SCIENCE CHINA
Information Sciences

A novel approach framework based on statistics for reconstruction and heartrate estimation from PPG with heavy motion artifacts

Bo PANG, Ming LIU, Xu ZHANG[1], Peng LI[1] and Hongda CHEN[1]

Citation: SCIENCE CHINA Information Sciences **61**, 022312 (2018); doi: 10.1007/s11432-017-9168-2

View online: https://engine.scichina.com/doi/10.1007/s11432-017-9168-2

View Table of Contents: https://engine.scichina.com/publisher/scp/journal/SCIS/61/2

Published by the Science China Press

SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

February 2018, Vol. 61 022312:1–022312:12
doi: 10.1007/s11432-017-9168-2

A novel approach framework based on statistics for reconstruction and heartrate estimation from PPG with heavy motion artifacts

Bo PANG[1,2], Ming LIU[1*], Xu ZHANG[1], Peng LI[1] & Hongda CHEN[1]

[1]*State Key Laboratory on Integrated Optoelectronics, Institute of Semiconductors, Chinese Academy of Sciences, Beijing* 100083*, China;*
[2]*University of Chinese Academy of Sciences, Beijing* 100049*, China*

Received 2 May 2017/Accepted 27 June 2017/Published online 13 November 2017

Abstract One of the most important applications of photoplethysmography (PPG) signal is heartrate (HR) estimation. For its applications in wearable devices, motion artifact (MA) may be the most serious challenge for randomness both in format and temporal distribution. This paper proposes an advanced time-frequency analysis framework based on empirical mode decomposition (EMD) to select specific time slices for signal reconstruction. This framework operates with a type of pre-processing called variance characterization series (VCS), EMD, singular value decomposition (SVD), and a precise and adaptive 2-D filtration reported first. This filtration is based on Harr wavelet transform (HWT) and 3rd order cumulant analysis, to make it have resolution in both the time domain and different components. The simulation results show that the proposed method gains 1.07 in absolute average error (AAE) and 1.87 in standard deviation (SD); AAEs' 1st and 3rd quartiles are 0.12 and 1.41, respectively. This framework is tested by the PhysioBank MIMIC II waveform database.

Keywords photoplethysmography (PPG), motion artifact, empirical mode decomposition (EMD), singular value decomposition, discrete wavelet transform, higher-order statistics

Citation Pang B, Liu M, Zhang X, et al. A novel approach framework based on statistics for reconstruction and heartrate estimation from PPG with heavy motion artifacts. Sci China Inf Sci, 2018, 61(2): 022312, doi: 10.1007/s11432-017-9168-2

1 Introduction

Photoplethysmography (PPG) is an optical method to measure periodic change in peripheral blood volume. It has the advantages of small size, low power consumption, noninvasiveness and convenience. With PPG signals, we can obtain several important indicators showing the health condition of a human body, such as heartrate (HR) [1, 2], oxyhemoglobin saturation, and blood press. Among them, the HR estimation by PPG signals has received great attention. However, the most serious challenge of PPG processing is motion artifact (MA), which can be involved by finger movement, tremble, and probe mismatch, etc. Typically, the frequency band of PPG is under 10 Hz [3]. Compared with baseline drift and power line interference, MA cannot be removed by traditional frequency domain methods because the frequency band of the PPG signals may be completely overlapped by MAs, as shown in Figure 1. The lower sample in Figure 1(a) is a length of PPG with MA, and Figure 1(b) is the corresponding

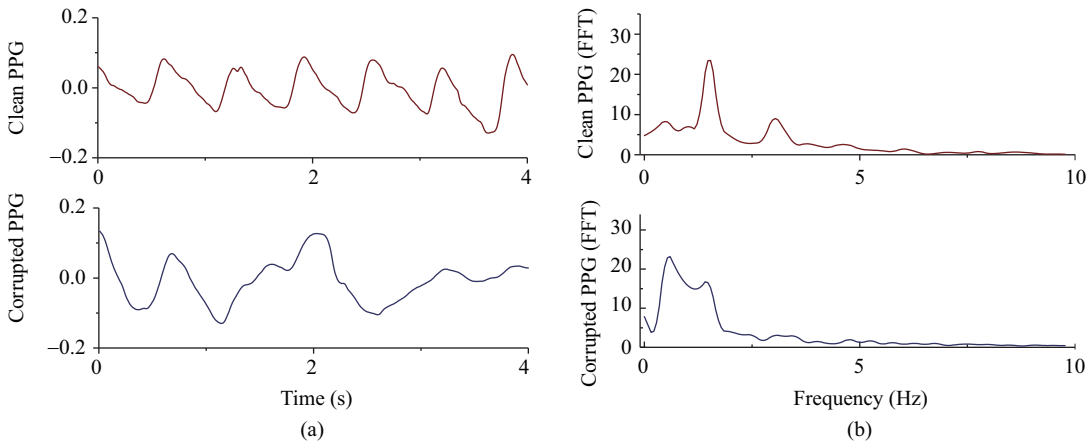

Figure 1 (Color online) (a) Clean PPG siganl and PPG with MA corruption; (b) corresponding frequency spectrums.

frequency spectrum. It is clear that MA can cover up the original signal in the frequency domain. Hence, time-frequency methods are applied to PPG processing against the effects of MAs.

Some work has been performed on MA removal and HR estimation. Zhang et al. [4] has reported a new framework named TROIKA, which consists of signal decomposition, reconstruction and spectral peak tracking (SPT). The signal decomposition is generally based on singular spectrum analysis (SSA), and sparse signal reconstruction (SSR) is used for signal reconstruction. In this way, most of MAs are removed from raw PPG signals, and SPT can be feasibly applied. The MISPT method is largely based on the SPT method of TROIKA [5], and the main improvement is to perform SPT both in forward and backward directions. Hence, more trajectories would be acquired and the final detection result would be more accurate.

Noting that even the uncorrupted PPG signal is non-stable and non-continuous, common time-frequency methods such as wavelet transform would not be optimal. Thus, empirical mode decomposition (EMD) is employed to solve this problem. EMD is a powerful feature extractor and adaptively decomposes the input signal into a sort of intrinsic mode functions (IMFs). Different orders of IMFs contain different components of the original signal [6–8]. By recombining those IMFs indicating the PPG signal, noises including MAs would be removed or reduced. As mentioned above, EMD is a type of adaptive decomposition. The chosen IMFs would vary not only from time to time but also from sample to sample. The primary idea to solve this is to characterize each IMF by statistics or frequency-domain methods and then classify them. Some impressive work has been reported such as ensemble EMD (EEMD) and further strategies. When a length of Gaussian noise (GN) is decomposed by EMD, EMD acts like a series of dyadic band-pass filters [9]. When the raw signal is buried by random GN and this process is repeated, the IMFs of the valuable band would be selected using certain methods. Furthermore, EEMD also helps remove the original GN from inputs. The EEMD-PCA method is an extension of EEMD in PPG processing [10]. It first decomposes PPG into a sort of IMFs by EEMD and then uses band-pass filters for selection. Finally, principal component analysis (PCA) is used to extract the principal component from the seleted IMFs. In general, all the above frameworks or methods consist of three parts: decomposition, filtration and reconstruction. However, in heavy MA situations, the distribution of distortions caused by MAs would be random. The former characterization methods may be confused as statistical feature changes in the inputs.

To solve this problem and the uncertainty of EMD itself, this study created a novel framework based on statistical characteristics from raw signals and deuterogenic vectors. This framework focused on the reconstruction of heavily corrupted PPG signals and the following HR estimation. Compared with TROIKA, this framework applied SSA to the components of the same time slice rather than a set of continuous time slices. Hence, the influence of variation in the input signal was maximally avoided. SSR was also unnecessary, as these components could be directly used for reconstruction. Compared with

* Corresponding author (email: liuming@semi.ac.cn)

traditional EMD and EEMD, this method applied variance characterized series (VCS) [11] to reduce the computing cost. We also introduced a type of two-dimensional filtration strategy in signal reconstruction, based on VCS, Harr wavelet transform and three-order cumulant analysis.

This paper is organized as follows. In Section 2, the main methods used are introduced in brief. In Section 3, this proposed framework is introduced in detail. Section 4 reports the experimental settings and discusses the results, including the parameter settings and performance test. The up-to-date HR estimation methods in PPG processing are also carefully compared. The final section presents the conclusion.

2 Theory in use

2.1 EMD

EMD is part of Hilbert-Huang transform and is a typical adaptive time-frequency method [12]. The most attractive aspect of EMD is that it does not require certain basis for signal decomposition. Hence, it is confirmed that the EMD method is highly effective and versatile for noisy non-linear and non-stationary signal processing, including physiological signals. The EMD flow can be summarized as follows.

(1) Find all the local maximums and minimums, connect all the local maximums, then minimums by cubic spline interpolation to gain the upper envelope $M(t)$ and lower envelope $m(t)$.

(2) Let $C(t) = (M(t) + m(t))/2$, then extract the original signal $S(t) = S(t) - C(t)$.

(3) Repeat the above two steps until $S(t)$ becomes an IMF, thus we get the first order IMF, i.e., IMF1.

(4) Make $S(t) = S(t) - \text{IMF1}$ and and repeat the above 3 steps until no more IMF can be extracted while the original signal turn into a monotonic function $\gamma(t)$. After the EMD process, the original signal is decomposed into

$$S(t) = \Sigma\text{IMF} + \gamma(t). \tag{1}$$

Different orders of IMFs contain different components from the raw PPG signal. In general, the first several orders indicate the high-frequency parts, and the last orders refer to low-frequency components. The basic principle of using EMD in PPG signal reconstruction is to decompose the inputs into IMFs, recombine the IMFs indicating PPG signals, and then abandon the noisy IMFs. It functions well when the noise is not complex. In this study, we aimed to manage serious noise. In such cases, the IMFs might be badly disturbed. As illustrated in Figure 2, two PPG slices are both decomposed by EMD. The corresponding first 6 orders of IMFs are also plotted in Figure 2(b) and (c). Figure 2 shows that when the input is clean, EMD functions well, whereas when MA corruption occurs, each IMF might be disturbed to different degrees. It would not be easy to regain a clean PPG by only selecting out the relevant orders.

2.2 Variance characterization series

In our former work, using variance characterization series (VCS) to recognize deviant slices in statistics was a method we have proposed to locate abnormal variations from a length of stationary signals [11]. As MA does not occur along the whole time domain, the computing cost can be reduced by selecting the periods requiring processing. The flow of VCS is shown below.

(1) Pick out all the local maximums M_i and minimums m_i;

(2) Computing the variance σ_{Mi} of each maximum and the following seven maximums, then the minimums' σ_{mi}. The last eight extrema take the same value;

(3) For each variance, make

$$\delta_{Mi} = |\sigma_{Mi}|/(\text{location}(M_{i+1}) - \text{location}(M_i))^2, \tag{2}$$

$$\delta_{mi} = |\sigma_{mi}|/(\text{location}(m_{i+1}) - \text{location}(m_i))^2. \tag{3}$$

The $\delta_{Mi/mi}$ series are the VCS of the signal $S(t)$. For each point in VCS, the value and the distances between its consecutive points are both taken into consideration in the following discussion. The VCS

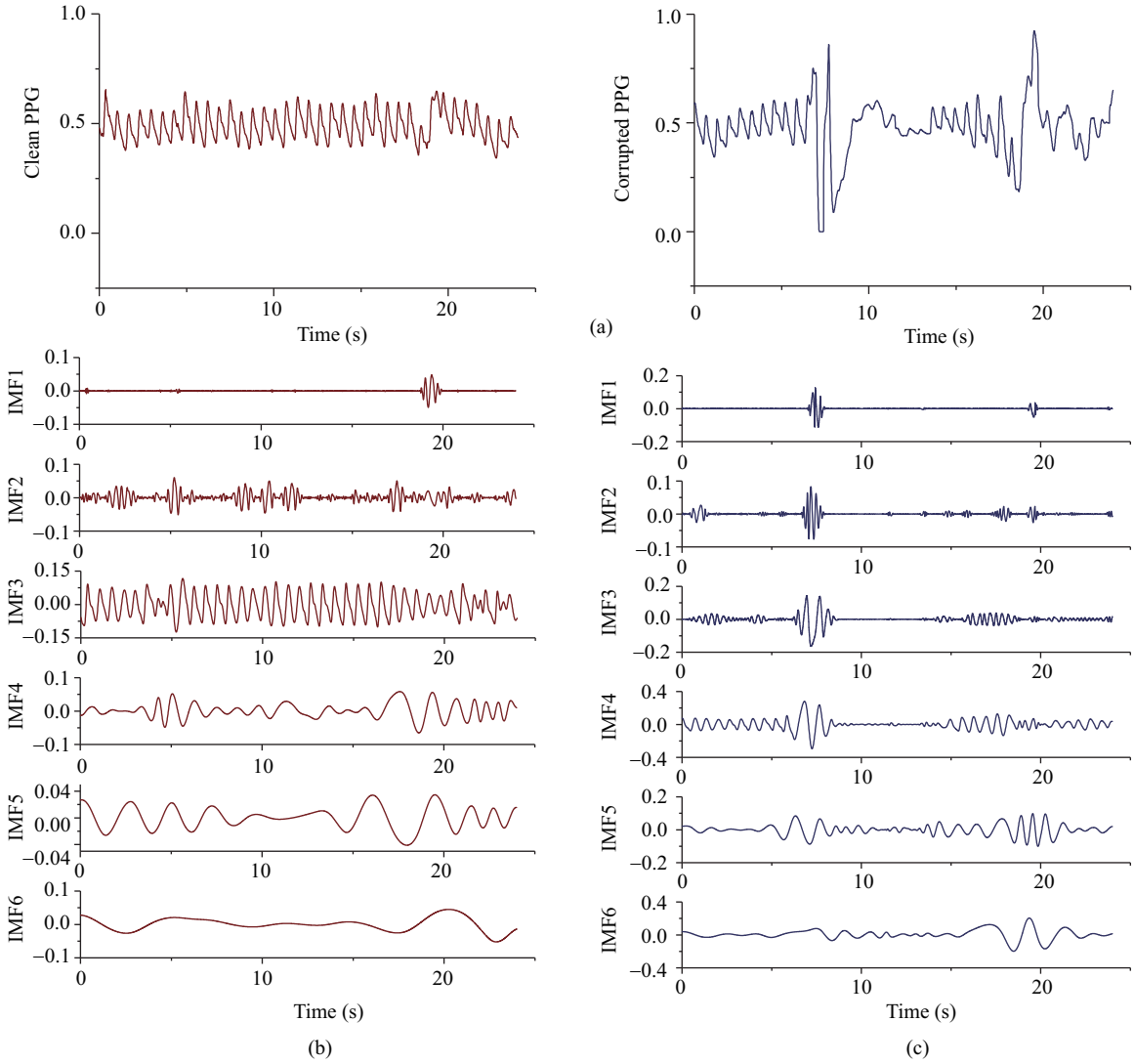

Figure 2 (Color online) (a) Clean PPG signal and PPG with MA corruption; (b) IMFs extracted from clean signal; (c) IMFs extracted from PPG with MA corruption.

in the uncorrupted period performed stably. When MA occurred, abrupt changes would show up in the values or the distances in VCS. The threshold values $\{\epsilon_{\text{upper/lower}}, \text{dis}_{\text{upper/lower}}\}$ were user-defined at first. The ϵ stands for the value threshold and 'dis' for distance. If a VCS unsatisfied constrained condition occured in a length of signal, such a slice would be regarded as corrupted. In this way, the corrupted slices could be selected rapidly and precisely from the whole signal. Then, four cycles of contiguous signals were added to both the head and tail of each selected slice to reduce spectrum leakage. This also helped to reduce the risk of overfitting caused by EMD's adaptive characteristics. The VCS could be regarded as a rapid pre-process to estimate the stationarity of signals. This method provides a rapid way to reduce costs in time and power consumption.

VCS can help reduce the length of input data. Furthermore, according to the former step, the decomposition of EMD should be ended when the number of maxima or minima is reduced to less than 8 in the current IMF vector.

2.3 Harr wavelet transform

Harr wavelet transform (HWT) is a type of wavelet transform. It is named after the mother wavelet function [13]:

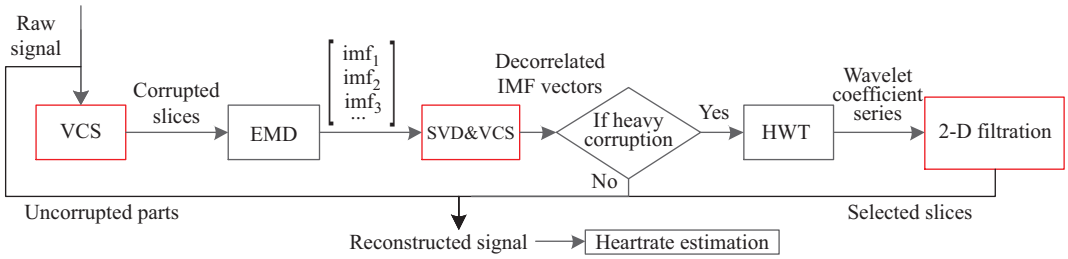

Figure 3 (Color online) Flowchart for the proposed method. Each processing is shown in module format. Variables passed among modules are shown.

$$\psi(t) = \begin{cases} 1, & 0 \leqslant t \leqslant 0.5, \\ -1, & 0.5 < t \leqslant 1. \end{cases} \tag{4}$$

A standard wavelet transform (WT) is

$$\mathrm{WT}_x(a,b) = \frac{1}{\sqrt{a}} + \int f(t)\psi\left(\frac{t-b}{a}\right)\mathrm{d}t, \tag{5}$$

where a and b are scale and translation factors respectively. HWT is an orthogonal transform and a type of powerful time-frequency analysis. One of its shortcomings is the inability to handle non-Gaussian noises. Note that one of the difficulties in this paper is that MA has strong non-Gaussian component. Compared with EMD, the decomposition of HWT is limited by the mother wavelet function, resulting in poor resolution of complex noise cases. However, wavelet de-noising was taken into account as an important procedure because we found it could be an indispensable addition for statistically analyzing IMF. The details are discussed below.

3 Filtration framework

3.1 Corrupted slices detection

A raw PPG signal was sent to VCS as pre-processing to reduce the cost of EMD and later processing, as shown in Figure 3. In our former work, the thresholds mentioned in Subsection 2.1 were completely user-defined. A small promotion conducted in this paper is that these thresholds would be updated by the uncorrupted parts in the current frame by averaging the values in VCS after processing. This dynamic threshold' update strategy was proposed to handle more drastic changes caused by types of MAs.

3.2 Signal decomposition and primary filtration

As shown in Figure 3, the selected slices were sent to EMD. The first two orders were removed because they were largely the high-frequency noise parts. The remaining N-orders IMFs were mapped into an $L * N$ matrix (L is the length of input corrupted slice and $*$ stands for multiplication of numbers in this paper),

$$Y = \begin{bmatrix} \mathrm{imf}_2 \\ \mathrm{imf}_3 \\ \cdots \\ \mathrm{imf}_{N+1} \end{bmatrix}^{\mathrm{T}}. \tag{6}$$

By PCA, the principal component of Y matrix could be extracted. Obviously, this component was most likely the cardiac beats [10]. However, there were still two problems to be solved.

(1) PCA must calculate the covariance matrix and there were many small values in Y in the practical experiment. This calculation would lead to reduced accuracy.

(2) If the involved MA was too strong, it would become the principal component instead of PPG. We defined this type of MA as a heavy MA, which requires further processing below.

To solve the first problem, SVD was chosen to replace standard PCA. This Y matrix was decomposed into three parts: $Y = U\Sigma V^{\mathrm{T}}$ by SVD. In this step, the IMFs from the same time slice were under processing instead of different time slices in TROIKA's SSA. According to the dimensionality reduction function of SVD, the first several columns in the matrix U would be the principal components. Thus this $N * L$ part formed the new matrix U discussed below. These columns were also decorrelated. In this way, mode mixing and the mutual noise in each IMF caused by MA were further removed [14].

As for the second problem, we chose VCS once again as a rapid evaluation of each IMF in stationarity. In this step, the VCS of each column in U was calculated. Noting that VCS contains 4 series in all, each series's variance was calculated as follows, denoted by $\{\lambda_{uv}, \lambda_{lv}, \lambda_{ud}, \lambda_{ld}\}$, u denoted upper, l denoted lower, v denoted value and d denoted distance. Each abnormal high value in this set indicated a corruption in the corresponding IMF. The first $[N/3]$ columns in the matrix U were under testing by the following criterion:

$$\lambda_i \in [\lambda_{i,\mathrm{ave}} * \{1-\omega\}, \lambda_{i,\mathrm{ave}} * \{1+\omega\}], \quad i = uv, lv, ud, ld, \tag{7}$$

where ω was a user-defined parameter and was set to 0.3 and ave stood for the average value calculated by all N columns, not just the first $[N/3]$ columns ([ ] means to round down to the nearest whole unit). If there was one λ that did not meet this criterion, the relevant slice would be regarded as the occurrence of heavy MA. Otherwise, the first $[N/2]$ columns would be added up to form the reconstructed signal as shown in Figure 3. Based on the extraction of statistical features, we had proposed a novel high-efficiency and universal ranking method for MAs. This would also help us reduce the general cost of computing.

3.3 Two-dimension filtration

To confront the adaptivity of EMD and the randomness of MA, a novel adaptive two-dimension filtration was first proposed in this step. This filtration took the third-order cumulant as an important measurement for matrix U's columns in cyclicality [15]. Remarking that the third-order cumulant for a Gaussian process is zero, those columns need to be preprocessed at first [9]. Consider the following processing:

$$S(t) = \mathrm{IMF}(t) + N(t) * \epsilon, \tag{8}$$

in which $S(t)$ was the observed signal, $\mathrm{IMF}(t)$ in this case was the PPG components with MA non-Gaussian portion, $N(t)$ stood for the Gaussian noise portion and ϵ denoted the noise level. Note that the outputs from EMD satisfy the Gaussian distribution (GD), hence, we had a basis for assuming that ϵ was large enough to make $S(t)$ a Gaussian distributed vector. The wavelet transform (WT) was chosen. Through the wavelet transform, the modulus maxima of $S_0(t)$ were enhanced by wavelet de-noising. Make the transfer matrix of WT as W and

$$W \times S(t) = W \times \mathrm{IMF}(t) + W \times N(t) * \epsilon. \tag{9}$$

As WT is a smooth transformation to extract signals discussed in the former work [16], in which it has been proved that noises are "smoothed". This means the energy of Gaussian noise is spread equably in wavelet coefficients. WT is also an orthogonal transformation, which means W is unitary. Define P as the covariance matrix of $W \times N(t) * \epsilon$,

$$E(P) = E(W \times NN^{\mathrm{T}} \times W^{\mathrm{T}}) * \varepsilon^2 = W \times W^{\mathrm{T}} * \varepsilon^2 * E(NN^{\mathrm{T}}) = \varepsilon^2 * \delta_N^2 * I. \tag{10}$$

Obviously, the GD noise is still a Gaussian noise after transformation. However its energy is dispersed, i.e., the GD portion would be weakened. As non-GD parts have no contributions to GD parts, the output $W \times S(t)$ would no longer be subject to GD.

Harr wavelet transform (HWT) was chosen. The 3 dB bandwidth of it at scale 2^5 is from 1.75 Hz to 4.75 Hz with a sample rate of 125 Hz, which was sufficient for PPG analysis and HR estimation [3, 17].

Daubechies wavelet seems to be more proper because its frequency response is low-pass, but the calculation would be more complex. Harr wavelet has been validated to be feasible for this framework by test.

The distribution of PPG in each IMF might be quite different, even in length as shown in Figure 2(c). Consequently, our next goal was to locate them and join them together. The detailed filtration strategy consists of four steps as follows.

(1) Apply HWT at scale 2^5 on each row of matrix U^{T}.

(2) Use VCS to part the coefficient series into several slices, denoted by

$$U^{\mathrm{T}} = \begin{bmatrix} \xi_{11}, \xi_{12}, \ldots, \xi_{1r_1} \\ \xi_{21}, \xi_{22}, \ldots, \xi_{2r_2} \\ \cdots \end{bmatrix}, \tag{11}$$

where r_1, r_2 stand for the numbers of slices in each row.

(3) For each slices, calculate the 3rd order cumulants, denoted by

$$\begin{bmatrix} \gamma_{11}, \gamma_{12}, \ldots, \gamma_{1r_1} \\ \gamma_{21}, \gamma_{22}, \ldots, \gamma_{2r_2} \\ \cdots \end{bmatrix}, \tag{12}$$

and then the whole series', denoted by

$$\begin{bmatrix} \tau_1 \\ \tau_2 \\ \cdots \end{bmatrix}. \tag{13}$$

(4) For each slice, defined

$$\varphi_{ij} = \ln\left(\frac{1 + |\gamma_{ij}/\tau_i|}{|\gamma_{ij}/\tau_i|}\right), \tag{14}$$

and got the adjudication matrix

$$S = \begin{bmatrix} \varphi_{11}, \varphi_{12}, \ldots, \varphi_{1r_1} \\ \varphi_{21}, \varphi_{22}, \ldots, \varphi_{2r_2} \\ \cdots \end{bmatrix}. \tag{15}$$

So far the diversities between each row had been removed and the results in matrix S were also normalized. We considered three cases.

Case 1. φ_{ij} was smaller than the threshold $\varphi_i * 3$. The φ_i was the smallest value among $\{\varphi_{i1}, \varphi_{i2}, \ldots, \varphi_{ir_i}\}$. In this case, the relevant slice in matrix U was retained. The small φ_{ij} indicated good cyclicity and stationarity of the current slice.

Case 2. φ_{ij} was larger than the threshold $\varphi_i'/3$. The φ_i' was the largest value among $\{\varphi_{i1}, \varphi_{i2}, \ldots, \varphi_{ir_i}\}$. In this case, the relevant slice in matrix U was set to zero. The large φ_{ij} indicated that the shape of the current slice was aberrant.

Case 3. There were also vague cases among the φ_{ij}, and thus, we took the mean of the adjacent φ_{ij-1} and φ_{ij+1} as the reference. Unless φ_{ij} was not larger than this mean value, this slice would be retained.

After this filtration strategy, the retained slices in matrix U was added up to be the reconstructed signal. This 2-D filtration is both low-complex in computing and precise in time domain.

3.4 Heartrate estimation

After all the above steps, the reconstructed slices were stitched back to constitute the reformed PPG signal. Because the slicing in Subsection 3.1 cut the slices from the nearest minimums, there was no need to do further processing for recombination. Considering that even the reconstructed signal could contain some devastations because of the MAs' high intensity, VCS was applied again to find such slices and abandon them, using the same method as in Subsection 3.1. Then, FFT was applied to those slices

whose the length was not under 3 s. In addition, the peak frequencies among 0.4 to 5 Hz were chosen for HR estimation, by

$$\mathrm{HR}_{\mathrm{cal}} = 60 * \sum f_{\mathrm{peak},k}/k, \tag{16}$$

where k stands for the number of peak frequencies in the current frame. The final result is the estimation for HR in the current frame.

Tips. The threshold for VCS processing was defined in Subsection 2.2 and the updating rule was mentioned. In addition, there was a parameter defined as ratio ϕ, which was the ratio between detection boundaries and referenced thresholds $\{\epsilon, \mathrm{dis}\}$, i.e., the upper boundaries were $\{\epsilon, \mathrm{dis} * 0.1\} * \phi$, and the lower were $\{\epsilon, \mathrm{dis} * 0.1\}/\phi$. The experiment for setting the $\phi's$ value will be discussed in Section 4.

4 Experimental results and discussion

4.1 Dataset

The proposed algorithm was tested under the multiparameter intelligent monitoring in intensive care (MIMIC II) database. This database is a part of the PhysioNet archive of biomedical signal databases. The physiologic signals and vital signs time series in this database are all captured from patient monitors, as well as comprehensive clinical data obtained from hospital medical information systems. In other words, the data used for test were all real data for medical diagnosis. The sampling rate of the PPG signals we used for the analysis is 125 Hz, and there is a real time ECGV signal for algorithm evaluation along with it. The whole length of signals for testing is 120 min in 24 slices and they are selected to have different types of corruptions. The test was divided into two parts in Subsections 4.2 and 4.3. In Subsection 4.2, the setting of ratio ϕ for VCS detection was discussed and the algorithm performance tests were in Subsection 4.3.

4.2 Ratio ϕ

The ratio ϕ was defined in the former section and it had two influences on estimation accuracy. On one hand, the value of ϕ determined the precision of corrupted slice detection. The higher its value was, the fewer corruptions would be classified and further processed. On the other hand, if the ratio was too small, more corrupted slices would be involved by proposed 2-D filtration, leading to a deteriorative result. Hence, two values of ϕ were required, i.e., ϕ_1 for corrupted slice detection and ϕ_2 for 2-D filtration. At first, set $\phi_1 = \phi_2$, and we made this value range from 12 to 22 with a step size of 1. Twelve samples of the signal were under tested in this step. The length of each sample was 5 min and each sample was divided into 10 frames. The samples were chosen to contain types of MAs. The absolute average error (AAE) between the calculated heartrate ($\mathrm{HR}_{\mathrm{cal}}$) and the exact value ($\mathrm{HR}_{\mathrm{ref}}$) from ECGV signal was chosen for evaluation:

$$\mathrm{AAE} = \frac{1}{N}\sum_{i=1} |\mathrm{HR}_{\mathrm{cal}} - \mathrm{HR}_{\mathrm{ref}}|_i. \tag{17}$$

The result is shown in Figure 4. The line chart shows the valley of AAE is approximately 18, which means that the accuracy is highest here.

Then, we set $\phi_1 = 18$ and made ϕ_2 vary from 18 to 26 with stepsize 1. The result is show in Figure 5(a), AAE stops decreasing significantly after 21. Then, we set $\phi_2 = 21$ and varied ϕ_1 from 10 to 17. The result in Figure 5(b) shows that the value of ϕ_1 should be 14.

4.3 Performance measurement and results

The standard deviation (SD) was also calculated over all the frames to evaluate the stability of the proposed algorithm. We noted that the SDs were larger than the averages, and the results seem not to content the normal distribution. Therefore, he boxplot was also used for assessment. EEMD-PCA [10] and our method without 2-D filtration were compared this time using the same dataset. The results in

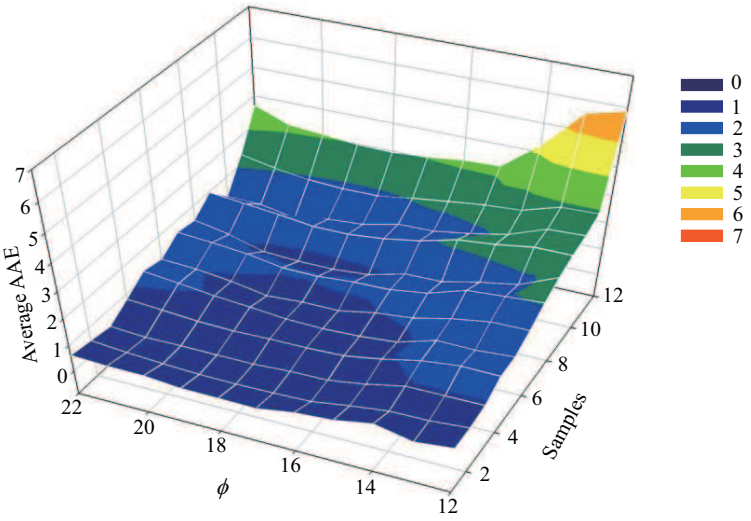

Figure 4 (Color online) AAE varies along ratio ϕ changing on 12 samples.

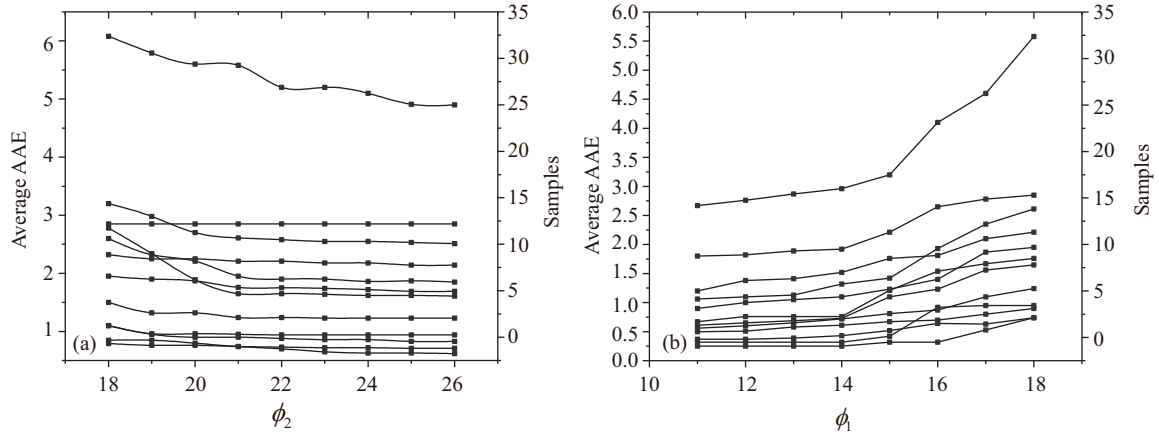

Figure 5 AAE varies along ratio (a) ϕ_2 and (b) ϕ_1 changing on 12 samples.

Figure 6 show that EEMD-PCA narrowly wins the proposed method without 2-D filtration, while the proposed method is outstanding in both average value and deviation distribution of precision. Additionally, the accuracy of the proposed method per frame is illustrated in Figure 7, in which the estimated HR ($HR_{estimation}$) and the reference HR are both plotted. Table 1 also shows the 1st and 3rd quartiles of the AAEs, which further confirms that this proposed framework is much more stable.

To evaluate our strategy's function of HR estimation, it was compared with those popular algorithms. TROIKA [4], MISPT [5], JOSS [18] algorithms were chosen and the result is in Table 2. Compared with the other three, the dataset we chose is more authoritative.

4.4 Discussion

In this study, we proposed a novel approach for HR estimation against heavy MA corruptions. The accuracy was validated using MIMIC datasets. This method consists of VCS, EMD, SVD and a novel 2-D filtration strategy. EMD is an adaptive decomposition; hence, there is no concern regarding its effect. However, the IMF itself and the MAs' effects on IMF are both random. This is the main difficulty we met in the EMD applications. In this proposed method, we avoided setting many parameters. Instead, we chose the signal itself as a reference by VCS filtration. The stationarity of each slice is the main criterion we needed. The slices were compared with each other, rather than configured thresholds. In this way, the loss of most of the information from original inputs is avoided. The novel 2-D filtration is also based on the

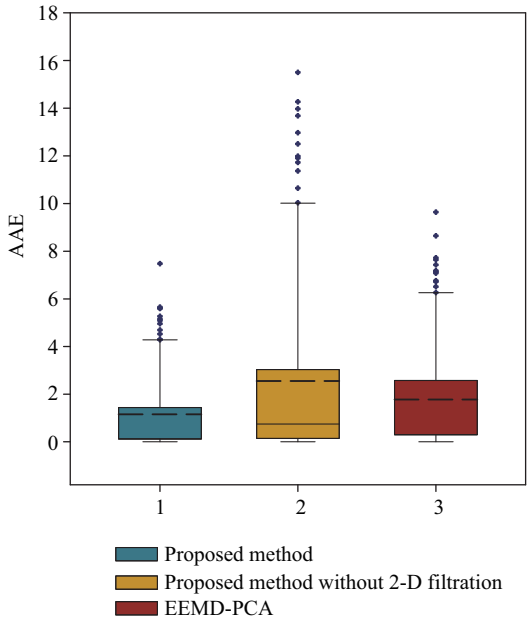

Figure 6 (Color online) AAE of EEMD-PCA, proposed methods and that without 2-D filtration shown in boxplot view.

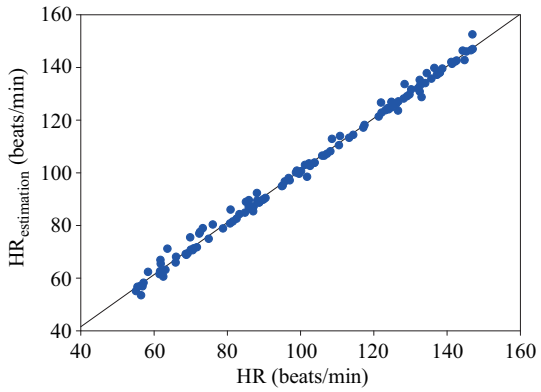

Figure 7 (Color online) Pearson correlation between referenced HR and proposed method based PPG derived HR.

Table 1 Average AAE and the 1st and 3rd quartiles of the three EMD-based methods

Method	Average AAE	1st quartile	3rd quartile
EEMD-PCA [10]	1.78	0.32	2.56
Proposed method without 2-D filtration	2.56	0.14	3.02
Proposed method	**1.07**	**0.12**	**1.41**

Table 2 AAE (in beat per minute) for each dataset using the proposed method as well as using MISPT, JOSS and TROIKA. Average (SD) is calculated by the test results of each slice

Strategy	Principal methods	Dataset	Average AAE	SD
TROIKA [4]	SSA, SSR, SPT	Experimental data	2.34	–
MISPT [5]	SSA, SSR, SPT	Experimental data	1.11	2.33
JOSS [18]	Joint SSR	Experimental data	1.28	2.61
Proposed method	EMD, VCS, HWT, 2-D filtration	MIMIC II	**1.07**	**1.87**

signals' statistical characteristics, and therefore, it has the ability to contend against the nondeterminacy of both EMD and MAs. Noise is removed by EMD first and then is further reduced by SVD. However, we must extract PPG while reserving the details maximally in heavily corrupted situations. Frequency analysis methods were given up because they all strictly rely on the length of inputs and their precisions cannot be guaranteed.

There are several user-defined parameters in this framework. However, all these values for the threshold calculation are proportionality coefficients, and the eventual thresholds would be based on statistical parameters extracted from the inputs. This would insure that the proposed framework has excellent universality. With the help of VCS for corruption detection and the MA ranking method, the computing cost is decreased. This guarantees that the framework is also much efficient.

An experimental parameter denoted by ratio ϕ was used for partitions via VCS. There were two ratios, ϕ_1 and ϕ_2. ϕ_1 was for corruption detection, whereas ϕ_2 was used in 2-D filtration. The two-step approach of their values is resented in Subsection 4.2. The smaller the value of ϕ_1 is, the better result we gain. The computing cost of entire processing can be reduced by large ϕ_1. Similarly, ϕ_2 cannot be too small because all of the noise slices should be weeded out.

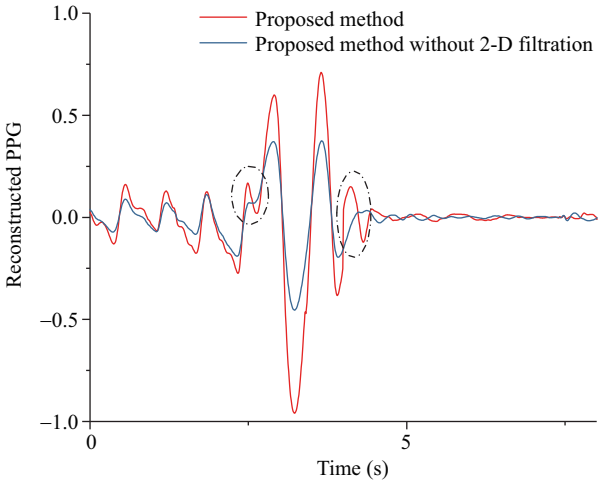

Figure 8 (Color online) Effect of 2-D filtration strategy in PPG reconstruction.

The test results in Table 2 show that the proposed method is both more precise and more stable than the other three. The average AAE is 1.27, 0.04 and 0.21 smaller than the compared algorithms, respectively. The SD 1.87 also presents our strategy functions pretty well against types of corruptions. Then, a boxplot is used to evaluate this framework because the results do not satisfy a normal distribution. Figure 6 reveals that the 2-D filtration is useful against heavy corruptions. There are few abnormal values in proposed strategy than EEMD-PCA, which means that this algorithm is more stable against variations in data. From the holistic perspective, the AAE of the proposed method and median error are much lower than the EEMD-PCA method and the method without the novel filtration. Hence, our method also has an advantage in precision. As shown Figure 8, the peaks surrounded by dotted lines are related to heartbeats, which are recovered by the novel 2-D filtration strategy. These peaks largely improve the accuracy of this proposed framework.

5 Conclusion

With the novel detection method based on VCS, the corrupted periods are selected from the whole input, and thus, the computing cost would be reduced. Then, SVD helps weaken the mode mixing and heavily corrupted slices are identified for further processing. In the third part, HWT and higher-order statistics analysis are combined to form a novel 2-D filtration to select periods related to PPG in each decorrelated vector. FFT is applied to the reconstructed signals to estimate the heartrate. The test results (average AAE of 1.07 and SD of 1.87) show that our method has an accurate and highly stable performance against MA corruptions. This proposed method uses statistical characteristics calculated from the original signal as dynamic criteria; thus, the majority of the stationary parts can be remained. In conclusion, this paper provides a highly feasible framework of PPG processing based on EMD against the randomness caused by both MA and EMD itself.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61634006, 61372060, 61335010, 61474107, 81300803), National Key Technologies R&D Program (Grant Nos. 2016YFB0401303, 2016YFB0402405), Basic Research Project of Shanghai Science and Technology Commission (Grant No. 16JC1400101), and Key Research Program of Frontier Science, Chinese Academy of Sciences (Grant No. QYZDY-SSW-JSC004).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Selvaraj N, Jaryal A, Santhosh J, et al. Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography. J Med Eng Technol, 2008, 32: 479–484
2 Petersen C L, Chen T P, Ansermino J M, et al. Design and evaluation of a low-cost smartphone pulse oximeter. Sensors, 2013, 13: 16882–16893
3 Lee J, Jung W. Design the filter to reject motion artifact of pulse oximetry. Elsevier Comput Sci, 2003, 26: 241–249
4 Zhang Z, Pi Z, Liu B. TROIKA: a general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise. IEEE Trans Biomed Eng, 2015, 62: 522–531
5 Murthy N K L, Madhusudana P C, Suresha P, et al. Multiple spectral peak tracking for heart rate monitoring from photoplethysmography signal during intensive physical exercise. IEEE Signal Process Lett, 2015, 22: 2391–2395
6 Komaty A, Boudraa A O, Augier B, et al. EMD-based filtering using similarity measure between probability density functions of IMFs. IEEE Trans Instrum Meas, 2014, 63: 27–34
7 Yang G, Liu Y, Wang Y, et al. EMD interval thresholding denoising based on similarity measure to select relevant modes. Signal Process, 2015, 109: 95–109
8 Sung P, Syed Z, Guttag J. Quantifying morphology changes in time series data with skew. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2009, Taipei, 2009. 477–480
9 Wu Z, Huang N E. Ensemble empirical mode decomposition: a noise-assisted data analysis method. Adv Adapt Data Anal, 2009, 1: 1–41
10 Motin M A, Karmakar C, Palaniswami M. Ensemble empirical mode decomposition with principal component analysis: a novel approach for extracting respiratory rate and heart rate from photoplethysmographic signal. IEEE J Biomed Health Inf, 2017. doi: 10.1109/JBHI.2017.2679108
11 Pang B, Liu M, Zhang X, et al. Advanced EMD method using variance characterization for PPG with motion artifact. In: Proceedings of 2016 IEEE Biomedical Circuits and Systems Conference (BioCAS), Shanghai, 2016. 196–199
12 Huang N E, Shen Z, Long S R, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc Math Phys Eng Sci, 1998, 454: 903–995
13 Li P, Liu M, Zhang X, et al. A low-complexity ECG processing algorithm based on the Haar wavelet transform for portable health-care devices. Sci China Inf Sci, 2014, 57: 122303
14 Gao Y, Ge G, Sheng Z, et al. Analysis and solution to the mode mixing phenomenon in EMD. Congress Image Signal Process, 2008, 5: 223–227
15 Dimitrakopoulos R, Mustapha H, Gloaguen E. High-order statistics of spatial random fields: exploring spatial cumulants for modeling complex non-Gaussian and non-linear phenomena. Math Geosci, 2010, 42: 65–99
16 Donoho D L. De-noising by soft-thresholding. IEEE Trans Inf Theory, 1995, 41: 613–627
17 Li X, Zhang X, Li P, et al. An 8.12 W wavelet denoising chip for PPG detection and portable heart rate monitoring in 0.18 μm CMOS. J Semicond, 2016, 37: 055006
18 Zhang Z. Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction. IEEE Trans Biomed Eng, 2015, 62: 1902–1910