense, strong correct peaks emerge quite frequently even when the PPG signals are heavily corrupted by MA, and these peaks are even more frequent if the cleansing process is effective.

This way, the reliable peaks can stand as checkpoints for the estimated HR track. Fig. 8 illustrates the effectiveness of this criterion which shows that even though we deliberately initialized the HR estimate to be 0 BPM, the algorithm after a while finds one correct reliable peak to meet the AC. So it takes the huge leap to move onto the correct track. In this paper, we use EEMD-based denoising technique to “suitably cleanse” the PPG signals for the AC condition, which is discussed at length in Section IV. However, the process of suitably cleansing the signal and the condition of identifying a peak as reliable may vary from algorithm to algorithm.

### D. RLS Filter

When the AC in our algorithm is not fulfilled, that is, when EEMD fails to provide IMFs with a reliable HR peak, we resort back to the raw PPG signals  $y_{i,\text{raw}}(n)$  and try to denoise their average  $\bar{y}_{\text{raw}}(n)$  using the raw acceleration data. Toward this end, we model the noise  $v(n)$  present in  $\bar{y}_{\text{raw}}(n)$  by

$$v(n) = v_x(n) + v_y(n) + v_z(n)$$

where  $v_\gamma(n)$  is the MA contribution correlated with the acceleration data  $a_{\gamma,\text{raw}}(n)$  ( $\gamma = x, y$ , and  $z$ ). We find that RLS adaptive FIR filters with length  $s$  work quite well to obtain the estimates  $\hat{v}_\gamma(n)$  for the noise components  $v_\gamma(n)$  from the reference signals  $a_{\gamma,\text{raw}}(n)$ .

Fig. 9 illustrates the whole process of noise cancellation. First, we estimate for  $v_x(n)$  from  $a_{x,\text{raw}}(n)$  by minimizing the least squares error

$$\xi(n) = \sum_{q=0}^n \lambda^{n-q} |e_x(q)|^2$$

where  $\lambda$  is the forgetting factor (we use  $\lambda = 1$ ) and  $e_x(n)$  is the error signal  $e_x(n) = \bar{y}_{\text{raw}}(n) - \hat{v}_x(n) = \bar{y}_{\text{raw}}(n) - a_{x,\text{raw}}(n) * w_x(n)$  with  $w_x(n)$  being the RLS filter coefficients. The coefficients are obtained adaptively by using a set of update equations given in [17]. Note that the error signal found from this stage  $e_x(n)$  is assumed to not contain MA contribution correlated

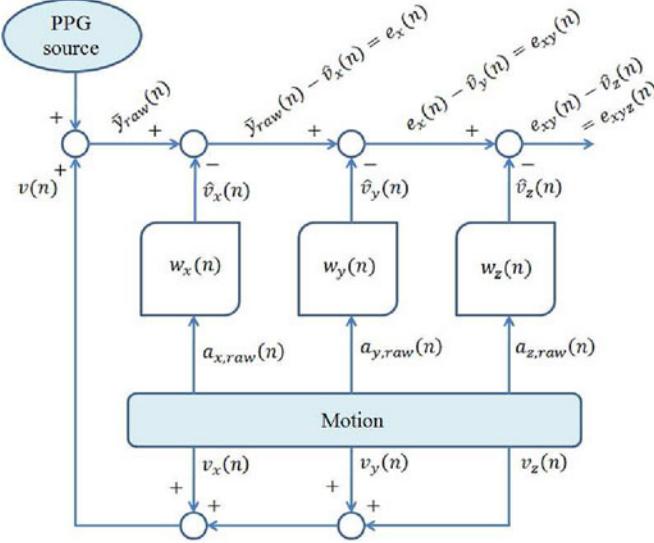


Fig. 9. RLS filter implementation to remove the different noise components correlated with the three-axis acceleration data.

with  $a_{x,\text{raw}}(n)$ . Next, another RLS filter  $w_y(n)$  is implemented to estimate for  $v_y(n)$  by treating  $a_{y,\text{raw}}(n)$  as the reference signal and  $e_x(n)$  as the desired signal. The error signal obtained in this stage  $e_{xy}(n) = \bar{y}_{\text{raw}}(n) - \hat{v}_x(n) - \hat{v}_y(n)$  now supposedly contains MA contribution coming from only  $a_{z,\text{raw}}(n)$ . In order to get rid of it too,  $a_{z,\text{raw}}(n)$  is now treated as the reference signal, while  $e_{xy}(n)$  as the desired signal to obtain another set of RLS filter coefficients  $w_z(n)$ . The final error signal  $e_{xyz}(n) = \bar{y}_{\text{raw}}(n) - \hat{v}_x(n) - \hat{v}_y(n) - \hat{v}_z(n)$  can be regarded as a denoised signal  $r_{\text{raw}}(n)$  which is assumed to have no correlation with the acceleration. This signal is then investigated further for HR peaks which we elaborate in detail in Section IV.

Fig. 10 illustrates the effectiveness of using RLS filters when EEMD fails to provide with a correct reliable peak. Fig. 10(a) shows an ECG segment ( $F_s = 125$  Hz), and the corresponding PPG segment is given in (b) along with its periodogram in (c). Note that the PPG periodogram is quite similar to the typical case illustrated in Fig. 1. The periodogram of the dominant IMF containing the clean PPG mode (second IMF for  $F_s = 125$  Hz) is shown in Fig. 10(d). Clearly, it fails to provide with correct HR since its maximum peak coincides with an MA peak as given by the periodogram of the acceleration data in Fig. 10(e). Finally, Fig. 10(f) shows the recovery of the true HR peak by the use of RLS filters as described above.

### E. Time-Domain Extraction

In our investigation of many sample datasets, we noted that for certain time windows, not all part of the signal segment is corrupted equally by MA. In fact, some part of the window might be clean enough that just cropping out that portion and applying the algorithm on it gives close estimates of HR. But, if the whole data window is taken for consideration, then MA is allowed to dominate the PPG spectrum and it gets difficult to detect the correct peak.

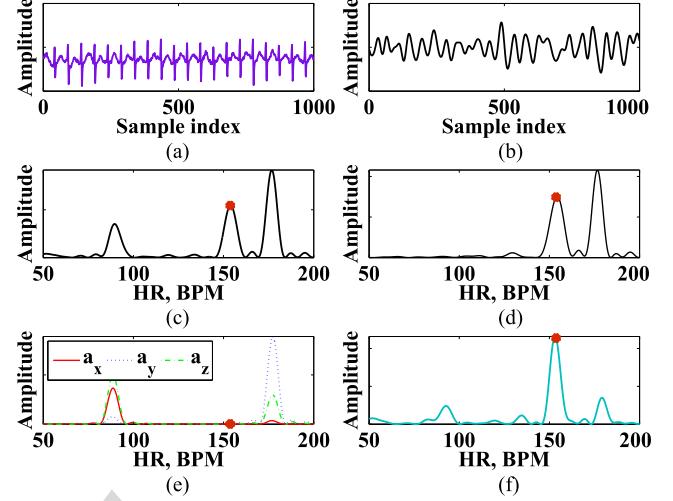


Fig. 10. Illustration of the effectiveness of using RLS filters when EEMD fails. Here, the clean PPG mode IMF (second IMF for  $F_s = 125$  Hz) does not contain the HR at its maximum peak location. Instead, it coincide with an MA peak as can be clearly seen by the periodograms of the accelerations. If RLS filters are applied using the acceleration signals as references, then maximum peak of the filter output coincides with the HR. (a) Original ECG. (b) Original PPG. (c) Periodogram of original PPG. (d) Periodogram of 2nd IMF of PPG. (e) Periodogram of acceleration. (f) Periodogram of PPG after RLS.

In order to decide on the location and the length of such window portions, the three-axis acceleration data can be used because they give a fair estimate of the amount of noise being injected. We found out that the net acceleration

$$a(n) = [a_x^2(n) + a_y^2(n) + a_z^2(n)]^{1/2}, \quad n = 0, 1, \dots, M-1$$

quite vividly marks out these portions: First, an interval  $I \subset \{0, 1, \dots, M-1\}$  is looked for so that

$$a(n) \leq \tau A \quad \forall n \in I$$

where  $A$  is the maximum value of  $a(n)$  in the whole time window and  $\tau$  is a suitable threshold usually kept around 0.3 [18]. Among all such intervals, if the largest one  $I_0$  is long enough to capture at least two heart beats (nearly 1 s long), then we crop the corresponding portion out of the original bandpass-filtered PPG signals  $y_i(n)$  by constructing

$$y_{i,\text{cropped}}(n) = y_i(I_0).$$

Though such small portion gives very smeared periodogram and its peak is not exactly the actual HR throughout the entire time window under consideration, it gives very close and robust estimates.

Fig. 11 illustrates the effectiveness of this idea. Fig. 11(a) shows an ECG signal segment (sampled at 125 Hz) for a particular time window along with its periodogram shown in (b). Fig. 11(c) shows the corresponding PPG signal, whereas (d) shows its periodogram. Note that the PPG periodogram is quite similar to the case illustrated in Fig. 2 in that there is no peak corresponding to HR at all. In fact, applying a high-resolution denoising technique like singular spectrum analysis [13], for example, fails to provide any peak at the HR as shown in the next Fig. 11(e). This is partly because of the wide spectrum of

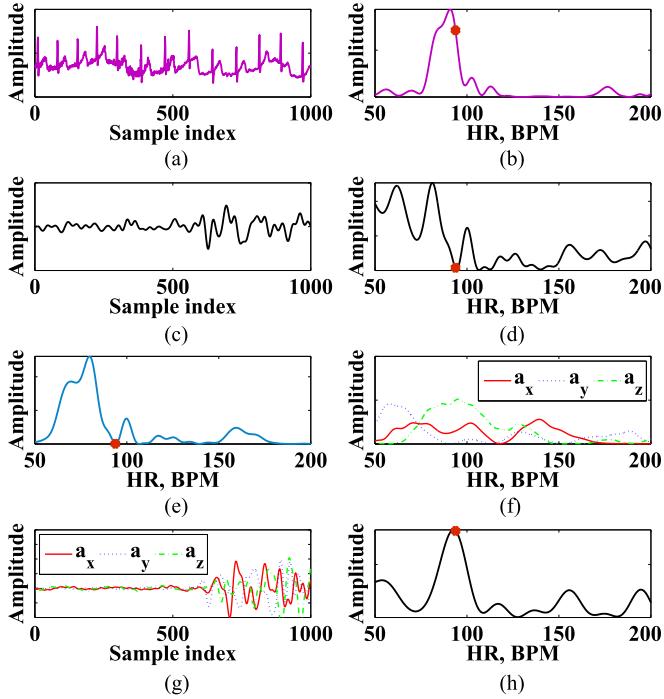


Fig. 11. Illustration of the effectiveness of time-domain extraction. (a) Original ECG. (b) Periodogram of ECG. (c) Original PPG. (d) Periodogram of PPG. (e) Periodogram of PPG after SSA. (f) Periodogram of acceleration. (g) Acceleration in time domain. (h) Periodogram of cropped PPG.

the MA, as shown by the periodograms of the three-axis acceleration data in Fig. 11(f). Now, if we observe the acceleration data in the time domain [see Fig. 11(g)], we immediately notice that for this particular time window we have a large portion of the signal where the MA is less intense compared to the rest of the window. Extracting that part of the PPG signal and just taking its periodogram yields a very strong peak corresponding to the true HR as shown in Fig. 11(h).

Aperiodic exercises, like boxing, can give large erroneous peaks if the entire window is considered since here the small amount of time where the hand plunges forward degrades the whole PPG spectrum. But if we extract only that portion of the time window where the hands are stowed to the chest between two successive punches, then quite correct estimates can be obtained. As mentioned earlier, estimating the HR during resting phase is often more difficult than in running phase. In these cases, this approach works quite well because sudden jolts of hands in rest phase can be filtered out by this approach.

#### IV. PROPOSED ALGORITHM

In this section, we attempt to encompass all the ideas illustrated above and give details of our proposed algorithm. The algorithm estimates HR through the following steps.

**Signal Workspace:** In this stage, we produce all the signals necessary to enter into our main algorithm. We are given two channels of raw PPG segments  $y_{i,\text{raw}}(n)$  ( $i = 1, 2$ ) and the corresponding raw acceleration segments  $a_{\gamma,\text{raw}}(n)$  ( $\gamma = x, y$  and  $z$ ) at a sampling frequency  $F_s$ . We apply the RLS filter block

(as illustrated in Fig. 9) on the average of  $y_{i,\text{raw}}(n)$  by using  $a_{\gamma,\text{raw}}(n)$  as the reference signals. In our experiments, we set the RLS filter length  $s$  around  $8 - 10$  for  $F_s = 25$  Hz, and it is increased for higher sampling frequencies. The partially denoised output  $r_{\text{raw}}(n)$  along with the PPG signals  $y_{i,\text{raw}}(n)$  and acceleration signals  $a_{\gamma,\text{raw}}(n)$  are then passed through an infinite impulse response bandpass filter with lower and upper passband cut off at 40 and 200 BPM, respectively (typical frequency range for human HR), and lower and upper stopband cut off at 35 and 205 BPM, respectively, and passband ripple 0.01, stopband attenuation 80 dB to obtain  $r(n)$ ,  $y_i(n)$ , and  $a_{\gamma}(n)$ , respectively.

**Initialization:** The main algorithm estimates HR by making decisions based on the previous HR estimate  $f_{\text{prev}}$ . But if the data window under consideration is the first one, there is no  $f_{\text{prev}}$  and we need to estimate it first. We do so as follows.

- 1)  $r(n)$  is inspected in Fourier domain and its dominant peaks' locations (peaks larger than 80% of the maximum peak) are arrayed together in a set  $S_{\text{rls}}$ . If  $|S_{\text{rls}}| = 1$  (here  $|\cdot|$  represents the cardinality of a set), which means  $S_{\text{rls}}$  contains the location of only the maximum peak, then we assign that to  $f_{\text{prev}}$ .
- 2) If  $|S_{\text{rls}}| \neq 1$ , then there are multiple strong peaks present. Therefore, we look for their second harmonics in  $S_{\text{rls}}$  and build a new set  $H_{\text{rls}} \subset S_{\text{rls}}$  so that if  $x \in H_{\text{rls}}$ , then there exists a  $y \in S_{\text{rls}}$  such that  $|2x - y| < 5$  BPM. If  $H_{\text{rls}}$  is not empty, then the strongest peak is assigned to  $f_{\text{prev}}$ .
- 3) If  $|H_{\text{rls}}| = 0$ , then we first construct a set  $S_{a, 0.8}$  by taking all the dominant peaks (80% of the maximum peak) from the acceleration signals  $a_{\gamma}(n)$ , and then cross out all the peaks in  $S_{\text{rls}}$  which are close to the acceleration peaks stored in  $S_{a, 0.8}$  within a 5 BPM range by forming  $S = S_{\text{rls}} \setminus^5 S_{a, 0.8}$ . Here for notational convenience, we define  $A \setminus^{\delta} B = \{x \in A : |x - y| > \delta \text{ BPM } \forall y \in B\}$ . Next, if  $S$  is not empty, the strongest peak in  $S$  is assigned to  $f_{\text{prev}}$ .
- 4) If  $f_{\text{prev}}$  is still not assigned (which is a rare case), then we assign its value to the strongest peak's location in  $S_{\text{rls}}$ . As we shall see later, even if  $f_{\text{prev}}$  is erroneously initialized, our main algorithm is designed in such a way that it can correct for it later and come back on track.

With  $f_{\text{prev}}$  assigned, we now enter the main algorithm.

**Main Algorithm:** Here, a hierarchy of steps is executed to obtain a crude estimate  $f$  for the HR as described below.

- 1) **AC:** First EEMD is applied to  $y_i(n)$  with  $N_E \simeq 5$  and SNR level around 30 dB. Among the obtained IMFs  $\bar{c}_{i,j}(n)$  ( $i$  being the channel index and  $j$  being the IMF index), only one IMF is considered per each channel. This selection is based on the sampling frequency (e.g.,  $j = 1$  for  $F_s = 25$  Hz,  $j = 2$  for  $F_s = 125$  Hz,  $j = 3$  for  $F_s = 250$  Hz, and  $j = 4$  for  $F_s = 500$  Hz). Then, the two selected IMFs obtained from both channels are inspected in Fourier domain and their maximum peaks' locations are put in a set  $S_{\text{imf}}$ . Similarly, we construct another set  $S_{a, 0.5}$  by taking all the dominant peaks (50% of the maximum peak, a low threshold to ensure capturing all MA peaks) from the acceleration signals  $a_{\gamma}(n)$ . Next,  $g$  groups are

constructed from the elements of  $S_{\text{imf}}$  such that they all contain peak locations in clusters of at most 2 BPM range and each of them has at least one peak that is not close to an acceleration peak. In other words, if  $G$  is such a group, then  $\max(G) - \min(G) \leq 2 \text{ BPM}$  and  $|G| \geq S_{a, 0.5} \neq 0$ . Groupwise averages are taken next, and from these  $g$  averages, the one closest to  $f_{\text{prev}}$ , say  $f_{\text{AC}}$  is considered. If  $|f_{\text{AC}} - f_{\text{prev}}| < \Delta_{\text{AC}}$  with  $\Delta_{\text{AC}}$  being a suitable range to be discussed shortly, then we are confident in assuming that this peak, which is present in one of the chosen IMFs and also does not lie close to any MA peak, is an intrinsic peak of the desired PPG signal. We further assume that the peak is reliable enough to meet the AC condition in this case and consider the peak as the crude estimate,  $f = f_{\text{AC}}$ . Each time the AC is not met, the range of  $\Delta_{\text{AC}}$  aforementioned is increased dynamically by an amount  $\Delta_d$  from its default value  $\Delta_0$ . That is, if  $f$  is not assigned any value from this step  $m$  times in a row, then we will have for the next time window  $\Delta_{\text{AC}} = \Delta_0 + m\Delta_d$ . However if assigned, then the range of  $\Delta_{\text{AC}}$  is reset to its default value  $\Delta_0$ . In our experiments, we set  $\Delta_0 \simeq 5 \text{ BPM}$  and  $\Delta_d \simeq 1 \text{ BPM}$ .

- 2) *Tracking from IMFs*: From this step onward, we employ different tracking mechanism. First, we construct  $S_{\text{imf}} \setminus S_{a, 0.5}$ , and from this set, we take the peak location nearest to  $f_{\text{prev}}$ . If its distance from  $f_{\text{prev}}$  is within  $\Delta_{\text{imf}} \simeq 7 \text{ BPM}$ , we assign it to  $f$ , and since this peak comes from IMFs corresponding to clean PPG modes, we also reduce the range of  $\Delta_{\text{AC}}$  set for the next time window by  $\Delta_d$ .

3) *Tracking from time-domain extraction*: In this step, we check whether the signals  $y_i(n)$  are suitable for time-domain extraction. If so, we take the respective clean portions  $y_i, \text{cropped}(n)$  and find their maximum peaks' locations in Fourier domain. The peak location nearer to  $f_{\text{prev}}$  is considered and checked to see if its distance from  $f_{\text{prev}}$  is also within a range  $\Delta_t$  (around  $7 - 12 \text{ BPM}$ ). If so, we assign it to  $f$ .

4) *Tracking from RLS signal*: Here, we consider  $r(n)$ , and similar to the initial step, we put its dominant peaks (80% of the maximum) in  $S_{\text{rls}}$ . We also construct  $S_{a, 0.6}$  by taking the dominant peaks (60% of the maximum, a moderate threshold for tracking purpose) from  $a_\gamma(n)$ . Now, if the set  $S_{\text{rls}} \setminus S_{a, 0.6}$  contains only one peak  $f_{\text{rls}}$  and if  $|f_{\text{rls}} - f_{\text{prev}}| < 25 \text{ BPM}$ , then we assign  $f = f_{\text{rls}}$ . However, if this is not the case, then the strongest peak in  $S_{\text{rls}}$  is looked for such that it lies close to  $f_{\text{prev}}$  within a range  $\Delta_{\text{rls}}$  ( $7 - 12 \text{ BPM}$ ). If found, we assign it to  $f$ .

5) *Tracking from the original signal*: If the above steps fail to provide with the crude estimate  $f$ , then we consider all the peak locations attainable from the periodograms of  $y_i(n)$  and array them together in a set  $S_{\text{org}}$ . We also construct another set  $S_{a, 0}$  by taking all the peaks (no threshold in this case) from  $a_\gamma(n)$ . Next, we take the peak  $f_{\text{org}} \in S_{\text{org}}$  which is closest to  $f_{\text{prev}}$  and also  $|f_{\text{org}} - f_{\text{prev}}| \leq 5 \text{ BPM}$ . Then, if  $f_{\text{org}}$  does not lie close to any acceleration peak stored in  $S_{a, 0}$  within a 3 BPM tolerance, that is, if  $f_{\text{org}} \in S_{\text{org}} \setminus S_{a, 0}$ , then we assign  $f = f_{\text{org}}$ . If this is not the case,

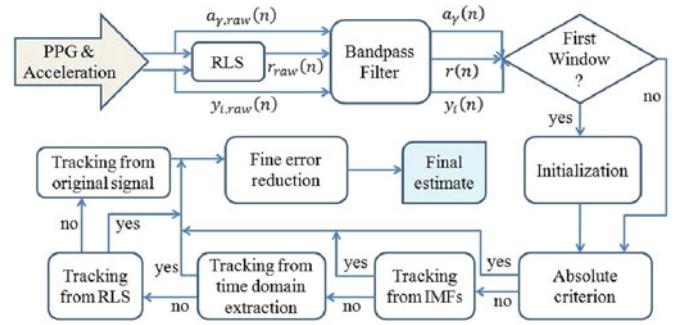


Fig. 12. Flowchart of the proposed algorithm.

then we reach the conclusion that the given PPG segments are highly corrupted, and in both the original segments and the denoised segments, there is no discernible peak.

Therefore, we discard the segment from our consideration by assigning  $f = f_{\text{prev}}$ .

*Fine Error Reduction:* This is the final step of our algorithm where we attempt to obtain a fine estimate  $f_{\text{est}}$  from the crude estimate  $f$  by working with the raw PPG data. To do so, we construct a set  $S_{\text{raw}}$  by taking all the peak locations attainable from the periodograms of  $y_{i,\text{raw}}$ . Then, we look for the peak  $f_{\text{raw}} \in S_{\text{raw}}$  which is closest to  $f$  and also  $|f_{\text{raw}} - f| \leq \Delta_{\text{fine}}$  with  $\Delta_{\text{fine}}$  around 3 – 4 BPM. If such a peak is found, then we assign our fine estimate  $f_{\text{est}} = f_{\text{raw}}$ . Otherwise, we consider our crude estimate as the final estimate  $f_{\text{est}} = f$ . Also, in this step, we assign  $f_{\text{prev}} = f_{\text{est}}$  for the next time window.

Therefore, the algorithm works in several key steps. In the first step of the main algorithm, robustness is ensured with the EEMD-based signal cleansing followed by the application of AC. Each time the AC is not met, the range of  $\Delta_{AC}$  is increased with no considerable limit and chances of meeting the AC in the next window increases, resulting in the algorithm's ability to catch up with the correct track and avoid a runaway scenario. Therefore, this step enables the algorithm to become robust to initial wrong estimates, and thereby, helps the algorithm to not require an initial resting phase. If this step is bypassed, then in the following series of steps tracking is emphasized. Here, previous HR estimate is given full priority and all nearby peaks attainable from IMFs, RLS output, time-domain extracts, and original bandpass-filtered signals are taken into consideration. In the last stage, the algorithm reduces the fine error by working with the raw PPG segments to better coincide with the ground truth. For ease of understanding, a flowchart of the algorithm is given in Fig. 12. Note that though the above algorithm is described explicitly for two channel PPG signals, it can be generalized for any number of channels. For single-channel implementation, simply letting  $y_{1,\text{raw}} = y_{2,\text{raw}}$  gives the corresponding framework.

## V. EXPERIMENTAL RESULTS

#### A. Performance Evaluation

Table I lists the average absolute error for each of the 12 datasets (sampled at  $F_s = 25$  Hz) for both the double-channel

TABLE I  
PERFORMANCE COMPARISON OF DIFFERENT ALGORITHMS USING AVERAGE ABSOLUTE ERRORS FOR THE 12 TRUNCATED DATASETS USED IN [14] FOR  $F_s = 25$  Hz

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	Subj 11	Subj 12	Mean $\pm$ SD
Proposed- <i>d</i>	1.70	0.84	0.56	1.15	0.77	1.06	0.63	0.53	0.52	2.56	1.05	0.91	1.02 $\pm$ 1.79
Proposed- <i>s</i>	1.77	1.94	0.73	1.19	0.51	1.09	0.52	0.43	0.36	3.43	0.89	0.98	1.15 $\pm$ 2.37
JOSS [14]	1.33	1.75	1.47	1.48	0.69	1.32	0.71	0.56	0.49	3.81	0.78	1.04	1.28 $\pm$ 2.61
TROIKA [13]	3.05	3.31	1.49	2.03	1.46	2.35	1.76	1.43	1.28	5.08	1.80	3.02	2.34 $\pm$ 2.86

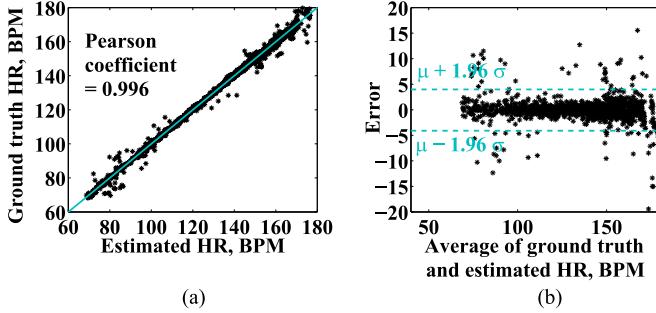


Fig. 13. (a) Pearson correlation between the ground truth and the estimated HRs over all the 12 datasets. (b) Bland-Altman plot over all the 12 datasets. Here,  $\mu$  is the mean of the error differences over all the estimates, and  $\sigma$  is the corresponding SD.

( $i = 1, 2$ ) and the single-channel ( $i = 1$ ) cases. In the same table, we also include the errors found by the methods of Zhang *et al.* [13], [14], for ease of comparison (HR estimates of JOSS and TROIKA were obtained through personal correspondence with the author). For the double-channel case (i.e., Proposed-*d*), averaged over all the estimates in the 12 datasets a gross average absolute error of 1.02 BPM is obtained with a standard deviation (SD) of 1.79 BPM. Here, the percentage error (absolute error/ground truth) gives an average of 0.79%. The Pearson correlation is calculated to be 0.996 as shown in Fig. 13(a). Fig. 13(b) gives the Bland-Altman plot [19] and the limit of agreement (LOA) is found to be  $[\mu - 1.96\sigma, \mu + 1.96\sigma] = [-4.10, 3.98]$  BPM ( $\mu$  being the mean of the error differences over all the estimates and  $\sigma$  being the corresponding SD).

Also note that for the single-channel case (i.e., Proposed-*s*), a gross average absolute error of 1.15 BPM is obtained (SD: 2.37 BPM) which is less than both JOSS and TROIKA's benchmarks. Here, the parameters are kept the same as for the double-channel case (obtained from the training phase). Since JOSS and TROIKA also work on single-channel PPG segments, the improvement in the single-channel result attests for the algorithmic superiority of the proposed method.

The error figures reported in Table I are for the datasets used in [14], which are not exactly the same as had been used in [13]. This is so because among the 12 datasets, 6 were very corrupted in the beginning and the method described in [14] did not initialize at these segments. Therefore, in [14], these initial corrupted segments were excluded and results were reported on a truncated version of the datasets. In Table II, we report the performance of our algorithm on the original untruncated datasets as well, along with the errors of TROIKA. Clearly, the

proposed algorithm, for both single channel and double channel, performs quite well even with noisy starts. This attests for the algorithm's needlessness of initial resting phase.

To illustrate the robustness of the algorithm against a change in exercise type, we applied the same algorithm along with the same set of parameters (obtained from the training phase) to a new test dataset where the subject performed boxing, hands shaking, jumps, and push-ups. The average absolute error was found to be 1.09 BPM [see Fig. 14(a)]. We also implemented TROIKA and found the corresponding error 1.97 BPM.

All the 12 datasets used previously were recorded for 12 healthy male subjects (age: 18–35). To illustrate the robustness of the algorithm against a change in subject profile, we ran our algorithm (with the same parameter set) for a female elderly subject (age: 58) with abnormal heart rhythm and blood pressure. We obtained an average absolute error of 0.66 BPM [see Fig. 14(b)], whereas the corresponding error for TROIKA was found 0.75 BPM. Therefore, it can be said that the algorithm performs well for these new test datasets as well.

### B. Sensitivity Analysis

To examine the performance of the algorithm for a change in the sampling frequency  $F_s$ , we apply the same algorithm for the two channel case at three more different sampling frequencies, e.g., 125, 250, and 500 Hz. The corresponding results are listed in Table III which shows that the algorithm performs equally well in higher sampling frequencies.

To illustrate the sensitivity of the algorithm's performance to its parameters, we run the simulations again (for double channels, at  $F_s = 25$  Hz) on all the 12 truncated datasets by altering the crucial parameters, namely  $N_E$  (number of realizations in an ensemble for performing EEMD),  $\Delta_0$  (default value of  $\Delta_{AC}$ , the BPM window length for AC),  $\Delta_d$  (dynamic increment of  $\Delta_{AC}$ ),  $\Delta_{imf}$  (tracking range for the IMFs),  $\tau$  (the threshold for time-domain extraction),  $\Delta_t$  (tracking range for the time-domain extracts),  $s$  (filter length for RLS),  $\Delta_{rls}$  (tracking range for RLS filter output), and  $\Delta_{fine}$  (fine error threshold). In this paper, we used  $N_E = 5$ ,  $\Delta_0 = 5$  BPM,  $\Delta_d = 1$  BPM,  $\Delta_{imf} = 7$  BPM,  $\tau = 0.3$ ,  $\Delta_t = 12$  BPM,  $s = 8$ ,  $\Delta_{rls} = 9$  BPM, and  $\Delta_{fine} = 4$  BPM. To test for sensitivity for the parameter  $N_E$ , we keep all the other parameters unchanged while setting  $N_E$  to 3 or 7 from its default value 5. For each of these cases, the gross average absolute error for the 12 datasets is recorded in Fig. 15. Similarly, sensitivity analysis for other parameters is performed by setting  $\Delta_0 = 4$  or 7 BPM,  $\Delta_d = 1.5$  BPM,  $\Delta_{imf} = 5$  or 10 BPM,  $\tau = 0.4$ ,  $\Delta_t = 7$  or 10 BPM,  $s = 9$  or 10,  $\Delta_{rls} = 7$  or 11 BPM, and  $\Delta_{fine} = 3$  BPM each in turn. As Fig. 15 shows the

TABLE II  
PERFORMANCE COMPARISON USING AVERAGE ABSOLUTE ERRORS FOR THE 12 UNTRUNCATED DATASETS USED IN [13] FOR  $F_s = 25$  Hz

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	Subj 11	Subj 12	Mean ± SD
Proposed- <i>d</i>	1.64	0.81	0.57	1.44	0.77	1.06	0.63	0.47	0.52	2.94	1.05	0.91	1.07 ± 2.17
Proposed- <i>s</i>	2.55	3.45	0.73	1.19	0.51	1.09	0.52	0.43	0.36	3.33	0.89	0.98	1.33 ± 3.32
TROIKA [13]	3.05	3.49	1.49	2.03	1.46	2.35	1.76	1.42	1.28	5.73	1.79	3.02	2.41 ± 3.45

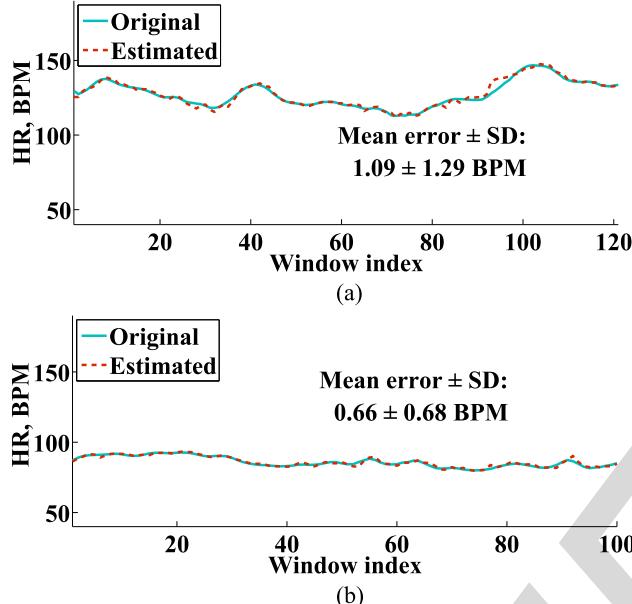


Fig. 14. (a) Performance of the proposed algorithm for a new set of datasets where the subject, instead of running, performed intensive forearm and upper arm movements (e.g., boxing). (b) Algorithm's performance for a subject (age 58, female) with abnormal heart rhythm and blood pressure.

performance remains more or less the same (with gross average absolute error around 1 BPM). This proves the algorithm's robustness to parameter values.

We also determine the algorithm's performance sensitivity to the various speed settings. Each of the 12 runs consisted of three types of phases: rest, moderate, and intense phases. The rest phase  $R$  was at the beginning and at the end of the run. The moderate phase  $M$  was the normal jogging condition at 6–8 km/h, whereas the intense phase  $I$  was at a more higher speed (12–15 km/h). Both the moderate and the intense phases were exerted twice in an alternating fashion. Therefore, the running protocol for speed sensitivity analysis is chosen to be

$$R \longrightarrow M \longrightarrow I \longrightarrow M \longrightarrow I \longrightarrow R.$$

Fig. 16(a) illustrates the algorithm's performance in each of these phases. Clearly, the estimated HR curve closely follows the ground truth HR in all the three phases in an identical manner. No discernible change in performance is noticed as the speeds are varied during the runs. Similar result also holds for the rest of the datasets.

Finally, Fig. 16(b) shows the algorithm's sensitivity to the presence or absence of one of its major subalgorithms, namely the use of AC. If the HR estimation goes wrong (in the figure, it is deliberately let carried away), clearly it is for this feature

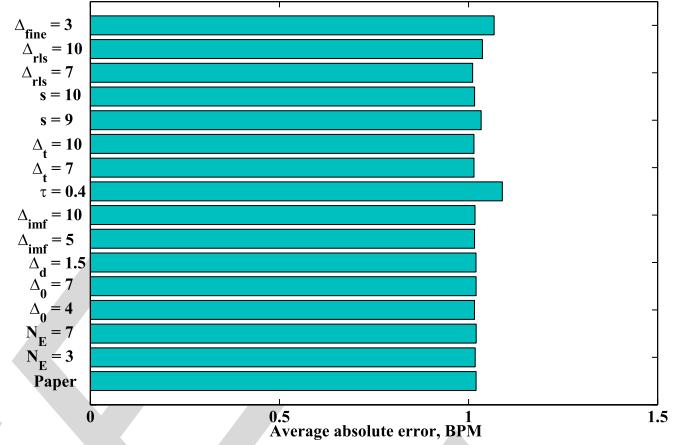


Fig. 15. Robustness of the proposed algorithm to parameter values. "Paper" indicates the result with parameter values given in Section V-B. " $N_E = 3$ " indicates the result with all parameter values unchanged except for  $N_E$  which is changed to 3 from its default value 5. Similar meanings go to other y-labels in the figure.

that the estimation is brought back ontrack and a total runaway situation is avoided.

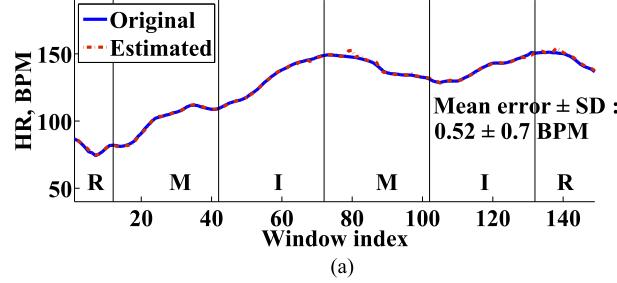
### C. Discussion

From the above figures and tables, it is quite clear that the proposed algorithm is better than the two recent works done in this field [13], [14]. It has a gross average absolute error superior to both TROIKA and JOSS, and the error variance is also significantly improved compared to both of the algorithms. TROIKA and JOSS reported LOA around  $[-7.26, 4.79]$  and  $[-5.94, 5.41]$ , respectively, whereas the proposed algorithm has a LOA of  $[-4.10, 3.98]$ . On the other hand, the Pearson correlation coefficient was found to be 0.992 in TROIKA and 0.993 in JOSS, while it is 0.996 in our proposed algorithm. But more importantly, increased robustness and stability together with no requirement of initial resting boot phase are the key features in the algorithm that we would like to highlight.

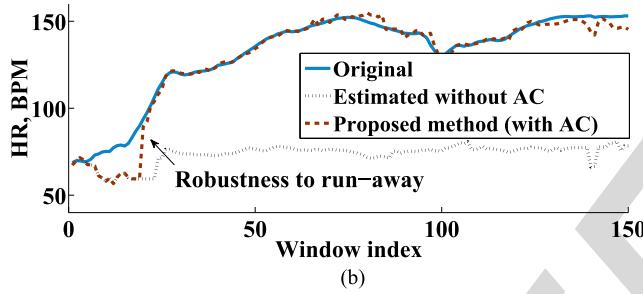
To better compare the performance of the proposed method with that of TROIKA and JOSS, Fig. 17 plots the histograms of the difference between the absolute error of the proposed method and the absolute error of either TROIKA [see Fig. 17(a)] and JOSS [see Fig. 17(b)] at each estimate over the 12 datasets, i.e., the histogram of  $e_{\text{proposed}}(l) - e_{\text{TROIKA}}(l)$  and  $e_{\text{proposed}}(l) - e_{\text{JOSS}}(l)$ , where  $e_{\text{proposed}}(l)$  indicates the absolute estimation error of the proposed method at the  $l$ th HR estimate, whereas  $e_{\text{TROIKA}}(l)$  and  $e_{\text{JOSS}}(l)$  indicate those of TROIKA and JOSS, respectively. From Fig. 17, it is clear that both histograms run downhill faster along the positive sides.

TABLE III  
AVERAGE ABSOLUTE ERRORS FOR THE 12 DATASETS USED IN [14] FOR DIFFERENT SAMPLING FREQUENCY  $F_s$

$F_s$ , Hz	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	Subj 11	Subj 12	Mean ± SD
25	1.70	0.84	0.56	1.15	0.77	1.06	0.63	0.53	0.52	2.56	1.05	0.91	1.02 ± 1.79
125	1.83	0.85	0.63	1.21	0.65	1.03	0.70	0.50	0.47	2.83	1.14	0.90	1.06 ± 2.02
250	1.96	0.93	0.64	1.19	0.68	1.00	0.73	0.49	0.50	2.95	1.21	0.95	1.10 ± 1.96
500	1.80	0.87	0.64	1.23	1.05	1.03	0.68	0.50	0.50	3.00	1.22	0.88	1.12 ± 2.11



(a)

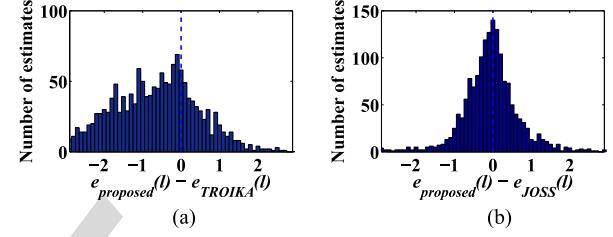


(b)

Fig. 16. (a) Performance of the proposed algorithm is illustrated at various speed settings. Here, R, M, and I keywords stand for “Rest,” “Moderate,” and “Intense” phases of the running protocol. (b) Algorithm shows quick robust turn back even though it is deliberately let carried away. This is due to the use of the AC, without which we have a runaway situation plotted in the same figure window.

This implies there are more estimates on the negative side; which in turn implies that for large number of estimates we have  $e_{\text{proposed}}(l) - e_{\text{TROIKA}}(l) < 0$  (75% of total estimates) and  $e_{\text{proposed}}(l) - e_{\text{JOSS}}(l) < 0$  (56% of the total estimates). Based on the  $t$ -test, the absolute estimation error of the proposed method is found to be significantly smaller than those of TROIKA and JOSS at the significance level  $\alpha = 0.01$ , and the corresponding  $p$  values are  $3.7415 \times 10^{-66}$  and  $3.1139 \times 10^{-5}$ , respectively.

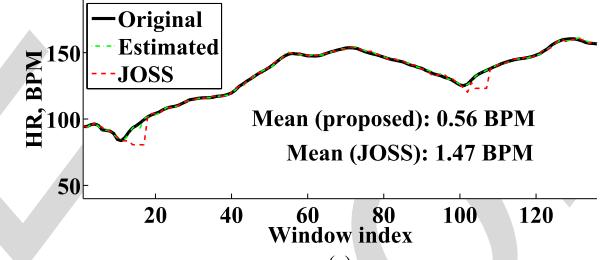
In Fig. 16(b), we see that even if the algorithm mistakes an MA peak as HR, as soon as a strong reliable peak is discovered later on, the correct HR is retained and a runaway situation is avoided. But even so, a considerable amount of offtrack error is produced thereby. In addition to making the algorithm robust to runaway situations, it is also imperative to make it robust against offtrack conditions. Now at the advent of offtracking, the HR peak and the MA peak (or its harmonics) usually get very close to each other. If proper spectral peak separation techniques are not used, then a wrongly chosen peak will give rise to an offtrack scenario. In our algorithm, we used, in addition to RLS filters, a novel technique to address this issue, namely time-domain extraction. Their combined performance is shown in Fig. 18. From Fig. 18(a), we can see the dataset is corrupted enough



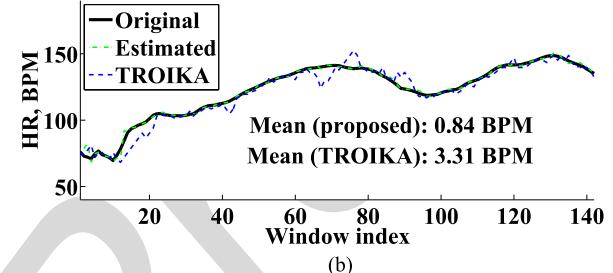
(a)

(b)

Fig. 17. Histograms of the difference between the absolute error of the proposed method and the absolute error of (a) TROIKA and (b) JOSS at each estimate over the 12 datasets. Based on the  $t$ -test, the absolute estimation error of the proposed method is found to be significantly smaller than those of TROIKA and JOSS at the significance level  $\alpha = 0.01$ , and the corresponding  $p$  values are  $3.7415 \times 10^{-66}$  and  $3.1139 \times 10^{-5}$ , respectively.



(a)



(b)

Fig. 18. Illustration of the proposed algorithm’s robustness against offtracking. (a) Proposed algorithm is successful in choosing correct peaks and avoid any offtracking pitfalls where JOSS failed. (b) Proposed algorithm is hardly offtracked, but TROIKA is derailed quite often.

to derail JOSS and incur some offtrack errors. But in the same pitfalls, the proposed algorithm was able to correctly choose the HR peaks and was not offtracked. In Fig. 18(b), we compare between the performances of the proposed algorithm and TROIKA. As we can see, TROIKA was offtracked quite frequently and had many kinks, whereas the proposed algorithm hardly lost track of HR.

Both TROIKA and JOSS estimate HR from a single channel of PPG. Though our algorithm for single channel (see Proposed-*s* in Table I) performs better than both of them, we find that

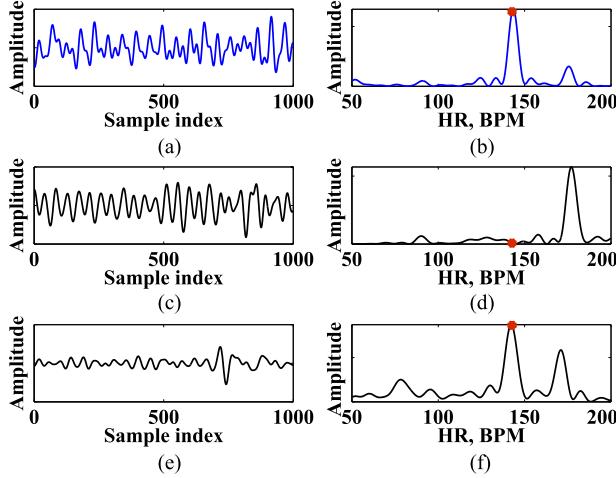


Fig. 19. Advantage of using two channels of PPG signals. Clearly channel 1 is distorted and does not contain HR information, whereas channel 2 readily gives the HR peak from its periodogram. (a) Original ECG. (b) Periodogram of ECG. (c) PPG Channel 1. (d) Periodogram of PPG Channel 1. (e) PPG Channel 2. (f) Periodogram of PPG Channel 2.

utilizing two channels of PPG gives even better results (see Proposed-*d* in Table I). The advantage of having an extra channel lies in the fact that even if one of the pulse oximeters loses contact with the skin or is heavily corrupted by internal shot noise, the other might still be functioning well and provide with the HR peak for that period of time. Since the proposed algorithm hardly mixes the two channels, instead it works with the peaks that they individually can provide with, the algorithm can estimate the HR even if one of the channels is malfunctioning. Obviously, any single-channel algorithm will be at a disadvantage in these situations. To illustrate this, Fig. 19(a) shows an ECG segment along with its periodogram in (b). Simultaneously recorded two segments of PPG signals are shown in Fig. 19(c) (channel 1) and (e) (channel 2) together with their periodograms shown in (d) and (f), respectively. Clearly, the PPG segment from channel 1 does not contain the HR information, but the PPG segment from channel 2 does.

Furthermore, both JOSS and TROIKA use signal sparsification techniques through the M-FOCUSS algorithm which is very much computationally expensive. For example, for  $F_s = 125$  Hz, TROIKA takes about 3.5 h to estimate HR for all the 12 datasets on a computer equipped with Intel Core-i7 4790 at 3.60 GHz, 8-GB RAM, Windows 7 64 bit, and MATLAB 2013a, whereas our proposed algorithm takes 668 s on the same computer. Similarly, on the same computer, for  $F_s = 25$  Hz, JOSS takes 300 s for all the datasets, whereas our proposed algorithm takes only 199 s.

Though all the algorithms discussed in this paper try to estimate HR instantaneously on the current time window, a significant improvement in the results can be achieved if the real-time implementation allows for a lag of around two time frames. With such a provision for delayed HR estimation, correct HR peaks can be chosen with more confidence. Fig. 20 illustrates this idea where our proposed algorithm is shown in two versions. In the original instantaneous version, there are some few inevitable

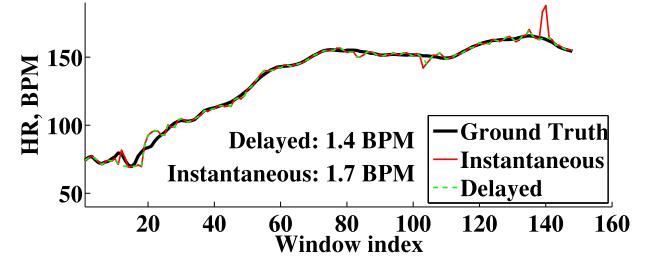


Fig. 20. Advantage of using a lag of just two time frames in real time. As opposed to the instantaneous algorithm, the delayed algorithm (delayed by 4 s) can identify outlier estimates and smooth them out.

kinks and sideways drifts in the HR estimation, whereas for the delayed version the small offtrack errors can be corrected in real time by identifying the outlier estimates and smoothing them out (in the figure we use delay of two frames which, in our present scenario, corresponds to a lag of only 4 s). Surely, such a small time lag is not as frightening to the user as displaying incorrect HR values.

## VI. CONCLUSION

This paper has proposed a new robust method for HR monitoring from dual-channel PPG signals corrupted by intense MAs. The algorithm mainly relies on EEMD-based clean PPG-mode extraction, adaptive filter-based denoising using three-axis accelerometer data, and some heuristic approaches for decision making. The salient feature of the proposed algorithm is that, unlike other recently reported well-known algorithms, it does not require an initial resting phase. Moreover, because of the use of two channel PPG signals, the algorithm has been found to be less susceptible to offtrack and runaway errors and, therefore, more robust and stable. The performance results using 14 PPG datasets of different subjects collected at various speed settings and motion activities have demonstrated superiority over other recent methods in terms of several figures of merit (e.g., average absolute error, error SD, limit of analysis, Pearson correlation coefficient etc.). In future works, we would like to investigate the algorithm's robustness in other kinds of athletic, gymnastic, aeronautical, military, and veterinary activities.

## REFERENCES

- [1] A. Kamal *et al.*, "Skin photoplethysmography review," *Comput. Methods Programs Biomed.*, vol. 28, no. 4, pp. 257–269, 1989.
- [2] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, no. 3, pp. R1–R39, 2007.
- [3] B. S. Kim and S. K. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 3, pp. 566–568, Mar. 2006.
- [4] 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.
- [5] R. Krishnan *et al.*, "Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 8, pp. 1867–1876, Aug. 2010.
- [6] M. Ram *et al.*, "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.

- [7] X. Sun *et al.*, "Robust heart beat detection from photoplethysmography interlaced with motion artifacts based on empirical mode decomposition," in *Proc. IEEE Int. Conf. Biomed. Health Informat.*, 2012, pp. 775–778.
- [8] F. Peng *et al.*, "Motion artifact removal from photoplethysmographic signals by combining temporally constrained independent component analysis and adaptive filter," *Biomed. Eng. Online*, vol. 13:50, no. 1, 2014.
- [9] K. Lam *et al.*, "Undergraduate students compete in the IEEE Signal Processing Cup: Part 1," *IEEE Signal Process. Mag.*, vol. 32, no. 4, pp. 123–125, Jul. 2015.
- [10] B. Lee *et al.*, "Improved elimination of motion artifacts from a photoplethysmographic signal using a Kalman smoother with simultaneous accelerometry," *Physiol. Meas.*, vol. 31, no. 12, pp. 1585–1603, 2010.
- [11] R. Yousefi *et al.*, "A motion-tolerant adaptive algorithm for wearable photoplethysmographic biosensors," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 2, pp. 670–681, Mar. 2014.
- [12] S. M. López *et al.*, "Heuristic algorithm for photoplethysmographic heart rate tracking during maximal exercise test," *J. Med. Biol. Eng.*, vol. 32, no. 3, pp. 181–188, 2012.
- [13] Z. Zhang *et al.*, "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.
- [14] 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.
- [15] K. T. Sweeney *et al.*, "The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 97–105, Jan. 2013.
- [16] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, 2009.
- [17] S. S. Haykin, *Adaptive Filter Theory*. Pearson Education, Delhi, India, 2008, ch. 9.
- [18] M. A. Arafat *et al.*, "A simple time domain algorithm for the detection of ventricular fibrillation in electrocardiogram," *Signal, Image Video Process.*, vol. 5, no. 1, pp. 1–10, 2011.
- [19] 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.



**Emroz Khan** was born in Bangladesh in 1992. He received the B.Sc. degree in electrical and electronic engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2015.

His research interest includes digital signal processing, electron microscopy, graphene and quantum information.

Mr. Khan participated in the IEEE Signal Processing Cup, Florence, Italy, in 2014, and his team secured the third place. He is currently a Student Member of

the IEEE Signal Processing Society and the IEEE Electron Devices Society. Since 2010, he has been associated with the academic team of Bangladesh Physics Olympiad.



**Forsad Al Hossain** was born in Bangladesh in 1992. He received the B.Sc. degree in electrical and electronic engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2015.

His research interest includes digital signal processing, electron microscopy and algorithm development.

Mr. Hossain was a participant at the IEEE Signal Processing Cup 2014, Florence, Italy, where his team achieved the first position. He is currently a Student Member of the IEEE Signal Processing Society. He was an Active Competitive Programming Contestant at the university.



**Shiekh Zia Uddin** was born in Bangladesh in 1992. He received the B.Sc. degree in electrical and electronic engineering from the Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 2015.

He is currently with the Nanophotonics Research Group, BUET. His research interest includes digital signal processing algorithms, inverse problems, plasmonic biosensors and SMD microscopy.

Mr. Uddin's team received the third place at the IEEE Signal Processing Cup, Florence, Italy, in 2014.



**S. Kaisar Alam** was born in Bangladesh. He received the Bachelor of Technology(Hons.) degree in electronics and electrical communication engineering from the Indian Institute of Technology Kharagpur, Kharagpur, India, in 1986, and the M.S. and Ph.D. degrees in electrical engineering from the University of Rochester, Rochester, NY, USA, in 1991 and 1996, respectively.

From 1986 to 1989, he was a Lecturer with the Department of Electrical and Electronic Engineering, Bangladesh Institute of Technology, Rajshahi, Bangladesh. While he was with the University of Rochester, his work on improving the ultrasonic estimation of blood velocity led to a patented technique for ultrasonic blood flow imaging. As a Postdoctoral Fellow with the University of Texas Health Science Center at Houston from 1995 to 1998, he worked on elastography, a new ultrasonic imaging modality to image elastic properties of tissue. He was a Principal Investigator with Riverside Research, New York, USA, for more than 15 years, where he worked on a variety of research topics in biomedical imaging. He is currently the Chairman and the Chief Research Officer at Imprlabs Pte Ltd., Singapore, an upcoming tech startup in Singapore. Since 2010, he has been a Visiting Faculty at the Department of Electrical and Electronic Engineering, Islamic University of Technology, Gazipur, Bangladesh. Since 2013, he has been a Visiting Research Professor at the Center for Computational Biomedicine Imaging and Modeling, Rutgers University, the State University of New Jersey, New Brunswick, NJ, USA. He has written more than 35 papers in international journals and holds several patents. He is the Coauthor of the textbook on Computational Health Informatics, scheduled to be published in 2016. His research interests include diagnostic and therapeutic applications of ultrasound and optics, and signal/image processing with applications to medical imaging.

Dr. Alam received the prestigious Fulbright Scholar Award in 2011 and 2012, respectively. He is a Member of the Sigma Xi, the Acoustical Society of America, and the Society of Photographic Instrumentation Engineers, and a Senior Member of the American Institute of Ultrasound in Medicine. He was with the AIUM Technical Standards Committee and the RSNA QIBA US SWS Technical Committee and is currently a Member of the RSNA QIBA Ultrasound Coordinating Committee. He is an Associate Editor of the *Ultrasonics and Ultrasonic Imaging* and a Member of the Editorial Board of the *Journal of Medical Engineering*.



**Md. Kamrul Hasan** received the B.Sc. and M. Sc. degrees in electrical and electronic engineering from the Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 1989 and 1991, respectively, and the M.Eng. and Ph.D. degrees in information and computer sciences from Chiba University, Chiba, Japan, in 1995 and 1997, respectively.

In 1989, he joined BUET as a Lecturer with the Department of Electrical and Electronic Engineering, where he is currently working as a Professor. He was a Postdoctoral Fellow and a Research Associate at Chiba University and Imperial College, London, U.K., respectively. He was a short-term invited Research Fellow at the University of Tokyo, Tokyo, Japan, and a Professor of International Scholars of Kyung Hee University, Seoul, Korea. He has authored or coauthored more than 100 scientific publications. His current research interests include digital signal processing, adaptive filtering, speech and image processing, and medical imaging.

Dr. Hasan is currently an Associate Editor for the IEEE ACCESS.

# A Robust Heart Rate Monitoring Scheme Using Photoplethysmographic Signals Corrupted by Intense Motion Artifacts

Emroz Khan, Forsad Al Hossain, Shiekh Zia Uddin, S. Kaisar Alam,  
and Md. Kamrul Hasan\*

**Abstract— Goal:** Although photoplethysmographic (PPG) signals can monitor heart rate (HR) quite conveniently in hospital environments, trying to incorporate them during fitness programs poses a great challenge, since in these cases, the signals are heavily corrupted by motion artifacts. **Methods:** In this paper, we present a novel signal processing framework which utilizes two channel PPG signals and estimates HR in two stages. The first stage eliminates any chances of a runaway error by resorting to an absolute criterion condition based on ensemble empirical mode decomposition. This stage enables the algorithm to depend very little on the previously estimated HR values and to discard the need of an initial resting phase. The second stage, on the other hand, increases the algorithm's robustness against offtrack errors by using recursive least squares filters complemented with an additional novel technique, namely time-domain extraction. **Results:** Using this framework, an average absolute error of 1.02 beat per minute (BPM) and standard deviation of 1.79 BPM are recorded for 12 subjects performing a run with peak velocities reaching as high as 15 km/h. **Conclusion:** The performance of this algorithm is found to be better than the other recently reported algorithms in this field such as TROIKA and JOSS. **Significance:** This method is expected to greatly facilitate the presently available wearable gadgets in HR computation during various physical activities.

**Index Terms**—Ensemble empirical mode decomposition (EEMD), photoplethysmography (PPG), recursive least squares (RLS) filter, wearable biomedical computing.

## I. INTRODUCTION

**P**HOTOPLETHYSMOGRAPHIC (PPG) signals [1], [2] are becoming a popular means of monitoring heart rate (HR) because of their wearable implementation compared to the conventional electrocardiography (ECG) technology. These signals are obtained through pulse oximeters which are embedded in a small wearable device to be put on at earlobes or fingertips (transmission type) or at wrists (reflection type). The pulse oximeters cast light on the wearer's skin through light-emitting diodes (LED) and obtain the transmitted or reflected light whose intensity depends on the amount of blood present in the arteries

Manuscript received March 4, 2015; revised July 8, 2015; accepted August 1, 2015. Date of publication; date of current version. Asterisk indicates corresponding author.

E. Khan, F. A. Hossain, and S. Z. Uddin are with the Bangladesh University of Engineering and Technology.

S. K. Alam is with Improlabs Pte, Ltd.

\*M. K. Hasan is with the Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka 1000, Bangladesh (e-mail: khasan@eee.buet.ac.bd).

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/TBME.2015.2466075

under the skin. Since the amount of blood varies with the cardiac cycle, the light intensity also varies with the cardiac rhythm and, thus, can be used to extract HR information.

Initially after its emergence, PPG signals were used to monitor patients' HRs 24 h in hospital beds, mainly using the fingertip-type devices. Even in these situations, slight movements of the patients' fingers would cause significant distortion in the signals and made the HR estimation difficult. Several attempts have been made to address this problem [3]–[8]. Ram *et al.* [6] used adaptive noise cancellation technique to remove the motion artifact (MA), though it suffered from high sensitivity to the choice of the reference signal. Sun *et al.* [7] performed artifact reduction using empirical mode decomposition (EMD), but considered all the modes with soft thresholding, instead of extracting only the correct PPG modes. Peng *et al.* [8] considered constrained-independent component analysis together with least mean square (LMS) adaptive filters to recover HR information. Though this algorithm and the rest others perform well for limited motion tolerance (e.g., horizontal and vertical movements of the finger, finger bending etc., with the subject sitting still), trying to apply them for significant MA (e.g., when the subject is in motion) does not yield satisfactory results.

Apart from clinical scenarios, many exercisers, sportsman, and elderly people require real-time mobile surveillance on their HR, so as to restrain them from excessive physical exertion and thereby, to prevent any chances of heart seizures. Following this line, many smart wearable gadgets have already been merchandised; however, these still require further robustness in their performance. Motivated by this future aspect of PPG signals, the IEEE Signal Processing Society has also arranged a research contest to invite attention to this problem [9].

To aid the signal processing task when PPG signals are being used during physical exercise, often three-axis accelerometry data are recorded in an attempt to model the MA [10]. Yousefi *et al.* [11] first attempted to bring the tissue effects under consideration and suggested a two-stage normalized-LMS adaptive noise canceller to remove MA. Though the method gives appreciable improvements over other previously reported algorithms, it requires pulse oximeters with different wavelengths. López *et al.* [12] on the other hand first pointed out the importance of heuristic approach to this problem and made interesting observations in the trend of PPG signals under intense running exercise. Using these observations, an algorithm was devised which yielded promising results emphasizing that previously estimated HR values can play a significant role in HR

estimation. But the algorithm does not work well in other types of exercises (e.g., boxing); in fact, even during rest, the algorithm faces difficulty in estimating HR.

Very recently, Zhang *et al.* [13] proposed a three-stage general framework and obtained benchmark results for substantial MA. But it still suffers from robustness to initial estimates. As a result, the subjects are required to stay still initially so that the initial estimates are correct and the tracker can perform accurately. As a sequel to this study, Zhang proposed another algorithm in [14] which though manages to further improve the previous results by a more vigorous signal sparsification technique, that is still heavily dependent on the tracking mechanism. In fact in his new algorithm, the HR monitoring process is not initialized at all if the MA is significant at the beginning. This was to make sure that HR estimates are on the right track from the start since the algorithm is not immune to initial wrong estimates.

In this paper, we consider the aforementioned shortcomings of the previous methods and propose a novel technique which works essentially in two stages. In the first stage, it is ensured that the algorithm is not suffering from a *runaway* error; that is, if for some reason the algorithm loses track of the true HR, it can still manage to get back on track no matter how far the ground truth is from the previous estimates. To accomplish this, we applied a noise-assisted ensemble empirical mode decomposition (EEMD) as a signal denoising technique together with an *absolute criterion* (AC) which dictates the algorithm to overthrow the dependence on previous estimates. Consequently, this stage also makes no need of an initial resting boot phase. The second stage works when the estimates are assumed to be already on the right track and deals with close proximity of MA and PPG peaks and spectral shadowing, so that the algorithm does not choose wrong estimates and get *offtrack*. To do so, we use a recursive least squares (RLS) adaptive filter complemented with an additional novel technique, namely *time-domain extraction*. Using this two-stage method, HR computation is performed on both single-channel and double-channel PPG signals on a number of sample datasets, and its performance is compared with the previously reported algorithms in the literature such as [13], [14].

This paper is organized as follows: Section II discusses the various cases where HR estimation is quite difficult, and Section III suggests ways to come around them. Section IV puts all the ideas together and establishes a well-structured algorithm. Section V illustrates the experimental results and compares with other benchmark algorithms in this field. Finally, Section VI concludes this paper.

## II. PROBLEM INSIGHTS

HR monitoring from motion-corrupted PPG signals is a challenging problem for a number of reasons. Some of them are listed below:

- 1) During intense hand movement (running, for example), the space between the wrist and the oximeter changes, and as a result, the measured intensity of the PPG signal varies. This variation is often correlated with the frequency of the hand motions. To illustrate this, Fig. 1(a) shows a segment

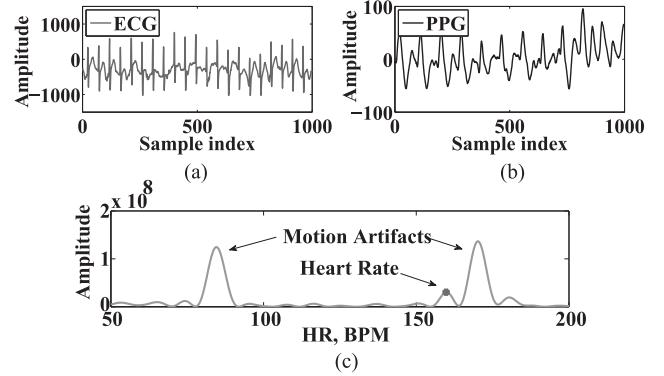


Fig. 1. Segment of the (a) ECG signal and the (b) corresponding PPG signal in time domain. (c) Periodogram of the PPG signal shows that there are large spurious MA peaks near the HR peak.

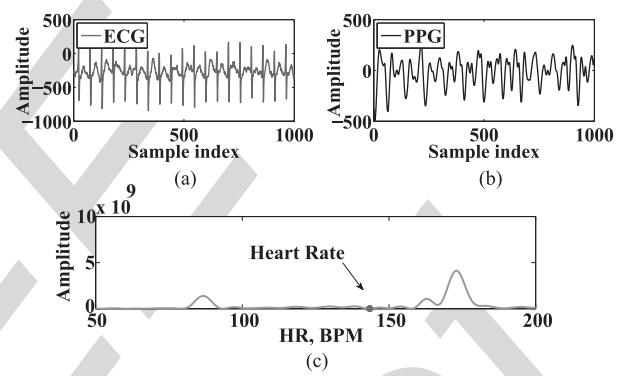


Fig. 2. Segment of the (a) ECG signal and the (b) corresponding PPG signal in time domain. (c) Periodogram of the PPG signal shows that there is no peak corresponding to the HR.

of the ECG signal, whereas (b) shows the corresponding PPG signal in time domain. As given by the periodogram of the PPG signal [see Fig. 1(c)], we can see that there are large spurious MA peaks near the HR peak.

- 2) Sometimes the sensors might be so far apart from the skin due to exhaustive motion that the original peak corresponding to the HR might be totally absent in the spectrum. Fig. 2(a) and (b) shows such ECG and PPG segments, respectively, together with the periodogram of the PPG in (c). This situation might persist for several seconds and poses a great difficulty in HR estimation.
- 3) Sometimes the peak corresponding to MA might be so close to the HR peak that they may get indistinguishable for a given resolution. Fig. 3(a) and (b) gives an example of such ECG–PPG pair along with the periodogram of PPG in (c). Increasing the data length will not circumvent the problem because that would require the HR process to be stationary for an extended amount of period, which seldom is the case. To make things worse, often the side-lobe of a MA peak is so strong that it masks the original HR peak.
- 4) As pointed out in [11], intense motion like boxing, jumping, etc., may cause so much acceleration in the blood at the arteries that even correct intensity might not correspond to actual HR.

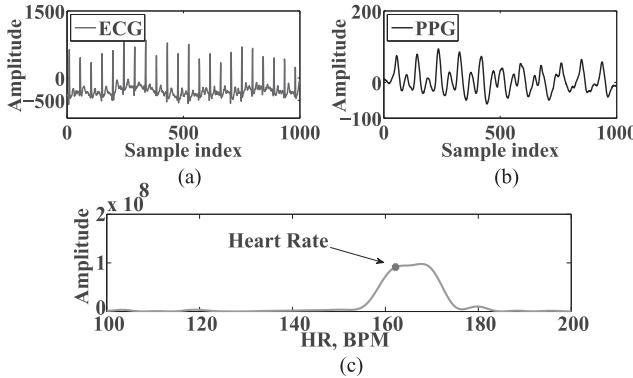


Fig. 3. Segment of the (a) ECG signal and the (b) corresponding PPG signal in time domain. (c) Periodogram of the PPG signal shows that the MA peak and the HR peak are almost indistinguishable from each other.

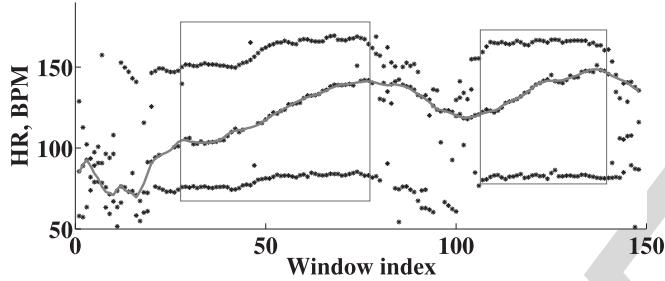


Fig. 4. Plot (in blue line) of the ground truth HRs as a subject starts running with an initial rest. After a while, the subject pauses to take a brief rest and starts running again at a different speed. The overlay dots (in black) are the strongest PPG peaks found in each time window. Clearly, the trend mentioned in [12] is observed, while the subject runs steadily (inside the red rectangles), but the pattern is broken elsewhere.

- 5) HR is a periodic process (if the time window is not too long) and, therefore, can have higher order harmonics. But the presence of harmonics is not sufficient enough to differentiate between HR and MA peaks (as was done in [13]), since MA can also be aperiodic or periodic. In rest, the random movements of the hands cause the MA to be aperiodic, whereas the free swinging of the hands in a running exercise renders the MA periodic.
- 6) The observation made in [12] shows that when a subject performs a run, then for each time window the HR is usually among the top three dominant peaks with the other two coming from MA (similar to the example given in Fig. 1). In fact, if we plot the overlays of the top three dominant peak locations over all the time windows (as shown in Fig. 4), then a certain pattern is observed. In this pattern, the highest and the lowest frequency peak locations remain almost constant and correspond to MA, whereas the middle frequency peak location shows a rising or falling trend corresponding to the actual HR. Though this holds to a good approximation when the subject runs steadily (inside the red rectangles in Fig. 4), the pattern is destroyed at rest or at change in running speed (similar to the cases illustrated in Figs. 2 and 3). In fact, estimating the HR during running is easier sometimes than the rest

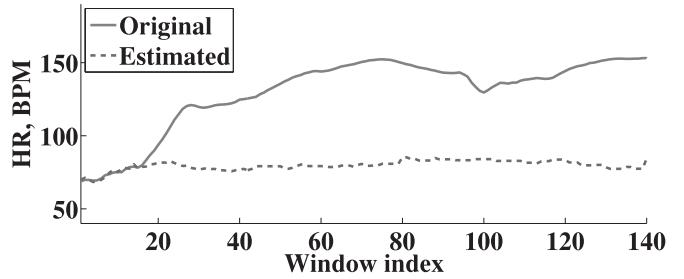


Fig. 5. Example of a runaway situation which arises from too much dependence upon the tracking mechanism. When a MA peak is close to the HR peak and is mistaken to be a HR estimate, the tracking mechanism would keep following the MA peaks in successive windows, even if the original HR peak gets stronger later in time.

position, since in the latter case, the MA peaks are large in number and are scattered all over the spectrum.

- 7) Since actual HRs do not show a sudden discontinuity, this provides with an important *a priori* information which allows for a provision of tracking mechanism. This gives good results most of the time, but for intense MA, this can, in fact, mislead the HR estimates go astray. For example, sometimes when a MA peak is close to the HR peak and is strong in magnitude and, thus, is mistaken to be a HR estimate, the tracking mechanism would follow the MA peaks in successive time windows, and would not allow for a come around to the correct track, even if the original HR peak gets stronger later in time. Thus, stability of the algorithm is sacrificed as too much confidence is placed upon the tracking mechanism resulting in a large runaway error as shown in Fig. 5.

Though the difficulties jotted above can certainly be reduced by applying various signal processing techniques, there are always chances for them to arise in certain time windows. In the next section, we discuss ways to minimize their occurrences and countermeasures to overcome them should they arise anyway.

### III. MATERIALS AND METHODS

In this section, we illustrate the key features to be used in our algorithm and expound on their effectiveness and functionality.

#### A. Data Acquisition and Modeling

For benchmarking purpose, we run our algorithm on the same datasets used in [13] and [14] where 12 healthy male subjects (age ranging from 18 to 35) ran on a treadmill with speeds reaching upto 15 km/h. Among the 12 datasets, we find datasets 1, 2, 4, and 10 the most corrupted ones and take them as our training data. The remaining datasets were used as testing purpose once the algorithm along with its parameters was derived from the training phase. Moreover, we run the same algorithm on a new dataset where the subject, instead of running, performed intensive forearm and upper arm movements (e.g., boxing). We also run the algorithm for a subject (age 58, female) with abnormal heart rhythm and blood pressure. For all datasets, the PPG signals were recorded from the wrist by two pulse oximeters with green LEDs (wavelength: 515 nm). Their distance

(from center to center) was 2 cm. The acceleration signal was also recorded from the wrist by a triaxis accelerometer. Both the pulse oximeters and the accelerometer were embedded in a wristband, which was comfortably worn. In order to calculate the ground truth HR for evaluating the algorithm's performance, ECG signal was recorded simultaneously from the chest using wet ECG sensors.

The proposed algorithm can estimate HR from either single-channel or two-channel PPG signals. Let us describe the algorithm for the two channel case and let each channel come with successive time windows of length  $M$  with some overlap  $L$ . Then, we can model the motion-corrupted raw PPG segments  $y_{i,\text{raw}}(n)$  ( $i = 1, 2$ ) by

$$y_{i,\text{raw}}(n) = d_i(n) + v_i(n), \quad n = 0, 1, \dots, M - 1$$

where  $d_i(n)$  is the desired clean PPG signal and  $v_i(n)$  is the noise introduced by motion. Let the three-axis raw acceleration data for the same time window be denoted by  $a_{x,\text{raw}}(n)$ ,  $a_{y,\text{raw}}(n)$ , and  $a_{z,\text{raw}}(n)$ . Since human HRs usually reside between 40 and 200 beat per minute (BPM), we bandpass filter the PPG and the acceleration segments in this frequency range to obtain  $y_i(n)$ ,  $a_x(n)$ ,  $a_y(n)$ , and  $a_z(n)$ , respectively. Also, let us assume that the sampling frequency of all the signals is  $F_s$ . Like Zhang *et al.* [13], [14], we also consider 8-s-long data windows with 75% overlap (this implies  $M = 1000$  and  $L = 750$  for  $F_s = 125$  Hz).

### B. Ensemble Empirical Mode Decomposition

In this paper, we use EEMD as a tool for signal decomposition to separate the clean PPG signal  $d_i(n)$  from the MA  $v_i(n)$ . It is reported that EEMD works very well for cleansing biomedical signals for postprocessing [15], [16]. EMD is a nonlinear signal processing technique that decomposes a given signal into a number of intrinsic mode functions (IMF). The IMFs satisfy the following two conditions: 1) the number of extrema and the number of zero crossings over the whole signal length must either be the same or differ at most by 1, and 2) the mean value of the envelope defined by the maxima and the envelope defined by the minima must be zero at all points in the signal.

For a given bandpass-filtered PPG signal  $y(n)$  (let us drop the channel index  $i$  in this section), its EMD is implemented using a sifting process that first finds all the local maxima and minima, and then connects them using cubic splines to create two corresponding envelopes: One upper envelope for the maxima and one lower envelope for the minima. The average of the two envelopes  $m(n)$  is then subtracted from the original signal  $y(n)$  to obtain a new signal

$$h(n) = y(n) - m(n).$$

This new signal is then put in place of  $y(n)$  and the above process is repeated until  $h(n)$  satisfies the two conditions of an IMF. When this occurs,  $h(n)$  becomes the first IMF  $c_1(n)$ .

Another round of sifting process is repeated on the residual signal  $y_1(n) = y(n) - c_1(n)$  to sift the next IMF. The whole process is halted when the residual signal  $y_K(n)$  after the extraction of the  $K$ th IMF becomes a monotonic function. The

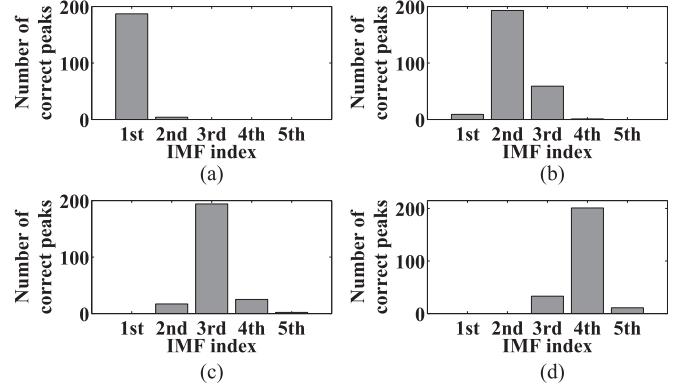


Fig. 6. Bar plot of the number of time windows of dataset 4 for which the maximum peak of the periodogram of an IMF gives the correct HR at sampling frequencies (a)  $F_s = 25$  Hz, (b) 125 Hz, (c) 250 Hz, and (d) 500 Hz. From this plot, it is evident that the dominant IMF for the clean PPG mode changes with  $F_s$ .

IMFs have the property that they, when added together with the final residual signal, can reconstruct the original signal

$$y(n) = \sum_{j=1}^K c_j(n) + y_K(n).$$

The EMD algorithm is, however, quite sensitive to mode mixing where an IMF includes uncontrolled oscillations or transient spectral content. To address mode mixing, an extension of the EMD algorithm, known as EEMD is proposed in [16]. Here, an ensemble of  $N_E$  signals is created from the given signal  $y(n)$  by adding white Gaussian noise  $w_p(n)$  of the same variance ( $p = 1, \dots, N_E$ ). That is, the ensemble contains

$$\tilde{y}_p(n) = \sum_{j=1}^K c_{pj}(n) + y_{pK}(n), \quad p = 1, \dots, N_E$$

$y_{pK}(n)$  being the residual function for  $\tilde{y}_p(n)$ . Finally, the optimum choice of the IMF is taken as the ensemble average

$$\bar{c}_j(n) = \frac{1}{N_E} \sum_{p=1}^{N_E} c_{pj}(n), \quad j = 1, \dots, K.$$

In the context of PPG, we find that if the PPG signal contains HR information, then it is usually in a specific IMF for a particular sampling frequency. Fig. 6 illustrates this point where we bar plot the number of time windows of dataset 4 for which the maximum peak of the periodogram of a certain IMF coincides with the ground truth HR. From this figure, we see that the dominant PPG mode is the first IMF for sampling frequency  $F_s = 25$  Hz [see Fig. 6(a)], the second IMF for  $F_s = 125$  Hz [see Fig. 6(b)], the third IMF for  $F_s = 250$  Hz [see Fig. 6(c)], and the fourth IMF for  $F_s = 500$  Hz [see Fig. 6(d)].

Fig. 7 illustrates the use of EEMD to extract HR information from ECG and PPG signals (here,  $F_s = 125$  Hz is used). Fig. 7(a) shows an ECG segment along with its periodogram in Fig. 7(b).

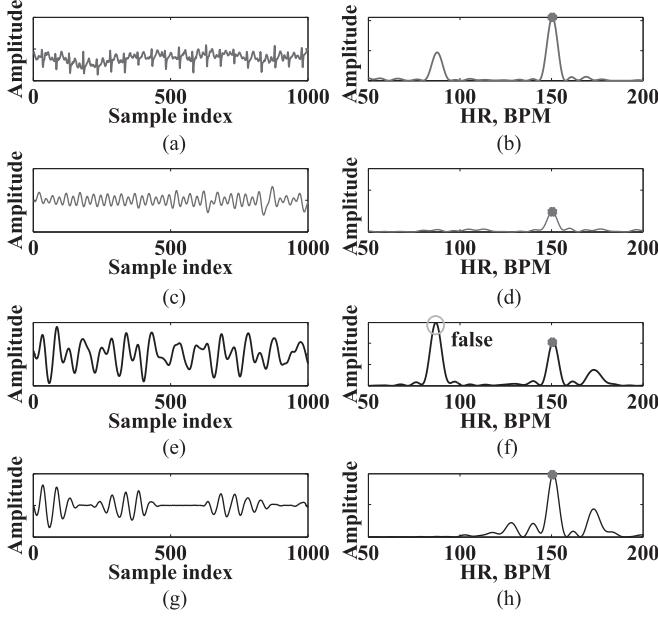


Fig. 7. Illustration of the effectiveness of using EEMD to extract the HR information from the noisy PPG signal. Here, both the ECG and the noise-corrupted PPG signal segments contain the HR information in their second IMFs (sampling frequency  $F_s = 125$  Hz). (a) Original ECG. (b) Periodogram of original ECG. (c) Second IMF of ECG. (d) Periodogram of second IMF of ECG. (e) Original PPG. (f) Periodogram of original PPG. (g) Second IMF of PPG. (h) Periodogram of second IMF of PPG.

If EEMD is performed on this segment, then we get the second IMF as shown in Fig. 7(c). Clearly, its maximum peak coincides with the ground truth HR shown as a red dot in the figure. Now, if we look at the PPG signal for the same time window, we see that it is corrupted by MA in both time [see Fig. 7(e)] and frequency domain [see Fig. 7(f)]. Its maximum peak (yellow circled in the figure) does not give the correct HR since it comes from MA. But if we perform EEMD on it to extract its second IMF [see Fig. 7(g)], we find that its maximum peak now coincides with the HR.

### C. Absolute Criterion

As discussed in Section II, stability and tracking mechanism work in opposite directions. To make sure that the proposed algorithm is stable, it should have a definite criterion which can allow to overthrow the previous track of HR estimates and jump to a distant estimate. This way, if for some reason, the HR estimates get derailed, it can still make it back to the original track and, thus, avoid any chance of runaway error. Also, no initialization would be required in this case, since no matter where it started from, it can still have the opportunity to catch up with the track should the criterion is fulfilled. We propose this AC to be as follows:

*If the PPG signal, after being suitably cleansed, shows a reliable peak, then this peak is assumed to correspond to the original HR.*

The existence of such an AC is vouched by the fact that a given signal no matter how much contaminated by MA, after suitably

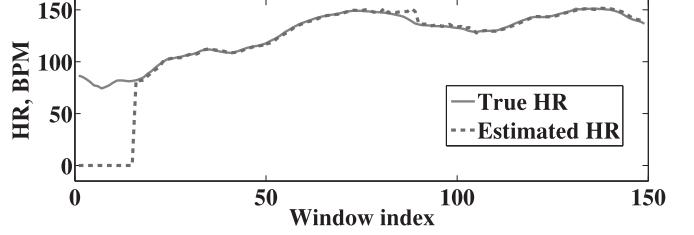


Fig. 8. Illustration of the effectiveness of using an AC which enables the algorithm in catching up with the original HR, and therefore, discards the requirement of an initial resting boot phase.

cleansed, is very unlikely to not give a strong reliable peak ever in any of its numerous time windows. In fact, our observation has shown that, even in statistical sense, strong correct peaks emerge quite frequently even when the PPG signals are heavily corrupted by MA, and these peaks are even more frequent if the cleansing process is effective.

This way, the reliable peaks can stand as checkpoints for the estimated HR track. Fig. 8 illustrates the effectiveness of this criterion which shows that even though we deliberately initialized the HR estimate to be 0 BPM, the algorithm after a while finds one correct reliable peak to meet the AC. So it takes the huge leap to move onto the correct track. In this paper, we use EEMD-based denoising technique to “suitably cleanse” the PPG signals for the AC condition, which is discussed at length in Section IV. However, the process of suitably cleansing the signal and the condition of identifying a peak as reliable may vary from algorithm to algorithm.

### D. RLS Filter

When the AC in our algorithm is not fulfilled, that is, when EEMD fails to provide IMFs with a reliable HR peak, we resort back to the raw PPG signals  $y_{i,\text{raw}}(n)$  and try to denoise their average  $\bar{y}_{\text{raw}}(n)$  using the raw acceleration data. Toward this end, we model the noise  $v(n)$  present in  $\bar{y}_{\text{raw}}(n)$  by

$$v(n) = v_x(n) + v_y(n) + v_z(n)$$

where  $v_\gamma(n)$  is the MA contribution correlated with the acceleration data  $a_{\gamma,\text{raw}}(n)$  ( $\gamma = x, y$ , and  $z$ ). We find that RLS adaptive FIR filters with length  $s$  work quite well to obtain the estimates  $\hat{v}_\gamma(n)$  for the noise components  $v_\gamma(n)$  from the reference signals  $a_{\gamma,\text{raw}}(n)$ .

Fig. 9 illustrates the whole process of noise cancellation. First, we estimate for  $v_x(n)$  from  $a_{x,\text{raw}}(n)$  by minimizing the least squares error

$$\xi(n) = \sum_{q=0}^n \lambda^{n-q} |e_x(q)|^2$$

where  $\lambda$  is the forgetting factor (we use  $\lambda = 1$ ) and  $e_x(n)$  is the error signal  $e_x(n) = \bar{y}_{\text{raw}}(n) - \hat{v}_x(n) = \bar{y}_{\text{raw}}(n) - a_{x,\text{raw}}(n) * w_x(n)$  with  $w_x(n)$  being the RLS filter coefficients. The coefficients are obtained adaptively by using a set of update equations given in [17]. Note that the error signal found from this stage  $e_x(n)$  is assumed to not contain MA contribution correlated

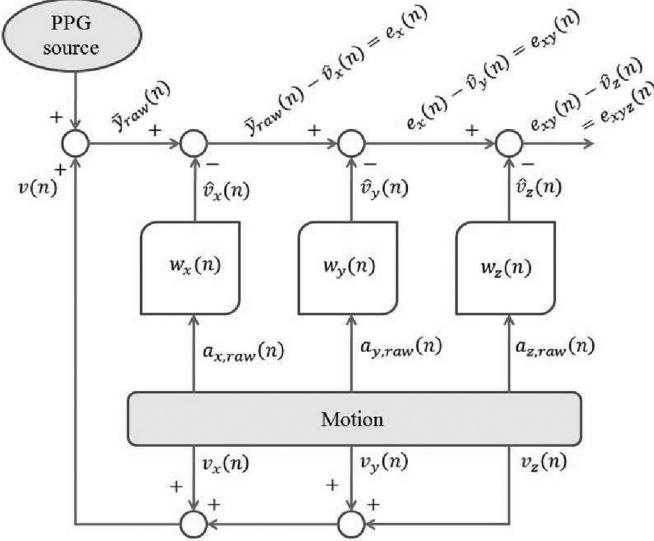


Fig. 9. RLS filter implementation to remove the different noise components correlated with the three-axis acceleration data.

with  $a_{x,\text{raw}}(n)$ . Next, another RLS filter  $w_y(n)$  is implemented to estimate for  $v_y(n)$  by treating  $a_{y,\text{raw}}(n)$  as the reference signal and  $e_x(n)$  as the desired signal. The error signal obtained in this stage  $e_{xy}(n) = \bar{y}_{\text{raw}}(n) - \hat{v}_x(n) - \hat{v}_y(n)$  now supposedly contains MA contribution coming from only  $a_{z,\text{raw}}(n)$ . In order to get rid of it too,  $a_{z,\text{raw}}(n)$  is now treated as the reference signal, while  $e_{xy}(n)$  as the desired signal to obtain another set of RLS filter coefficients  $w_z(n)$ . The final error signal  $e_{xyz}(n) = \bar{y}_{\text{raw}}(n) - \hat{v}_x(n) - \hat{v}_y(n) - \hat{v}_z(n)$  can be regarded as a denoised signal  $r_{\text{raw}}(n)$  which is assumed to have no correlation with the acceleration. This signal is then investigated further for HR peaks which we elaborate in detail in Section IV.

Fig. 10 illustrates the effectiveness of using RLS filters when EEMD fails to provide with a correct reliable peak. Fig. 10(a) shows an ECG segment ( $F_s = 125$  Hz), and the corresponding PPG segment is given in (b) along with its periodogram in (c). Note that the PPG periodogram is quite similar to the typical case illustrated in Fig. 1. The periodogram of the dominant IMF containing the clean PPG mode (second IMF for  $F_s = 125$  Hz) is shown in Fig. 10(d). Clearly, it fails to provide with correct HR since its maximum peak coincides with an MA peak as given by the periodogram of the acceleration data in Fig. 10(e). Finally, Fig. 10(f) shows the recovery of the true HR peak by the use of RLS filters as described above.

### E. Time-Domain Extraction

In our investigation of many sample datasets, we noted that for certain time windows, not all part of the signal segment is corrupted equally by MA. In fact, some part of the window might be clean enough that just cropping out that portion and applying the algorithm on it gives close estimates of HR. But, if the whole data window is taken for consideration, then MA is allowed to dominate the PPG spectrum and it gets difficult to detect the correct peak.

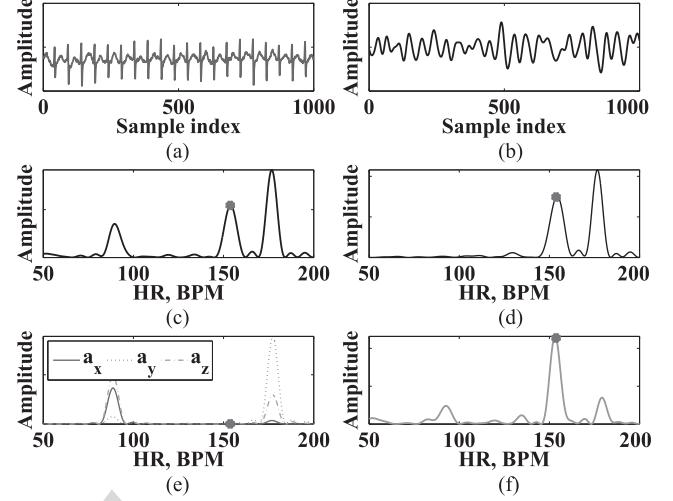


Fig. 10. Illustration of the effectiveness of using RLS filters when EEMD fails. Here, the clean PPG mode IMF (second IMF for  $F_s = 125$  Hz) does not contain the HR at its maximum peak location. Instead, it coincide with an MA peak as can be clearly seen by the periodograms of the accelerations. If RLS filters are applied using the acceleration signals as references, then maximum peak of the filter output coincides with the HR. (a) Original ECG. (b) Original PPG. (c) Periodogram of original PPG. (d) Periodogram of 2nd IMF of PPG. (e) Periodogram of acceleration. (f) Periodogram of PPG after RLS.

In order to decide on the location and the length of such window portions, the three-axis acceleration data can be used because they give a fair estimate of the amount of noise being injected. We found out that the net acceleration

$$a(n) = [a_x^2(n) + a_y^2(n) + a_z^2(n)]^{1/2}, \quad n = 0, 1, \dots, M-1$$

quite vividly marks out these portions: First, an interval  $I \subset \{0, 1, \dots, M-1\}$  is looked for so that

$$a(n) \leq \tau A \quad \forall n \in I$$

where  $A$  is the maximum value of  $a(n)$  in the whole time window and  $\tau$  is a suitable threshold usually kept around 0.3 [18]. Among all such intervals, if the largest one  $I_0$  is long enough to capture at least two heart beats (nearly 1 s long), then we crop the corresponding portion out of the original bandpass-filtered PPG signals  $y_i(n)$  by constructing

$$y_{i,\text{cropped}}(n) = y_i(I_0).$$

Though such small portion gives very smeared periodogram and its peak is not exactly the actual HR throughout the entire time window under consideration, it gives very close and robust estimates.

Fig. 11 illustrates the effectiveness of this idea. Fig. 11(a) shows an ECG signal segment (sampled at 125 Hz) for a particular time window along with its periodogram shown in (b). Fig. 11(c) shows the corresponding PPG signal, whereas (d) shows its periodogram. Note that the PPG periodogram is quite similar to the case illustrated in Fig. 2 in that there is no peak corresponding to HR at all. In fact, applying a high-resolution denoising technique like singular spectrum analysis [13], for example, fails to provide any peak at the HR as shown in the next Fig. 11(e). This is partly because of the wide spectrum of

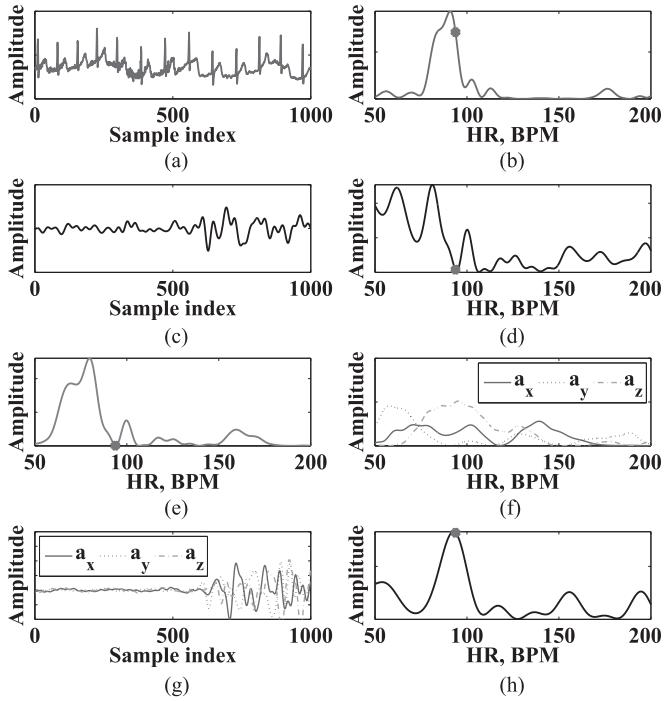


Fig. 11. Illustration of the effectiveness of time-domain extraction. (a) Original ECG. (b) Periodogram of ECG. (c) Original PPG. (d) Periodogram of PPG. (e) Periodogram of PPG after SSA. (f) Periodogram of acceleration. (g) Acceleration in time domain. (h) Periodogram of cropped PPG.

the MA, as shown by the periodograms of the three-axis acceleration data in Fig. 11(f). Now, if we observe the acceleration data in the time domain [see Fig. 11(g)], we immediately notice that for this particular time window we have a large portion of the signal where the MA is less intense compared to the rest of the window. Extracting that part of the PPG signal and just taking its periodogram yields a very strong peak corresponding to the true HR as shown in Fig. 11(h).

Aperiodic exercises, like boxing, can give large erroneous peaks if the entire window is considered since here the small amount of time where the hand plunges forward degrades the whole PPG spectrum. But if we extract only that portion of the time window where the hands are stowed to the chest between two successive punches, then quite correct estimates can be obtained. As mentioned earlier, estimating the HR during resting phase is often more difficult than in running phase. In these cases, this approach works quite well because sudden jolts of hands in rest phase can be filtered out by this approach.

#### IV. PROPOSED ALGORITHM

In this section, we attempt to encompass all the ideas illustrated above and give details of our proposed algorithm. The algorithm estimates HR through the following steps.

**Signal Workspace:** In this stage, we produce all the signals necessary to enter into our main algorithm. We are given two channels of raw PPG segments  $y_{i,\text{raw}}(n)$  ( $i = 1, 2$ ) and the corresponding raw acceleration segments  $a_{\gamma,\text{raw}}(n)$  ( $\gamma = x, y$  and  $z$ ) at a sampling frequency  $F_s$ . We apply the RLS filter block

(as illustrated in Fig. 9) on the average of  $y_{i,\text{raw}}(n)$  by using  $a_{\gamma,\text{raw}}(n)$  as the reference signals. In our experiments, we set the RLS filter length  $s$  around  $8 - 10$  for  $F_s = 25$  Hz, and it is increased for higher sampling frequencies. The partially denoised output  $r_{\text{raw}}(n)$  along with the PPG signals  $y_{i,\text{raw}}(n)$  and acceleration signals  $a_{\gamma,\text{raw}}(n)$  are then passed through an infinite impulse response bandpass filter with lower and upper passband cut off at 40 and 200 BPM, respectively (typical frequency range for human HR), and lower and upper stopband cut off at 35 and 205 BPM, respectively, and passband ripple 0.01, stopband attenuation 80 dB to obtain  $r(n)$ ,  $y_i(n)$ , and  $a_{\gamma}(n)$ , respectively.

**Initialization:** The main algorithm estimates HR by making decisions based on the previous HR estimate  $f_{\text{prev}}$ . But if the data window under consideration is the first one, there is no  $f_{\text{prev}}$  and we need to estimate it first. We do so as follows.

- 1)  $r(n)$  is inspected in Fourier domain and its dominant peaks' locations (peaks larger than 80% of the maximum peak) are arrayed together in a set  $S_{\text{rls}}$ . If  $|S_{\text{rls}}| = 1$  (here  $|\cdot|$  represents the cardinality of a set), which means  $S_{\text{rls}}$  contains the location of only the maximum peak, then we assign that to  $f_{\text{prev}}$ .
- 2) If  $|S_{\text{rls}}| \neq 1$ , then there are multiple strong peaks present. Therefore, we look for their second harmonics in  $S_{\text{rls}}$  and build a new set  $H_{\text{rls}} \subset S_{\text{rls}}$  so that if  $x \in H_{\text{rls}}$ , then there exists a  $y \in S_{\text{rls}}$  such that  $|2x - y| < 5$  BPM. If  $H_{\text{rls}}$  is not empty, then the strongest peak is assigned to  $f_{\text{prev}}$ .
- 3) If  $|H_{\text{rls}}| = 0$ , then we first construct a set  $S_{a,0.8}$  by taking all the dominant peaks (80% of the maximum peak) from the acceleration signals  $a_{\gamma}(n)$ , and then cross out all the peaks in  $S_{\text{rls}}$  which are close to the acceleration peaks stored in  $S_{a,0.8}$  within a 5 BPM range by forming  $S = S_{\text{rls}} \setminus^5 S_{a,0.8}$ . Here for notational convenience, we define  $A \setminus^{\delta} B = \{x \in A : |x - y| > \delta \text{ BPM } \forall y \in B\}$ . Next, if  $S$  is not empty, the strongest peak in  $S$  is assigned to  $f_{\text{prev}}$ .
- 4) If  $f_{\text{prev}}$  is still not assigned (which is a rare case), then we assign its value to the strongest peak's location in  $S_{\text{rls}}$ . As we shall see later, even if  $f_{\text{prev}}$  is erroneously initialized, our main algorithm is designed in such a way that it can correct for it later and come back on track.

With  $f_{\text{prev}}$  assigned, we now enter the main algorithm.

**Main Algorithm:** Here, a hierarchy of steps is executed to obtain a crude estimate  $f$  for the HR as described below.

- 1) **AC:** First EEMD is applied to  $y_i(n)$  with  $N_E \simeq 5$  and SNR level around 30 dB. Among the obtained IMFs  $\bar{c}_{i,j}(n)$  ( $i$  being the channel index and  $j$  being the IMF index), only one IMF is considered per each channel. This selection is based on the sampling frequency (e.g.,  $j = 1$  for  $F_s = 25$  Hz,  $j = 2$  for  $F_s = 125$  Hz,  $j = 3$  for  $F_s = 250$  Hz, and  $j = 4$  for  $F_s = 500$  Hz). Then, the two selected IMFs obtained from both channels are inspected in Fourier domain and their maximum peaks' locations are put in a set  $S_{\text{imf}}$ . Similarly, we construct another set  $S_{a,0.5}$  by taking all the dominant peaks (50% of the maximum peak, a low threshold to ensure capturing all MA peaks) from the acceleration signals  $a_{\gamma}(n)$ . Next,  $g$  groups are

constructed from the elements of  $S_{\text{imf}}$  such that they all contain peak locations in clusters of at most 2 BPM range and each of them has at least one peak that is not close to an acceleration peak. In other words, if  $G$  is such a group, then  $\max(G) - \min(G) \leq 2 \text{ BPM}$  and  $|G \setminus S_{a,0.5}| \neq 0$ . Groupwise averages are taken next, and from these  $g$  averages, the one closest to  $f_{\text{prev}}$ , say  $f_{\text{AC}}$  is considered. If  $|f_{\text{AC}} - f_{\text{prev}}| < \Delta_{\text{AC}}$  with  $\Delta_{\text{AC}}$  being a suitable range to be discussed shortly, then we are confident in assuming that this peak, which is present in one of the chosen IMFs and also does not lie close to any MA peak, is an intrinsic peak of the desired PPG signal. We further assume that the peak is reliable enough to meet the AC condition in this case and consider the peak as the crude estimate,  $f = f_{\text{AC}}$ . Each time the AC is not met, the range of  $\Delta_{\text{AC}}$  aforementioned is increased dynamically by an amount  $\Delta_d$  from its default value  $\Delta_0$ . That is, if  $f$  is not assigned any value from this step  $m$  times in a row, then we will have for the next time window  $\Delta_{\text{AC}} = \Delta_0 + m\Delta_d$ . However if assigned, then the range of  $\Delta_{\text{AC}}$  is reset to its default value  $\Delta_0$ . In our experiments, we set  $\Delta_0 \simeq 5 \text{ BPM}$  and  $\Delta_d \simeq 1 \text{ BPM}$ .

- 2) *Tracking from IMFs:* From this step onward, we employ different tracking mechanism. First, we construct  $S_{\text{imf}} \setminus S_{a,0.5}$ , and from this set, we take the peak location nearest to  $f_{\text{prev}}$ . If its distance from  $f_{\text{prev}}$  is within  $\Delta_{\text{imf}} \simeq 7 \text{ BPM}$ , we assign it to  $f$ , and since this peak comes from IMFs corresponding to clean PPG modes, we also reduce the range of  $\Delta_{\text{AC}}$  set for the next time window by  $\Delta_d$ .
- 3) *Tracking from time-domain extraction:* In this step, we check whether the signals  $y_i(n)$  are suitable for time-domain extraction. If so, we take the respective clean portions  $y_{i,\text{cropped}}(n)$  and find their maximum peaks' locations in Fourier domain. The peak location nearer to  $f_{\text{prev}}$  is considered and checked to see if its distance from  $f_{\text{prev}}$  is also within a range  $\Delta_t$  (around 7 – 12 BPM). If so, we assign it to  $f$ .
- 4) *Tracking from RLS signal:* Here, we consider  $r(n)$ , and similar to the initial step, we put its dominant peaks (80% of the maximum) in  $S_{\text{rls}}$ . We also construct  $S_{a,0.6}$  by taking the dominant peaks (60% of the maximum, a moderate threshold for tracking purpose) from  $a_\gamma(n)$ . Now, if the set  $S_{\text{rls}} \setminus S_{a,0.6}$  contains only one peak  $f_{\text{rls}}$  and if  $|f_{\text{rls}} - f_{\text{prev}}| < 25 \text{ BPM}$ , then we assign  $f = f_{\text{rls}}$ . However, if this is not the case, then the strongest peak in  $S_{\text{rls}}$  is looked for such that it lies close to  $f_{\text{prev}}$  within a range  $\Delta_{\text{rls}}$  (7 – 12 BPM). If found, we assign it to  $f$ .
- 5) *Tracking from the original signal:* If the above steps fail to provide with the crude estimate  $f$ , then we consider all the peak locations attainable from the periodograms of  $y_i(n)$  and array them together in a set  $S_{\text{org}}$ . We also construct another set  $S_{a,0}$  by taking all the peaks (no threshold in this case) from  $a_\gamma(n)$ . Next, we take the peak  $f_{\text{org}} \in S_{\text{org}}$  which is closest to  $f_{\text{prev}}$  and also  $|f_{\text{org}} - f_{\text{prev}}| \leq 5 \text{ BPM}$ . Then, if  $f_{\text{org}}$  does not lie close to any acceleration peak stored in  $S_{a,0}$  within a 3 BPM tolerance, that is, if  $f_{\text{org}} \in S_{\text{org}} \setminus S_{a,0}$ , then we assign  $f = f_{\text{org}}$ . If this is not the case,

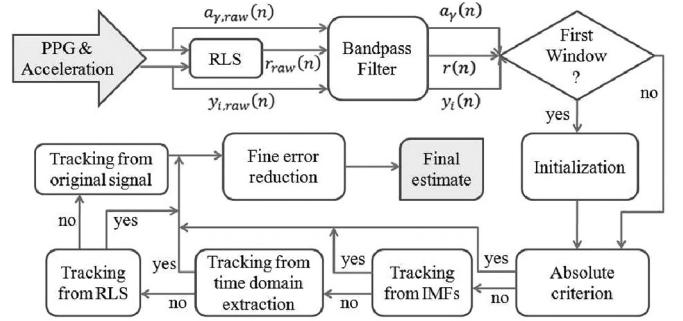


Fig. 12. Flowchart of the proposed algorithm.

then we reach the conclusion that the given PPG segments are highly corrupted, and in both the original segments and the denoised segments, there is no discernible peak. Therefore, we discard the segment from our consideration by assigning  $f = f_{\text{prev}}$ .

*Fine Error Reduction:* This is the final step of our algorithm where we attempt to obtain a fine estimate  $f_{\text{est}}$  from the crude estimate  $f$  by working with the raw PPG data. To do so, we construct a set  $S_{\text{raw}}$  by taking all the peak locations attainable from the periodograms of  $y_{i,\text{raw}}$ . Then, we look for the peak  $f_{\text{raw}} \in S_{\text{raw}}$  which is closest to  $f$  and also  $|f_{\text{raw}} - f| \leq \Delta_{\text{fine}}$  with  $\Delta_{\text{fine}}$  around 3 – 4 BPM. If such a peak is found, then we assign our fine estimate  $f_{\text{est}} = f_{\text{raw}}$ . Otherwise, we consider our crude estimate as the final estimate  $f_{\text{est}} = f$ . Also, in this step, we assign  $f_{\text{prev}} = f_{\text{est}}$  for the next time window.

Therefore, the algorithm works in several key steps. In the first step of the main algorithm, robustness is ensured with the EEMD-based signal cleansing followed by the application of AC. Each time the AC is not met, the range of  $\Delta_{\text{AC}}$  is increased with no considerable limit and chances of meeting the AC in the next window increases, resulting in the algorithm's ability to catch up with the correct track and avoid a runaway scenario. Therefore, this step enables the algorithm to become robust to initial wrong estimates, and thereby, helps the algorithm to not require an initial resting phase. If this step is bypassed, then in the following series of steps tracking is emphasized. Here, previous HR estimate is given full priority and all nearby peaks attainable from IMFs, RLS output, time-domain extracts, and original bandpass-filtered signals are taken into consideration. In the last stage, the algorithm reduces the fine error by working with the raw PPG segments to better coincide with the ground truth. For ease of understanding, a flowchart of the algorithm is given in Fig. 12. Note that though the above algorithm is described explicitly for two channel PPG signals, it can be generalized for any number of channels. For single-channel implementation, simply letting  $y_{1,\text{raw}} = y_{2,\text{raw}}$  gives the corresponding framework.

## V. EXPERIMENTAL RESULTS

### A. Performance Evaluation

Table I lists the average absolute error for each of the 12 datasets (sampled at  $F_s = 25 \text{ Hz}$ ) for both the double-channel

TABLE I  
PERFORMANCE COMPARISON OF DIFFERENT ALGORITHMS USING AVERAGE ABSOLUTE ERRORS FOR THE 12 TRUNCATED DATASETS USED IN [14] FOR  $F_s = 25$  Hz

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	Subj 11	Subj 12	Mean $\pm$ SD
Proposed- <i>d</i>	1.70	0.84	0.56	1.15	0.77	1.06	0.63	0.53	0.52	2.56	1.05	0.91	1.02 $\pm$ 1.79
Proposed- <i>s</i>	1.77	1.94	0.73	1.19	0.51	1.09	0.52	0.43	0.36	3.43	0.89	0.98	1.15 $\pm$ 2.37
JOSS [14]	1.33	1.75	1.47	1.48	0.69	1.32	0.71	0.56	0.49	3.81	0.78	1.04	1.28 $\pm$ 2.61
TROIKA [13]	3.05	3.31	1.49	2.03	1.46	2.35	1.76	1.43	1.28	5.08	1.80	3.02	2.34 $\pm$ 2.86

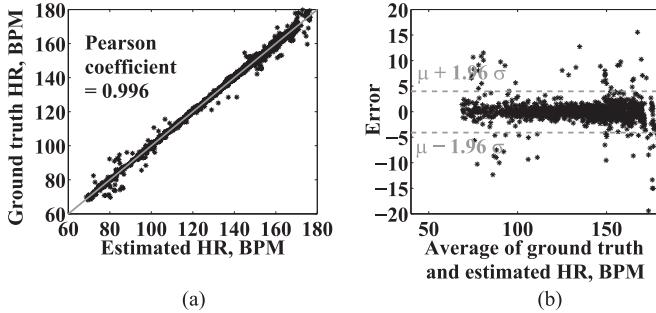


Fig. 13. (a) Pearson correlation between the ground truth and the estimated HRs over all the 12 datasets. (b) Bland-Altman plot over all the 12 datasets. Here,  $\mu$  is the mean of the error differences over all the estimates, and  $\sigma$  is the corresponding SD.

( $i = 1, 2$ ) and the single-channel ( $i = 1$ ) cases. In the same table, we also include the errors found by the methods of Zhang *et al.* [13], [14], for ease of comparison (HR estimates of JOSS and TROIKA were obtained through personal correspondence with the author). For the double-channel case (i.e., Proposed-*d*), averaged over all the estimates in the 12 datasets a gross average absolute error of 1.02 BPM is obtained with a standard deviation (SD) of 1.79 BPM. Here, the percentage error (absolute error/ground truth) gives an average of 0.79%. The Pearson correlation is calculated to be 0.996 as shown in Fig. 13(a). Fig. 13(b) gives the Bland-Altman plot [19] and the limit of agreement (LOA) is found to be  $[\mu - 1.96\sigma, \mu + 1.96\sigma] = [-4.10, 3.98]$  BPM ( $\mu$  being the mean of the error differences over all the estimates and  $\sigma$  being the corresponding SD).

Also note that for the single-channel case (i.e., Proposed-*s*), a gross average absolute error of 1.15 BPM is obtained (SD: 2.37 BPM) which is less than both JOSS and TROIKA's benchmarks. Here, the parameters are kept the same as for the double-channel case (obtained from the training phase). Since JOSS and TROIKA also work on single-channel PPG segments, the improvement in the single-channel result attests for the algorithmic superiority of the proposed method.

The error figures reported in Table I are for the datasets used in [14], which are not exactly the same as had been used in [13]. This is so because among the 12 datasets, 6 were very corrupted in the beginning and the method described in [14] did not initialize at these segments. Therefore, in [14], these initial corrupted segments were excluded and results were reported on a truncated version of the datasets. In Table II, we report the performance of our algorithm on the original untruncated datasets as well, along with the errors of TROIKA. Clearly, the

proposed algorithm, for both single channel and double channel, performs quite well even with noisy starts. This attests for the algorithm's needlessness of initial resting phase.

To illustrate the robustness of the algorithm against a change in exercise type, we applied the same algorithm along with the same set of parameters (obtained from the training phase) to a new test dataset where the subject performed boxing, hands shaking, jumps, and push-ups. The average absolute error was found to be 1.09 BPM [see Fig. 14(a)]. We also implemented TROIKA and found the corresponding error 1.97 BPM.

All the 12 datasets used previously were recorded for 12 healthy male subjects (age: 18–35). To illustrate the robustness of the algorithm against a change in subject profile, we ran our algorithm (with the same parameter set) for a female elderly subject (age: 58) with abnormal heart rhythm and blood pressure. We obtained an average absolute error of 0.66 BPM [see Fig. 14(b)], whereas the corresponding error for TROIKA was found 0.75 BPM. Therefore, it can be said that the algorithm performs well for these new test datasets as well.

### B. Sensitivity Analysis

To examine the performance of the algorithm for a change in the sampling frequency  $F_s$ , we apply the same algorithm for the two channel case at three more different sampling frequencies, e.g., 125, 250, and 500 Hz. The corresponding results are listed in Table III which shows that the algorithm performs equally well in higher sampling frequencies.

To illustrate the sensitivity of the algorithm's performance to its parameters, we run the simulations again (for double channels, at  $F_s = 25$  Hz) on all the 12 truncated datasets by altering the crucial parameters, namely  $N_E$  (number of realizations in an ensemble for performing EEMD),  $\Delta_0$  (default value of  $\Delta_{AC}$ , the BPM window length for AC),  $\Delta_d$  (dynamic increment of  $\Delta_{AC}$ ),  $\Delta_{imf}$  (tracking range for the IMFs),  $\tau$  (the threshold for time-domain extraction),  $\Delta_t$  (tracking range for the time-domain extracts),  $s$  (filter length for RLS),  $\Delta_{rls}$  (tracking range for RLS filter output), and  $\Delta_{fine}$  (fine error threshold). In this paper, we used  $N_E = 5$ ,  $\Delta_0 = 5$  BPM,  $\Delta_d = 1$  BPM,  $\Delta_{imf} = 7$  BPM,  $\tau = 0.3$ ,  $\Delta_t = 12$  BPM,  $s = 8$ ,  $\Delta_{rls} = 9$  BPM, and  $\Delta_{fine} = 4$  BPM. To test for sensitivity for the parameter  $N_E$ , we keep all the other parameters unchanged while setting  $N_E$  to 3 or 7 from its default value 5. For each of these cases, the gross average absolute error for the 12 datasets is recorded in Fig. 15. Similarly, sensitivity analysis for other parameters is performed by setting  $\Delta_0 = 4$  or 7 BPM,  $\Delta_d = 1.5$  BPM,  $\Delta_{imf} = 5$  or 10 BPM,  $\tau = 0.4$ ,  $\Delta_t = 7$  or 10 BPM,  $s = 9$  or 10,  $\Delta_{rls} = 7$  or 11 BPM, and  $\Delta_{fine} = 3$  BPM each in turn. As Fig. 15 shows the

TABLE II  
PERFORMANCE COMPARISON USING AVERAGE ABSOLUTE ERRORS FOR THE 12 UNTRUNCATED DATASETS USED IN [13] FOR  $F_s = 25$  Hz

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	Subj 11	Subj 12	Mean $\pm$ SD
Proposed- <i>d</i>	1.64	0.81	0.57	1.44	0.77	1.06	0.63	0.47	0.52	2.94	1.05	0.91	1.07 $\pm$ 2.17
Proposed- <i>s</i>	2.55	3.45	0.73	1.19	0.51	1.09	0.52	0.43	0.36	3.33	0.89	0.98	1.33 $\pm$ 3.32
TROIKA [13]	3.05	3.49	1.49	2.03	1.46	2.35	1.76	1.42	1.28	5.73	1.79	3.02	2.41 $\pm$ 3.45

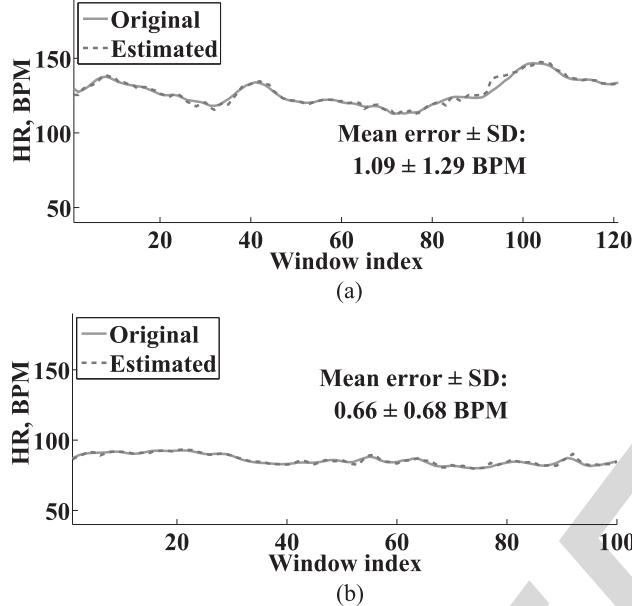


Fig. 14. (a) Performance of the proposed algorithm for a new set of datasets where the subject, instead of running, performed intensive forearm and upper arm movements (e.g., boxing). (b) Algorithm's performance for a subject (age 58, female) with abnormal heart rhythm and blood pressure.

performance remains more or less the same (with gross average absolute error around 1 BPM). This proves the algorithm's robustness to parameter values.

We also determine the algorithm's performance sensitivity to the various speed settings. Each of the 12 runs consisted of three types of phases: rest, moderate, and intense phases. The rest phase  $R$  was at the beginning and at the end of the run. The moderate phase  $M$  was the normal jogging condition at 6–8 km/h, whereas the intense phase  $I$  was at a more higher speed (12–15 km/h). Both the moderate and the intense phases were exerted twice in an alternating fashion. Therefore, the running protocol for speed sensitivity analysis is chosen to be

$$R \longrightarrow M \longrightarrow I \longrightarrow M \longrightarrow I \longrightarrow R.$$

Fig. 16(a) illustrates the algorithm's performance in each of these phases. Clearly, the estimated HR curve closely follows the ground truth HR in all the three phases in an identical manner. No discernible change in performance is noticed as the speeds are varied during the runs. Similar result also holds for the rest of the datasets.

Finally, Fig. 16(b) shows the algorithm's sensitivity to the presence or absence of one of its major subalgorithms, namely the use of AC. If the HR estimation goes wrong (in the figure, it is deliberately let carried away), clearly it is for this feature

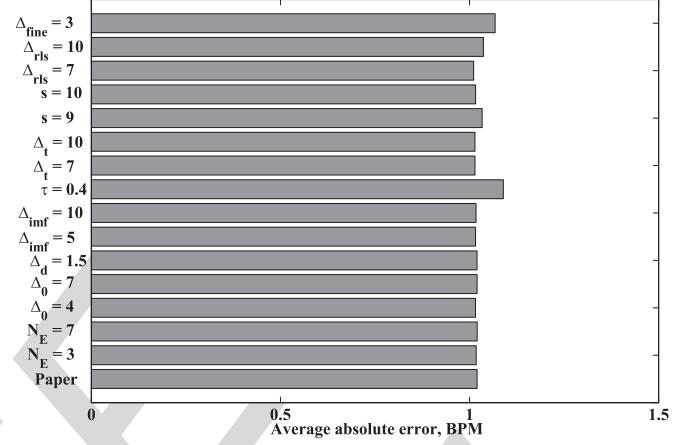


Fig. 15. Robustness of the proposed algorithm to parameter values. "Paper" indicates the result with parameter values given in Section V-B. " $N_E = 3$ " indicates the result with all parameter values unchanged except for  $N_E$  which is changed to 3 from its default value 5. Similar meanings go to other y-labels in the figure.

that the estimation is brought back ontrack and a total runaway situation is avoided.

### C. Discussion

From the above figures and tables, it is quite clear that the proposed algorithm is better than the two recent works done in this field [13], [14]. It has a gross average absolute error superior to both TROIKA and JOSS, and the error variance is also significantly improved compared to both of the algorithms. TROIKA and JOSS reported LOA around  $[-7.26, 4.79]$  and  $[-5.94, 5.41]$ , respectively, whereas the proposed algorithm has a LOA of  $[-4.10, 3.98]$ . On the other hand, the Pearson correlation coefficient was found to be 0.992 in TROIKA and 0.993 in JOSS, while it is 0.996 in our proposed algorithm. But more importantly, increased robustness and stability together with no requirement of initial resting boot phase are the key features in the algorithm that we would like to highlight.

To better compare the performance of the proposed method with that of TROIKA and JOSS, Fig. 17 plots the histograms of the difference between the absolute error of the proposed method and the absolute error of either TROIKA [see Fig. 17(a)] and JOSS [see Fig. 17(b)] at each estimate over the 12 datasets, i.e., the histogram of  $e_{\text{proposed}}(l) - e_{\text{TROIKA}}(l)$  and  $e_{\text{proposed}}(l) - e_{\text{JOSS}}(l)$ , where  $e_{\text{proposed}}(l)$  indicates the absolute estimation error of the proposed method at the  $l$ th HR estimate, whereas  $e_{\text{TROIKA}}(l)$  and  $e_{\text{JOSS}}(l)$  indicate those of TROIKA and JOSS, respectively. From Fig. 17, it is clear that both histograms run downhill faster along the positive sides.

TABLE III  
AVERAGE ABSOLUTE ERRORS FOR THE 12 DATASETS USED IN [14] FOR DIFFERENT SAMPLING FREQUENCY  $F_s$

$F_s$ , Hz	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	Subj 11	Subj 12	Mean ± SD
25	1.70	0.84	0.56	1.15	0.77	1.06	0.63	0.53	0.52	2.56	1.05	0.91	1.02 ± 1.79
125	1.83	0.85	0.63	1.21	0.65	1.03	0.70	0.50	0.47	2.83	1.14	0.90	1.06 ± 2.02
250	1.96	0.93	0.64	1.19	0.68	1.00	0.73	0.49	0.50	2.95	1.21	0.95	1.10 ± 1.96
500	1.80	0.87	0.64	1.23	1.05	1.03	0.68	0.50	0.50	3.00	1.22	0.88	1.12 ± 2.11

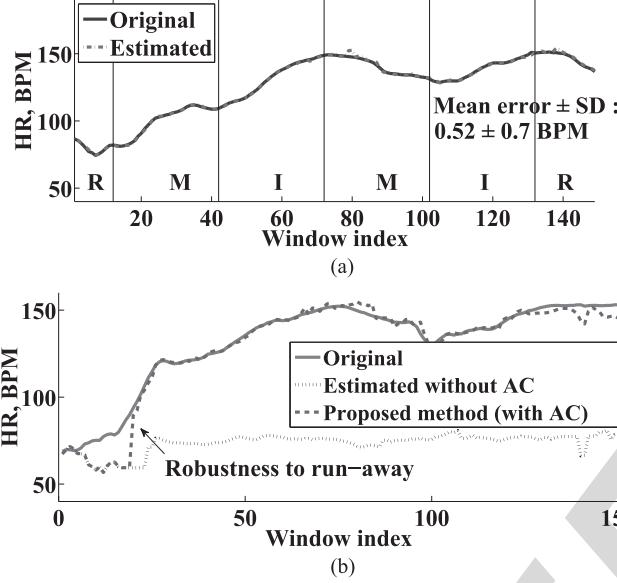


Fig. 16. (a) Performance of the proposed algorithm is illustrated at various speed settings. Here, R, M, and I keywords stand for “Rest,” “Moderate,” and “Intense” phases of the running protocol. (b) Algorithm shows quick robust turn back even though it is deliberately let carried away. This is due to the use of the AC, without which we have a runaway situation plotted in the same figure window.

This implies there are more estimates on the negative side; which in turn implies that for large number of estimates we have  $e_{\text{proposed}}(l) - e_{\text{TROIKA}}(l) < 0$  (75% of total estimates) and  $e_{\text{proposed}}(l) - e_{\text{JOSS}}(l) < 0$  (56% of the total estimates). Based on the  $t$ -test, the absolute estimation error of the proposed method is found to be significantly smaller than those of TROIKA and JOSS at the significance level  $\alpha = 0.01$ , and the corresponding  $p$  values are  $3.7415 \times 10^{-66}$  and  $3.1139 \times 10^{-5}$ , respectively.

In Fig. 16(b), we see that even if the algorithm mistakes an MA peak as HR, as soon as a strong reliable peak is discovered later on, the correct HR is retained and a runaway situation is avoided. But even so, a considerable amount of offtrack error is produced thereby. In addition to making the algorithm robust to runaway situations, it is also imperative to make it robust against offtrack conditions. Now at the advent of offtracking, the HR peak and the MA peak (or its harmonics) usually get very close to each other. If proper spectral peak separation techniques are not used, then a wrongly chosen peak will give rise to an offtrack scenario. In our algorithm, we used, in addition to RLS filters, a novel technique to address this issue, namely time-domain extraction. Their combined performance is shown in Fig. 18. From Fig. 18(a), we can see the dataset is corrupted enough

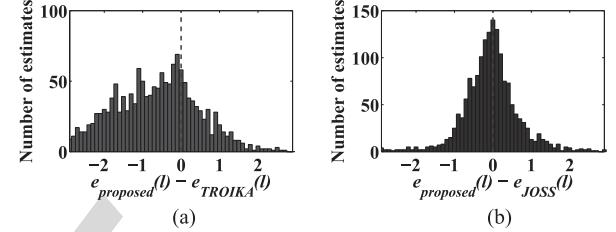


Fig. 17. Histograms of the difference between the absolute error of the proposed method and the absolute error of (a) TROIKA and (b) JOSS at each estimate over the 12 datasets. Based on the  $t$ -test, the absolute estimation error of the proposed method is found to be significantly smaller than those of TROIKA and JOSS at the significance level  $\alpha = 0.01$ , and the corresponding  $p$  values are  $3.7415 \times 10^{-66}$  and  $3.1139 \times 10^{-5}$ , respectively.

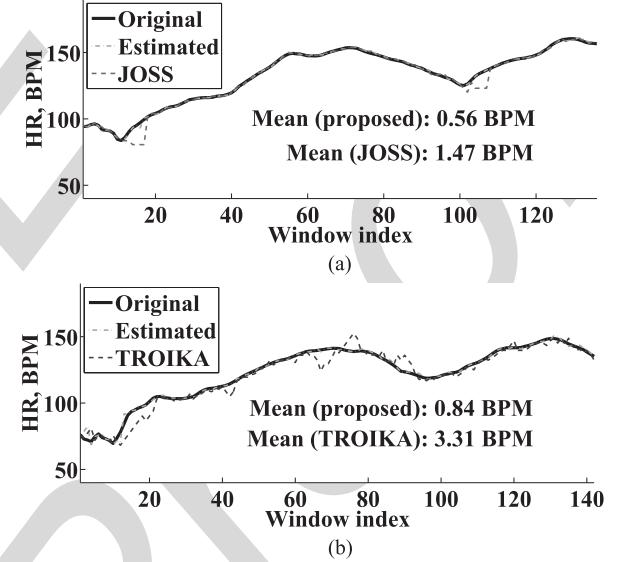


Fig. 18. Illustration of the proposed algorithm’s robustness against offtracking. (a) Proposed algorithm is successful in choosing correct peaks and avoid any offtracking pitfalls where JOSS failed. (b) Proposed algorithm is hardly offtracked, but TROIKA is derailed quite often.

to derail JOSS and incur some offtrack errors. But in the same pitfalls, the proposed algorithm was able to correctly choose the HR peaks and was not offtracked. In Fig. 18(b), we compare between the performances of the proposed algorithm and TROIKA. As we can see, TROIKA was offtracked quite frequently and had many kinks, whereas the proposed algorithm hardly lost track of HR.

Both TROIKA and JOSS estimate HR from a single channel of PPG. Though our algorithm for single channel (see Proposed-*s* in Table I) performs better than both of them, we find that

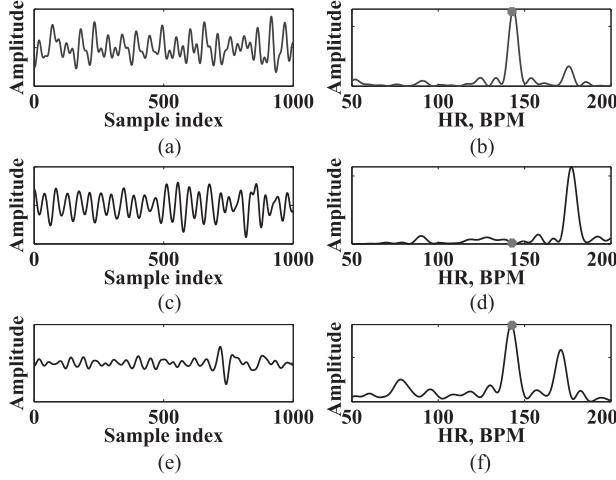


Fig. 19. Advantage of using two channels of PPG signals. Clearly channel 1 is distorted and does not contain HR information, whereas channel 2 readily gives the HR peak from its periodogram. (a) Original ECG. (b) Periodogram of ECG. (c) PPG Channel 1. (d) Periodogram of PPG Channel 1. (e) PPG Channel 2. (f) Periodogram of PPG Channel 2.

utilizing two channels of PPG gives even better results (see Proposed-*d* in Table I). The advantage of having an extra channel lies in the fact that even if one of the pulse oximeters loses contact with the skin or is heavily corrupted by internal shot noise, the other might still be functioning well and provide with the HR peak for that period of time. Since the proposed algorithm hardly mixes the two channels, instead it works with the peaks that they individually can provide with, the algorithm can estimate the HR even if one of the channels is malfunctioning. Obviously, any single-channel algorithm will be at a disadvantage in these situations. To illustrate this, Fig. 19(a) shows an ECG segment along with its periodogram in (b). Simultaneously recorded two segments of PPG signals are shown in Fig. 19(c) (channel 1) and (e) (channel 2) together with their periodograms shown in (d) and (f), respectively. Clearly, the PPG segment from channel 1 does not contain the HR information, but the PPG segment from channel 2 does.

Furthermore, both JOSS and TROIKA use signal sparsification techniques through the M-FOCUSS algorithm which is very much computationally expensive. For example, for  $F_s = 125$  Hz, TROIKA takes about 3.5 h to estimate HR for all the 12 datasets on a computer equipped with Intel Core-i7 4790 at 3.60 GHz, 8-GB RAM, Windows 7 64 bit, and MATLAB 2013a, whereas our proposed algorithm takes 668 s on the same computer. Similarly, on the same computer, for  $F_s = 25$  Hz, JOSS takes 300 s for all the datasets, whereas our proposed algorithm takes only 199 s.

Though all the algorithms discussed in this paper try to estimate HR instantaneously on the current time window, a significant improvement in the results can be achieved if the real-time implementation allows for a lag of around two time frames. With such a provision for delayed HR estimation, correct HR peaks can be chosen with more confidence. Fig. 20 illustrates this idea where our proposed algorithm is shown in two versions. In the original instantaneous version, there are some few inevitable

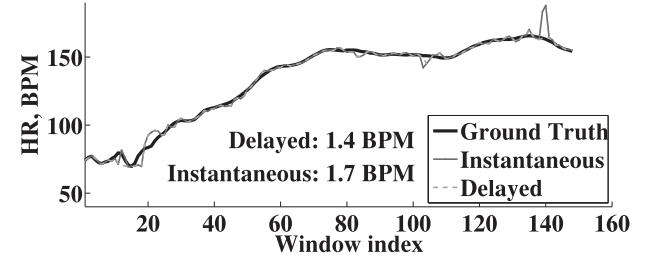


Fig. 20. Advantage of using a lag of just two time frames in real time. As opposed to the instantaneous algorithm, the delayed algorithm (delayed by 4 s) can identify outlier estimates and smooth them out.

kinks and sideways drifts in the HR estimation, whereas for the delayed version the small offtrack errors can be corrected in real time by identifying the outlier estimates and smoothing them out (in the figure we use delay of two frames which, in our present scenario, corresponds to a lag of only 4 s). Surely, such a small time lag is not as frightening to the user as displaying incorrect HR values.

## VI. CONCLUSION

This paper has proposed a new robust method for HR monitoring from dual-channel PPG signals corrupted by intense MAs. The algorithm mainly relies on EEMD-based clean PPG-mode extraction, adaptive filter-based denoising using three-axis accelerometer data, and some heuristic approaches for decision making. The salient feature of the proposed algorithm is that, unlike other recently reported well-known algorithms, it does not require an initial resting phase. Moreover, because of the use of two channel PPG signals, the algorithm has been found to be less susceptible to offtrack and runaway errors and, therefore, more robust and stable. The performance results using 14 PPG datasets of different subjects collected at various speed settings and motion activities have demonstrated superiority over other recent methods in terms of several figures of merit (e.g., average absolute error, error SD, limit of analysis, Pearson correlation coefficient etc.). In future works, we would like to investigate the algorithm's robustness in other kinds of athletic, gymnastic, aeronautical, military, and veterinary activities.

## REFERENCES

- [1] A. Kamal *et al.*, "Skin photoplethysmography review," *Comput. Methods Programs Biomed.*, vol. 28, no. 4, pp. 257–269, 1989.
- [2] J. Allen, "Photoplethysmography and its application in clinical physiological measurement," *Physiol. Meas.*, vol. 28, no. 3, pp. R1–R39, 2007.
- [3] B. S. Kim and S. K. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 3, pp. 566–568, Mar. 2006.
- [4] 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.
- [5] R. Krishnan *et al.*, "Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 8, pp. 1867–1876, Aug. 2010.
- [6] M. Ram *et al.*, "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.

- [7] X. Sun *et al.*, "Robust heart beat detection from photoplethysmography interlaced with motion artifacts based on empirical mode decomposition," in *Proc. IEEE Int. Conf. Biomed. Health Informat.*, 2012, pp. 775–778.
- [8] F. Peng *et al.*, "Motion artifact removal from photoplethysmographic signals by combining temporally constrained independent component analysis and adaptive filter," *Biomed. Eng. Online*, vol. 13:50, no. 1, 2014.
- [9] K. Lam *et al.*, "Undergraduate students compete in the IEEE Signal Processing Cup: Part 1," *IEEE Signal Process. Mag.*, vol. 32, no. 4, pp. 123–125, Jul. 2015.
- [10] B. Lee *et al.*, "Improved elimination of motion artifacts from a photoplethysmographic signal using a Kalman smoother with simultaneous accelerometry," *Physiol. Meas.*, vol. 31, no. 12, pp. 1585–1603, 2010.
- [11] R. Yousefi *et al.*, "A motion-tolerant adaptive algorithm for wearable photoplethysmographic biosensors," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 2, pp. 670–681, Mar. 2014.
- [12] S. M. López *et al.*, "Heuristic algorithm for photoplethysmographic heart rate tracking during maximal exercise test," *J. Med. Biol. Eng.*, vol. 32, no. 3, pp. 181–188, 2012.
- [13] Z. Zhang *et al.*, "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.
- [14] 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.
- [15] K. T. Sweeney *et al.*, "The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 1, pp. 97–105, Jan. 2013.
- [16] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, 2009.
- [17] S. S. Haykin, *Adaptive Filter Theory*. Pearson Education, Delhi, India, 2008, ch. 9.
- [18] M. A. Arafat *et al.*, "A simple time domain algorithm for the detection of ventricular fibrillation in electrocardiogram," *Signal, Image Video Process.*, vol. 5, no. 1, pp. 1–10, 2011.
- [19] 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.



**Emroz Khan** was born in Bangladesh in 1992. He received the B.Sc. degree in electrical and electronic engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2015.

His research interest includes digital signal processing, electron microscopy, graphene and quantum information.

Mr. Khan participated in the IEEE Signal Processing Cup, Florence, Italy, in 2014, and his team secured the third place. He is currently a Student Member of the IEEE Signal Processing Society and the IEEE Electron Devices Society. Since 2010, he has been associated with the academic team of Bangladesh Physics Olympiad.

the IEEE Signal Processing Society and the IEEE Electron Devices Society. Since 2010, he has been associated with the academic team of Bangladesh Physics Olympiad.



**Forsad Al Hossain** was born in Bangladesh in 1992. He received the B.Sc. degree in electrical and electronic engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2015.

His research interest includes digital signal processing, electron microscopy and algorithm development.

Mr. Hossain was a participant at the IEEE Signal Processing Cup 2014, Florence, Italy, where his team achieved the first position. He is currently a Student Member of the IEEE Signal Processing Society. He was an Active Competitive Programming Contestant at the university.



**Shiekh Zia Uddin** was born in Bangladesh in 1992. He received the B.Sc. degree in electrical and electronic engineering from the Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 2015.

He is currently with the Nanophotonics Research Group, BUET. His research interest includes digital signal processing algorithms, inverse problems, plasmonic biosensors and SMD microscopy.

Mr. Uddin's team received the third place at the IEEE Signal Processing Cup, Florence, Italy, in 2014.



**S. Kaisar Alam** was born in Bangladesh. He received the Bachelor of Technology(Hons.) degree in electronics and electrical communication engineering from the Indian Institute of Technology Kharagpur, Kharagpur, India, in 1986, and the M.S. and Ph.D. degrees in electrical engineering from the University of Rochester, Rochester, NY, USA, in 1991 and 1996, respectively.

From 1986 to 1989, he was a Lecturer with the Department of Electrical and Electronic Engineering, Bangladesh Institute of Technology, Rajshahi, Bangladesh. While he was with the University of Rochester, his work on improving the ultrasonic estimation of blood velocity led to a patented technique for ultrasonic blood flow imaging. As a Postdoctoral Fellow with the University of Texas Health Science Center at Houston from 1995 to 1998, he worked on elastography, a new ultrasonic imaging modality to image elastic properties of tissue. He was a Principal Investigator with Riverside Research, New York, USA, for more than 15 years, where he worked on a variety of research topics in biomedical imaging. He is currently the Chairman and the Chief Research Officer at Imprlabs Pte Ltd., Singapore, an upcoming tech startup in Singapore. Since 2010, he has been a Visiting Faculty at the Department of Electrical and Electronic Engineering, Islamic University of Technology, Gazipur, Bangladesh. Since 2013, he has been a Visiting Research Professor at the Center for Computational Biomedicine Imaging and Modeling, Rutgers University, the State University of New Jersey, New Brunswick, NJ, USA. He has written more than 35 papers in international journals and holds several patents. He is the Coauthor of the textbook on Computational Health Informatics, scheduled to be published in 2016. His research interests include diagnostic and therapeutic applications of ultrasound and optics, and signal/image processing with applications to medical imaging.

Dr. Alam received the prestigious Fulbright Scholar Award in 2011 and 2012, respectively. He is a Member of the Sigma Xi, the Acoustical Society of America, and the Society of Photographic Instrumentation Engineers, and a Senior Member of the American Institute of Ultrasound in Medicine. He was with the AIUM Technical Standards Committee and the RSNA QIBA US SWS Technical Committee and is currently a Member of the RSNA QIBA Ultrasound Coordinating Committee. He is an Associate Editor of the *Ultrasonics and Ultrasonic Imaging* and a Member of the Editorial Board of the *Journal of Medical Engineering*.



**Md. Kamrul Hasan** received the B.Sc. and M. Sc. degrees in electrical and electronic engineering from the Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh, in 1989 and 1991, respectively, and the M.Eng. and Ph.D. degrees in information and computer sciences from Chiba University, Chiba, Japan, in 1995 and 1997, respectively.

In 1989, he joined BUET as a Lecturer with the Department of Electrical and Electronic Engineering, where he is currently working as a Professor. He was a Postdoctoral Fellow and a Research Associate at Chiba University and Imperial College, London, U.K., respectively. He was a short-term invited Research Fellow at the University of Tokyo, Tokyo, Japan, and a Professor of International Scholars of Kyung Hee University, Seoul, Korea. He has authored or coauthored more than 100 scientific publications. His current research interests include digital signal processing, adaptive filtering, speech and image processing, and medical imaging.

Dr. Hasan is currently an Associate Editor for the IEEE ACCESS.