Combining Nonlinear Adaptive Filtering and Signal Decomposition for Motion Artifact Removal in Wearable Photoplethysmography

Yalan Ye, Yunfei Cheng, Wenwen He, Mengshu Hou, *Member, IEEE*, and Zhilin Zhang, *Senior Member, IEEE*

***Abstract*— Heart rate (HR) estimation using photoplethysmography (PPG) has drawn increasing attention in the field of wearable technology due to its advantages with higher degree of usability and lower cost than Electrocardiograph. It has been widely used in wearable devices, such as smart-watches for fitness tracking and vital sign monitoring. However, motion artifact is a strong interference, preventing accurate estimation of HR. Signal decomposition and adaptive filtering are two popular approaches for motion artifact removal, but each of them has inherent drawbacks. In this paper, a hybrid motion artifact removal method is proposed, which combines nonlinear adaptive filtering and signal decomposition, getting the best of both approaches. The method was evaluated on the PPG database used in the 2015 IEEE Signal Processing Cup. The experimental results showed that the method achieved the average absolute error of 1.16 beat per minutes (BPM) on the 12 training data sets, and 2.98 BPM on the ten testing data sets.**

***Index Terms*— Heart rate (HR), photoplethysmography (PPG), motion artifacts (MA), recursive least squares (RLS), nonlinear adaptive filtering, signal decomposition.**

I. Introduction

With people's growing demands for healthcare and fitness, there has been an increasing trend to research heart rate (HR) estimation [1], [2] to prevent heart troubles and control training load during exercise because HR is an important index to indicate exercise intensity. Over-higher heart rate value, namely over-greater exercise intensity, indicates that the exercisers are at risk for health problems [3]. Therefore, it is important to estimate HR especially during exercise in order to make exercisers' heart work at its optimal intensity.

Nowadays, two main techniques for HR estimation during exercise have been put forward. One is the conventional electrical technology Electrocardiograph (ECG), which has been used for HR estimation for a long time [4]. However, ECG is not convenient to use especially during movement because it always needs several electrodes tied to the body so as to capture the electrical signals. Another method is photoplethysmography (PPG) [5]. Since rhythms of PPG signals correspond to heartbeats, estimated HR from PPG signals can be achieved by detecting instantaneous fundamental frequency of PPG signals [6], [7]. Moreover, due to higher usability and lower cost of PPG sensors compared with ECG, PPG is widely used in modern wearable devices such as smart-watches and smart wristbands. However, voluntary or involuntary subject movement would then invariably disturb the contact between the sensor and the skin, corrupting the PPG signal with motion artifacts [8], [9], resulting in erroneous interpretation of PPG signals and degrading the accuracy and reliability of PPG-based algorithms for the estimation of HR. Consequently, removing MA effectively is a big challenge (Fig.1).

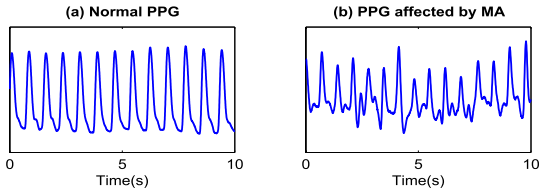

Fig. 1. The Photoplethysmography signals. (a) is the normal PPG signals. (b) is the PPG signals affected by motion artifacts during movement.

Many techniques were proposed to remove motion artifacts. One widely used technique is independent component analysis (ICA) [10]. However, ICA has a key assumption of statistical independence or uncorrelation, which corrupted PPG signals do not hold [11]. Thus, the result of separating MA from PPG signals is not satisfactory in many practical fitness tracking scenarios.

Adaptive filtering [12]–[16] is another widely used technique.With suitably designed reference signals, it can achieve satisfactory results in some cases. However, a poorly designed reference signal can cause poor MA removal performance. Besides, the relation between MA in raw PPG signals and a reference signal may not be linearly correlated, but nonlinearly correlated, which is often encountered when motions are complicated while simultaneous acceleration signals are chosen as reference signals.

Signal decomposition [17], [18] is recently shown to be a powerful method to remove MA. For example, in [17], the

Manuscript received September 2, 2015; revised July 18, 2016; accepted July 18, 2016. Date of publication August 2, 2016; date of current version September 1, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 61501096 and Grant 61472067, in part by the Chengdu Research Institute of UESTC under Grant RWS-CYHKF-02-20150005, in part by the International Science and Technology Cooperation and Exchange Program of Sichuan Province, China, under Grant 2016HH0020, and in part by the Fundamental Research Funds for the Central Universities. The associate editor coordinating the review of this paper and approving it for publication was Dr. Akshya Swain. *(Corresponding authors: Yalan Ye; Zhilin Zhang.)*

Y. Ye, Y. Cheng, W. He, and M. Hou are with the School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: yalanye@uestc.edu.cn; yunfeicheng@hotmail.com; hww1015@foxmail.com; mshou@uestc.edu.cn).

Z. Zhang is with Samsung Research America–Dallas, Richardson TX 75082 USA, and also with the School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: zhilinzhang@ieee.org).

Digital Object Identifier 10.1109/JSEN.2016.2597265

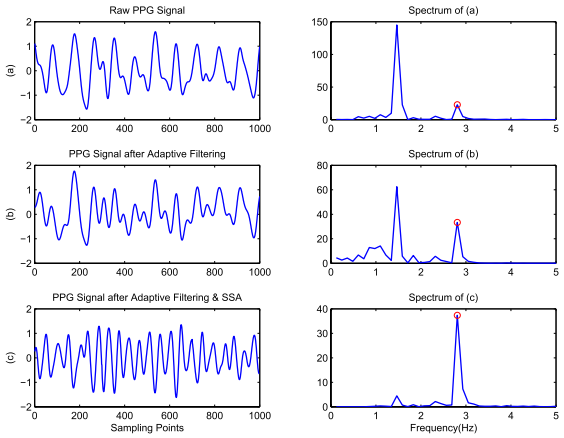

Fig. 2. Comparison of only nonlinear adaptive filtering, and RLS Volterra-based nonlinear adaptive filtering and SSA: (a) is the raw PPG signal, (b) is the signal after the nonlinear adaptive filtering, and (c) is the signal afer the nonlinear adaptive filtering and SSA, and red circle indicates the spectral peak associated with heart rate. From the example, we can see that the combination of two method can make the peak corresponding to HR more discernable than only using nonlinear adaptive filtering.

authors used singular spectrum analysis (SSA) to decompose a raw PPG into a number of components, and then used information from simultaneous acceleration signals to identify the components associated with MA. After removing these components and reconstructing the PPG signal using remained components, a cleaner PPG signal was obtained. Similarly, in [18], empirical mode decomposition (EMD) was used to decompose a raw PPG signal into many components, and spectrum subtraction was used to remove the components associated with MA using the spectra of simultaneous acceleration signals. However, signal decomposition generally has large computational load, which may prevent its use in low-power wearable devices.

In this study we propose a hybrid method combining adaptive filtering and signal decomposition, getting the best of the two approaches while overcoming their drawbacks. In particular, we first use the second order recursive least squares (RLS) Volterra-based nonlinear adaptive, a nonlinear adaptive filter, to perform adaptive noise cancelation, with simultaneous acceleration signals chosen as reference signals. The nonlinear adaptive filter aims to better capture the nonlinear relation between MA in raw PPG signals and simultaneous acceleration signals, thus improving MA removal effect. However, we do not solely rely on it, since the nonlinear adaptive filter may not achieve satisfactory results in some cases, illustrated by Fig.2. Thus, signal decomposition is applied on the output of the nonlinear adaptive filter, removing the remained MA components in the raw PPG signal. Experiments with three datasets were performed to demonstrate the efficacy of our proposed method.

The rest of the paper is organized as follows. Section II presents our proposed algorithm. Section III first describes three PPG datasets, and then analyses the experimental results on three datasets in detail. This section further validates that our proposed algorithm is of high accuracy. Conclusions are given in the last section.

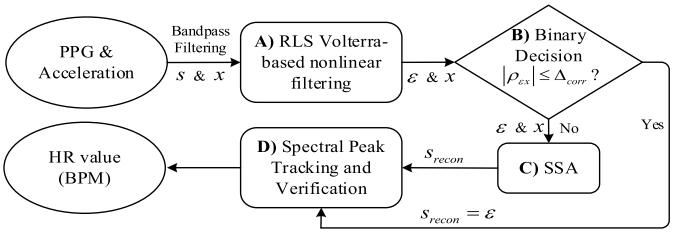

Fig. 3. The flowchart of our proposed algorithm.

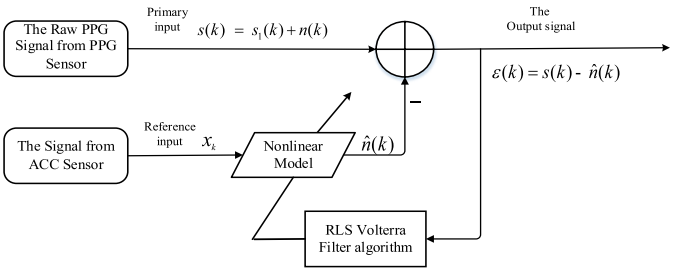

Fig. 4. The block diagram of RLS Volterra-based nonlinear adaptive filtering system. It has two inputs. One is the raw PPG signal $s(k)$, containing the signal of interest $s_1(k)$ and plus interference $n(k)$. The other input is the acceleration data x_k. The noise is filtered to produce an output $\hat{n}(k)$ that is as close a replica as possible of $n(k)$. This output is subtracted from the primary input $s(k)$ to produce the system output $\varepsilon(k)$.

II. Proposed Method

The flowchart of our proposed HR estimation algorithm is shown Fig.3. Raw PPG signals s and simultaneous acceleration signals x are input to the nonlinear adaptive filter, followed by SSA. Then binary decision would be made to decide whether the output of the nonlinear adaptive filter is clean in order to determine whether to use SSA to further reduce MA. Next, the spectral peak tracking with verification algorithm proposed in [17] is adopted to find the spectral peak corresponding to HR.

As in [17], a time window with T seconds is sliding on the signals with incremental step of M seconds (generally $M \leq T/2$) and the algorithm estimates HR in each time window. Since 0.4 Hz and 5 Hz are the frequency band we are interested in, all raw recordings including the PPG signals and the acceleration signals are band-pass filtered from 0.4 Hz to 5 Hz before the proposed algorithm starts.

A. Second Order RLS Volterra-Based Nonlinear Adaptive Filtering

This step adopts nonlinear adaptive filtering to suppress strong noise. In solving nonlinear problems, many kinds of nonlinear adaptive filtering methods have been established, such as the neural network method, bilinear filters, Volterra filter. Among these methods, Volterra filter is very useful for nonlinear system, since it can be computed using traditional signal processing algorithms in the same way as linear models [19]. In particular, recursive least-squares (RLS) Volterra algorithm can achieve fast convergence speed, which can bring our nonlinear system great convenience. Thus in this paper, RLS Volterra-based algorithm was used. Fig.4 gives the block diagram of the second order RLS Volterra-based nonlinear adaptive filtering system. The problem considered in

the step is that of finding the RLS solution for the linear and quadratic coefficients of the adaptive filtering that minimize the cost function

$$J(t) = \sum_{k=0}^{t} \lambda^{t-k} \varepsilon^2 = \sum_{k=0}^{t} \lambda^{t-k} (s(k) - \hat{n}(k))^2 \tag{1}$$

at each time instant t, where ε is the output error, λ is the forgetting factor, $k = 0, 1, \cdots, t$ and t represents the time. Here, the nonlinearity is modeled using a second order Volterra series expansion, thus $\hat{n}(k)$ is obtained as

$$\begin{aligned} \hat{n}(k) &= \sum_{i=0}^{M-1} \hat{a}_i(t) x(k-i) + \sum_{i=0}^{M-1} \sum_{j=i}^{M-1} \hat{b}_{i,j}(t) x(k-i) x(k-j) \\ &= \sum_{k=0}^{t} w_t^T x_k, \end{aligned} \tag{2}$$

where M is the number of delays involved, $x(t)$ and $\hat{n}(k)$ are the input and output sequences, $\hat{a}_i$ and $\hat{b}_{i,j} (i, j = 0, 1, \ldots, M-1)$ are possibly time-varying linear and quadratic coefficients of the nonlinear filter, and w_t is the coefficient vector.

By differentiating $J(t)$ with respect to w_t and equating the result to zero, the optimal vector w_t can be given by

$$w_{t,opt} = R_t^{-1} P_t, \tag{3}$$

where $R_t = \sum_{k=0}^{t} \lambda^{t-k} x_k x_k^T$ and $P_t = \sum_{k=0}^{t} \lambda^{t-k} x_k n(k)$. Then we can obtain the output vector of adaptive filter, namely, estimated noise vector $\hat{n}(k)$, by Eq.(2). And the PPG signal after filtering can be calculated by

$$\varepsilon(k) = s(k) - \hat{n}(k) = s(k) - w_{t,opt} x_k, \quad k = 0, 1, \cdots, t. \tag{4}$$

B. Binary Decision

In our method, for each time window, instead of always using both the nonlinear adaptive filter and SSA, we are based on the strength of the Pearson correlation between nonlinear adaptive filtered PPG signals ε and the acceleration signals x, to decide whether to use SSA to further reduce the MA. That is, here we measure whether the Pearson correlation value $\rho_{\varepsilon x}$ satisfies $|\rho_{\varepsilon x}| \leq \Delta_{corr}$, where Δ_{corr} is a preset threshold.

In fact, if the nonlinear adaptive filter does not yield a satisfactory result, which indicates that there still exists MA in the filtered PPG signals ε, the correlation between the two signals is still strong and then $|\rho_{\varepsilon x}| \leq \Delta_{corr}$ would not be satisfied. Then the nonlinear adaptive filtered PPG signal ε and the simultaneous acceleration signals x will be processed by SSA as stated in the next subsection. However, if the nonlinear adaptive filter yields a satisfactory result indicating that the correlation is weak and then the filtered PPG signals are almost clean, then there is no need to use SSA. Thus, calculating the Pearson correlation is important, which determines whether to use SSA to further reduce MA.

C. SSA

Periodogram is first used to get the spectrum of each channel of acceleration data x, from which we determine the dominant frequencies with amplitude larger than 50% of the maximum amplitude. Denote by L_{acc} the set of location indexes of selected dominant frequencies in spectra. Then the output signal ε of the first step is decomposed into some time series $\tilde{\varepsilon}_p$ by SSA [17], [20], [21]. Next, Periodogram is used again to calculate the spectrum of every time series, from which we select the dominant frequency with the maximum amplitude. Then we remove the noise time series whose dominant frequency has location indexes in L_{acc} [6]. The remained time series are used to reconstruct a cleansed PPG signal s_{recon}.

D. Spectral Peak Tracking and Verification

After the original raw PPG signal was processed by the nonlinear adaptive filter or SSA, the spectral peak tracking/verification algorithm in [17] is used to select the spectral peak corresponding to HR.

III. Experimental Results

A. Datasets

We evaluated the proposed algorithm based on the PPG database used in 2015 IEEE Signal Processing Cup.[1] The database consists of 12 training datasets and 10 testing datasets. In each set, there are two-channel PPG signals, tri-axis acceleration signals and one-channel ECG signal, recorded simultaneously from a subject. The two-channel PPG signals were recorded from the wrist by two pulse oximeters with green LEDs (wavelength: 515 nm). The acceleration signals were also recorded from the wrist by a tri-axis accelerometer. And the ECG signal was recorded from the chest using wet ECG sensors.

The 12 training datasets were initially used in [17], which were recorded during subjects' walking, jogging, and running on a treadmill. These subjects first walked with a low speed of 1-2 km/h for 0.5 minute, then ran with the speed of 6-8 km/h for 1 minute, next speeded up to 12-15 km/h for 1 minute, then slowed down to 6-8 km/h for 1 minute, next speeded up to 12-15 km/h for 1 minute, and finally changed to walk with the speed of 1-2 km/h for 0.5 minute.

The 10 testing datasets were recorded when subjects performed many actions including various forearm and upper arm exercises, running, jump, and push-up where MA is much stronger than the MA in the 12 training datasets.

In addition, we recorded extra 8 datasets using the same equipment in [17], but this time the wearers' hands performed more intensive and non-rhythmic movements, such as cooking and dancing. (Note that during recording the previous 12 training datasets and 10 testing datasets, the wearers' hands generally performed rhythmic movements.) So, these 8 datasets are more challenging to MA reduction algorithms.

[1] http://www.signalprocessingsociety.org/spcup2015/index.html

TABLE I

HR ESTIMATION RESULTS IN TERMS OF ERROR1 WHEN USING 12 PPG DATASETS WERE SHOWN IN THIS TABLE. IN ORDER TO PRESENT THE EFFICIENCY FOR HR ESTIMATION DURING MOVEMENT WHEN THE NONLINEAR RLS VOLTERRA ADAPTIVE FILTERING AND SSA WERE BOTH USED, THE FIRST STAGE WAS REPLACED BY RLS BASED ADAPTIVE FILTERING ALGORITHM AND SSA, ONLY BY RLS BASED NONLINEAR ADAPTIVE FILTERING, ONLY BY SSA, FASTICA, PEARSONICA, AND SOBI. AND THE UNIT WAS BEAT PER MINUTE (BPM).
ERROR1

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	Subj 11	Subj 12	Average
Proposed	1.66	1.56	0.65	1.48	0.77	1.12	0.72	0.91	0.42	2.35	1.45	0.78	1.16
TROIKA [17]	2.87	2.75	1.91	2.25	1.69	3.16	1.72	1.83	1.58	4.00	1.96	3.33	2.42
JOSS [23]	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
SPECTRAP [24]	1.18	2.42	0.86	1.38	0.92	1.37	1.53	0.64	0.60	3.65	0.92	1.25	1.50
RLS+SSA	2.64	1.74	0.63	1.34	0.75	1.13	0.74	0.61	0.82	13.83	0.72	2.15	2.26
NonlinearRLS	3.08	4.09	0.74	0.97	0.77	6.76	0.56	0.77	0.52	32.42	0.97	1.44	4.42
SSA	2.77	7.44	1.07	1.44	0.83	25.07	0.67	1.64	0.49	13.40	1.73	2.00	4.88
FastICA [25]	4.90	16.46	16.56	3.28	1.23	9.69	2.07	21.87	3.18	4.08	3.80	7.32	7.87
PearsonICA [26]	3.72	17.47	20.47	2.68	0.95	2.83	2.00	24.18	2.39	4.08	1.82	7.74	7.53
Sobi [27]	2.47	27.41	21.57	2.41	1.60	1.99	1.54	16.89	3.22	3.89	3.20	7.64	7.82

B. The Performance Indexes

The ground-truth of each time window is calculated from the simultaneous ECG signals, and it is now available in the PPG datasets for performance evaluation. In this paper, four performance indexes are used to evaluate the performance of our proposed algorithm.

(1) The first one is the average absolute error defined as:

$$Error1 = \frac{1}{W}\sum_{i=1}^{W}|BPM_{est}(i) - BPM_{true}(i)|. \quad (5)$$

(2) The second one is the average absolute error percentage defined as:

$$Error2 = \frac{1}{W}\sum_{i=1}^{W}\frac{|BPM_{est}(i) - BPM_{true}(i)|}{BPM_{true}(i)}. \quad (6)$$

where $BPM_{true}(i)$ represents the ground-truth of HR in the i-th time window, $BPM_{est}(i)$ denotes the estimated HR values, and W is the total number of time windows. The smaller the average absolute error (Error1) and the average absolute error percentage (Error2) are, the better estimation performance of the algorithm is.

(3) The third one is the Bland-Altman plot which is used to verify agreement between the ground-truth of HR and the estimated HR values. And the Limit of Agreement (LOA) expressed by $[\mu - 1.96\sigma, \mu + 1.96\sigma]$ is also calculated here, where μ is the average difference and σ is the standard deviation. In the Bland-Altman plot, the smaller the absolute value of difference is, the higher agreement between the ground-truth and the estimates.

(4) The last index is the Pearson Correlation Coefficient between the ground-truth and the estimates. High Pearson Correlation Coefficient indicates a good HR estimation.

C. Results

1) Experiments on the 12 training PPG datasets: In the first simulation, the first of the two-channel PPG signals of the 12 training PPG datasets recorded during exercise was used to test the performance of our proposed algorithm with the sampling rate of 125Hz. In the first stage, in order to make nonlinear adaptive filtering system work well, we set the time delay (between the corrupted PPG signal and the acceleration data when recording the signals) to 0.08s, which can make the acceleration signal x highly correlated with the noise n and help the nonlinear adaptive filtering work well [16]. Besides, the time window T was set to 8 seconds sweeping all the raw signals simultaneously, with the step of every two seconds. To show the promising performance of our proposed algorithm, we chose framework TROIKA [17] for comparison.

Table I and Table II present the average absolute error (Error1) and the average error percentage (Error2) on all 12 subjects' recordings, respectively, stored in 12 PPG datasets. Using our proposed algorithm can obtain accurate estimation of HR with the averages of 12 subjects' recordings about Error1 being 1.16 ± 2.23, and Error2 being 0.93%, respectively, which was better than several excellent algorithms proposed in the last two years, namely TROIKA [17], JOSS [22], and SPECTRAP [23]. Besides, to verify the superiority of our proposed MA removal algorithm, the first stage was replaced by RLS + SSA, only by RLS-based nonlinear adaptive filtering, only by SSA, respectively. From the two tables we can see that our proposed MA removal algorithm was better than others. The results reveal that using nonlinear adaptive filtering and SSA is better than only using the nonlinear adaptive filtering because in our experiment there exists strong Pearson correlation between the nonlinear adaptive filtered PPG signals and the acceleration data which indicates that the output of the nonlinear adaptive filter is not satisfactory. It also shows that the proposed method is better than RLS + SSA based adaptive filtering, which further proves that using our proposed MA removal method is a better choice to ensure accurate HR estimation than other MA removal method listed here.

In Table I and Table II, we also compared three ICA/BSS algorithms. ICA/BSS algorithms were used to remove MA in

TABLE II

RESULTS RELATED TO DIFFERENT MA REDUCTION METHOD IN TERMS OF ERROR2 WERE SHOW IN THE TABLE, AND THE MEANING OF EACH ROW WAS THE SAME AS TABLE I.

ERROR2

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	Subj 11	Subj 12	Average
Proposed	1.42%	1.44%	0.53%	1.51%	0.60%	0.90%	0.60%	0.80%	0.36%	1.45%	0.94%	0.60%	0.93%
TROIKA [17]	2.18%	2.37%	1.50%	2.00%	1.22%	2.51%	1.27%	1.47%	1.28%	2.49%	1.29%	2.30%	1.82%
JOSS [23]	1.19%	1.66%	1.27%	1.41%	0.51%	1.09%	0.54%	0.47%	0.41%	2.43%	0.51%	0.81%	1.01%
SPECTRAP [24]	1.04%	2.33%	0.66%	1.31%	0.74%	1.14%	1.36%	0.55%	0.52%	2.27%	0.65%	1.02%	1.12%
RLS+SSA	1.78%	2.07%	0.58%	1.27%	0.52%	1.03%	0.41%	0.90%	0.42%	4.43%	0.47%	1.86%	1.31%
nonlinearRLS	1.38%	1.46%	0.51%	1.17%	0.44%	0.96%	0.54%	0.77%	0.35%	4.16%	0.51%	0.46%	1.05%
SSA	2.31%	5.97%	0.82%	1.33%	0.63%	18.41%	0.54%	1.45%	0.41%	8.30%	1.17%	1.38%	3.26%
FastICA [25]	4.06%	13.10%	11.93%	2.87%	0.94%	7.82%	1.64%	17.39%	2.90%	2.56%	2.52%	5.00%	6.06%
PearsonICA [26]	3.16%	13.65%	14.94%	2.22%	0.73%	2.36%	1.67%	19.18%	2.14%	2.53%	1.14%	5.36%	5.76%
Sobi [27]	2.16%	21.87%	15.94%	2.17%	1.36%	1.70%	1.21%	13.06%	2.86%	2.40%	2.00%	5.28%	6.00%

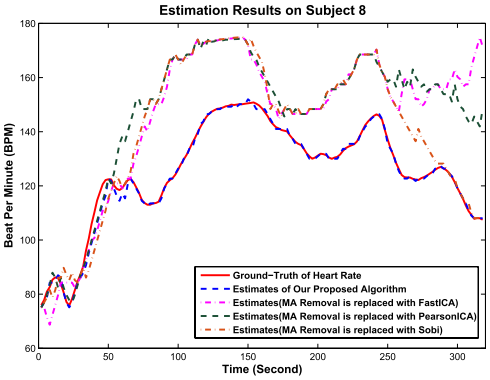

Fig. 5. Estimation results on recordings of Subject 8 of the 12 training PPG datasets. The traces of HR estimating with respect to four algorithms were plotted, i.e., using our proposed algorithm, replacing MA removal method with three independent component analysis (ICA) /blind source separation (BSS) algorithms, namely, FastICA [24], PearsonICA [25], and Sobi [26]. They were all compared to the ground-truth which was recorded simultaneously from ECG.

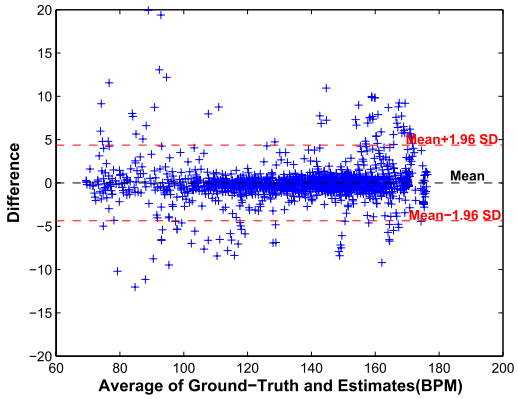

Fig. 6. The Bland-Altman plot of the estimation results of our proposed algorithm on the 12 training PPG datasets. The LOA was [-4.37,4.37] BPM (standard deviation σ = 2.23 BPM).

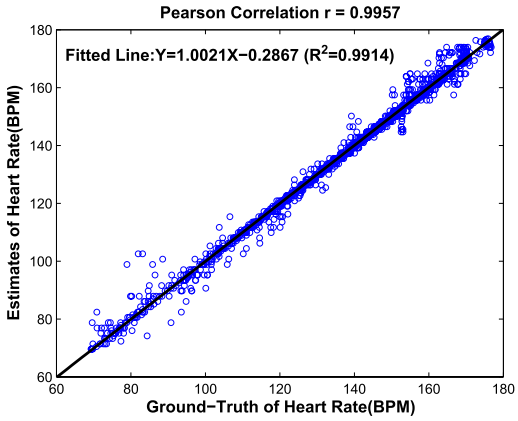

Fig. 7. Scatter plot on the 12 datasets between the ground-truth and the estimates of our proposed algorithm. The fitted line was $y = 1.0021x - 0.2876$ where x represents the ground-truth HR values, and y denotes the estimated HR values; the R^2 value was 0.9914; the Pearson Correlation Coefficient was 0.9957.

some published works. However, these algorithms were basic ICA/BSS algorithms, which cannot relax the assumption of independence/uncorrelation among underlying latent components. In this experiment, we used two basic ICA algorithms (FastICA [24], PearsonICA [25]) and one BSS algorithm (SOBI [26]). For these algorithms, both two-channel of PPG signals were used. And the detailed process of exploiting ICA/BSS algorithms to obtain the clean PPG signals was as follows. Two-channel output was first obtained by ICA/BSS algorithms, then the correlation coefficient of the output's each channel with the three-axis acceleration signals was calculated, and finally the channel signal which has small correlation with acceleration signals would be selected as the clean PPG signal. From the two tables, we can see that the HR estimation accuracy is very bad when using FastICA, PearsonICA, and Sobi. So we further confirm that both independence-required ICA algorithms and the uncorrelation-required BSS algorithm are not suitable to remove MA from the corrupted PPG signals.

In Fig.5, the HR traces of four algorithms on the recordings of subject 8 are plotted to compare with the ground-truth. From the figure we can see that, the HR trace of our proposed algorithm was closest to the ground-truth. And also we observe that the HR traces of FastICA [24], PearsonICA [25] and SOBI [26] were far from the ground-truth, showing that it failed when using FastICA, PearsonICA and Sobi in the MA removal of the first stage to estimate HR.

Then the Bland-Altman plot is illustrated in Fig.6. The LOA was [-4.37,4.37] BPM. And the Scatter Plot between

TABLE III
HR ESTIMATION RESULTS IN TERMS OF ERROR1 WHEN USING 10 PPG DATASETS WERE SHOWN IN THIS TABLE, AND THE MEANING OF EACH ROW WAS THE SAME AS TABLE I.
ERROR1

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	Average
Proposed	7.71	1.62	3.10	7.00	2.99	1.67	2.45	1.81	0.92	0.49	2.98
TROIKA [17]	6.63	1.94	1.35	7.82	2.46	1.73	3.33	3.41	2.68	0.51	3.19
JOSS [23]	8.07	1.61	3.10	7.00	2.99	1.67	2.80	1.88	0.92	0.49	3.05
SPECTRAP [24]	4.89	1.58	1.83	3.05	1.62	1.24	2.04	2.49	1.16	0.66	2.13
RLS+SSA	20.76	4.30	2.30	9.29	3.60	2.03	2.94	5.48	2.98	0.75	5.44
nonlinearRLS	47.96	13.39	2.94	25.77	4.36	1.46	3.53	2.82	1.55	0.49	10.43
SSA	40.92	2.38	32.15	17.10	4.88	1.96	3.68	3.04	1.28	0.54	10.81
FastICA [25]	47.73	11.25	15.00	4.64	17.34	9.53	4.83	10.83	11.13	1.48	13.38
PearsonICA [26]	62.65	46.58	6.31	4.91	15.54	24.56	24.97	15.31	11.79	1.03	21.37
Sobi [27]	51.49	16.94	8.99	5.45	16.74	10.68	6.79	11.23	6.21	1.60	13.61

TABLE IV
HR ESTIMATION RESULTS IN TERMS OF ERROR2 WHEN USING 10 PPG DATASETS WERE SHOWN IN THIS TABLE, AND THE MEANING OF EACH ROW WAS THE SAME AS TABLE II.
ERROR2

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8	Subj 9	Subj 10	Average
Proposed	10.58%	2.02%	2.68%	4.49%	2.52%	1.23%	3.00%	1.26%	0.74%	0.57%	2.91%
TROIKA [17]	8.76%	2.56%	1.04%	4.88%	2.00%	1.27%	3.90%	2.43%	2.12%	0.59%	2.96%
JOSS [23]	10.97%	2.01%	2.69%	4.49%	2.52%	1.23%	3.46%	1.32%	0.74%	0.57%	2.99%
SPECTRAP [24]	6.29%	1.98%	1.49%	2.0%	1.36%	0.92%	2.23%	1.81%	0.92%	0.79%	2.04%
RLS+SSA	18.34%	2.07%	2.79%	8.66%	5.94%	0.97%	2.77%	1.29%	0.96%	0.58%	4.44%
nonlinearRLS	13.29%	2.73%	2.00%	11.04%	2.54%	1.03%	2.16%	0.89%	0.87%	0.59%	3.71%
SSA	55.66%	3.04%	23.27%	10.92%	3.93%	1.45%	4.11%	2.21%	1.01%	0.63%	10.62%
FastICA [25]	63.95%	14.26%	11.53%	2.98%	15.14%	6.92%	5.69%	7.58%	8.80%	1.75%	13.86%
PearsonICA [26]	84.31%	57.66%	5.11%	3.14%	13.69%	17.87%	25.74%	10.75%	9.26%	1.20%	22.87%
Sobi [27]	68.56%	21.53%	7.15%	3.52%	14.64%	7.74%	7.76%	7.92%	4.78%	1.90%	14.55%

the ground-truth and the estimates is given in Fig.7, from which we obtained that: the fitted line was $y = 1.0021x - 0.2876$, where x indicates the ground-truth heart rate value, and y indicates the associated estimate; the R^2 value which was the measure of goodness of fit was 0.9914; Pearson Correlation Coefficient was 0.9957. Therefore, from Fig.6 and Fig.7, we found that the gap between the estimated HR values of our proposed algorithm and the ground-truth of HR was very little, indicating the excellent performance of our proposed algorithm again.

In order to further show the performance visually, compared to TROIKA, the estimation results on the recordings of Subject 3 are given in Fig.8 where we can see that the estimated HR values of our proposed algorithm were closer to the ground-truth compared to that of TROIKA, showing that our proposed algorithm outperformed TROIKA again.

Fig. 8. Estimation results on recordings of Subject 3 of the 12 training PPG datasets. The HR traces of two algorithms for HR estimation are plotted, one was our proposed algorithm, another one was TROIKA [17]. They were all compared to the ground-truth which was recorded simultaneously from ECG.

2) Experiments on the 10 Testing PPG Datasets: Table III and Table IV list the Error1 and Error2 of the 10 subjects' recordings stored in the 10 testing PPG datasets. From the two tables we observed that, the Error1 was 2.98 ± 5.44 and the Error2 was 2.91%. The result is similar to [17], [22] and [23], showing that even on 10 PPG datasets recorded with extremely strong MA when subjects doing strenuous exercise, the esti-

TABLE V
HR ESTIMATION RESULTS IN TERMS OF ERROR1 WHEN USING 8 PPG DATASETS WERE SHOWN IN THIS TABLE, AND THE MEANING OF EACH ROW WAS AS SAME AS TABLE III.
ERROR1

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8
Proposed	4.60	6.53	14.79	4.28	12.77	8.77	4.88	3.43
RLS+SSA	3.75	6.78	11.18	3.38	14.67	6.75	5.30	4.31
nonlinearRLS	39.65	4.99	15.88	3.06	24.02	19.49	5.40	4.15
SSA	3.13	9.93	21.46	3.08	15.95	12.85	5.55	5.08
FastICA [25]	47.70	4.57	19.34	22.75	20.74	10.73	20.42	9.89
PearsonICA [26]	41.58	22.64	22.13	3.43	9.20	9.84	12.81	48.04
Sobi [27]	27.24	21.51	21.74	6.06	11.40	13.11	26.57	35.36

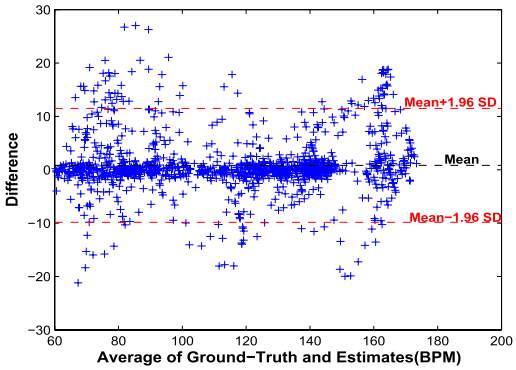

Fig. 9. The Bland-Altman plot of the estimates of our proposed algorithm on the 10 PPG datasets. The LOA was [-9.84,11.51] BPM (standard deviation σ = 5.44 (BPM).

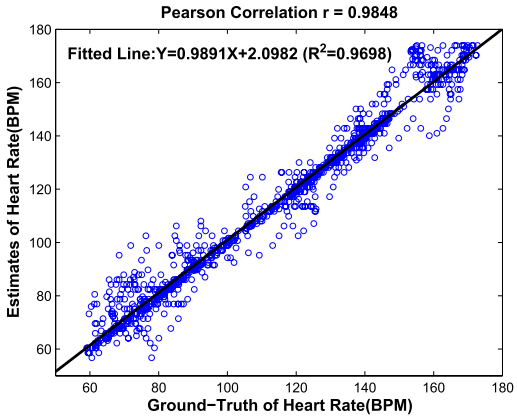

Fig. 10. Scatter plot on the 10 datasets between the ground-truth and the estimates of our proposed algorithm. Here mainly showed: the fitted line was $y = 0.9891x + 2.0982$; the R^2 value was 0.9698; the Pearson Correlation Correlation was 0.9848.

mation accuracy of our proposed algorithm was still very satisfactory. Similar to that of the 12 datasets, the first stage was replaced by RLS + SSA, only by RLS-based nonlinear adaptive filtering, only by SSA, FastICA [24], PearsonICA [25]) and one BSS algorithm (SOBI [26]), respectively. The results also reveal that our proposed MA removal method is better than others listed here and it is of high performance even when the subject doing complex motions where MA is extremely strong. It is worthy mentioning that this result beats the second place of the 2015 IEEE

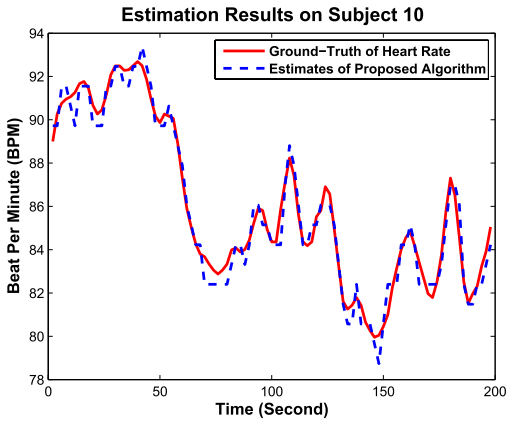

Fig. 11. Estimation results on recordings of Subject 10 of 10 PPG datasets. The HR traces of our proposed algorithm for HR estimating were plotted, and it was compared to the ground-truth which was recorded simultaneously from ECG.

Signal Processing Cup in terms of the average absolute error.[2]

To further observe the performance of HR estimating visually, the Bland-Altman is plotted in Fig.9 and the LOA was [-9.84,11.51] BPM. And the Scatter Plot is also given in Fig.10, which shows that the fitted line was $y = 0.9891x + 2.0982$ and the Pearson Correlation Coefficient was 0.9848. Obviously, Fig.9 and Fig.10 show that, the estimated HR values of our proposed algorithm were very close to the ground-truth when MA was extremely strong.

Fig.11 presents the HR trace of our proposed algorithm on the recordings of Subject 10 of 10 testing PPG datasets, again showing that the estimated HR values of our proposed algorithm were very close to the ground-truth HR values.

Clearly, from Table III, Table IV, Fig.9, Fig.10, and Fig.11, we observed that our algorithm can still obtain satisfactory estimation performance even on the 10 testing PPG datasets which were recorded when MA was extremely strong.

3) Experiments on the 8 Testing PPG Datasets: Table V and Table VI list the Error1 and Error2 of the 8 subjects' recordings stored in the 8 testing PPG datasets. Similar to

[2]The average absolute error (over the 10 testing datasets) of the first three places in the Cup were 2.27 BPM, 3.26 BPM, and 3.44 BPM, respectively. (http://www.signalprocessingsociety.org/spcup2015/.)

TABLE VI
HR ESTIMATION RESULTS IN TERMS OF ERROR1 WHEN USING 8 PPG DATASETS WERE SHOWN IN THIS TABLE, AND THE MEANING OF EACH ROW WAS AS SAME AS TABLE IV. ERROR2

	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6	Subj 7	Subj 8
Proposed	4.16%	10.12%	12.05%	4.12%	10.48%	7.49%	5.47%	4.13%
RLS+SSA	3.32%	10.48%	9.93%	3.31%	12.07%	6.00%	5.63%	4.26%
nonlinearRLS	30.53%	7.64%	12.52%	3.01%	19.91%	16.10%	6.03%	4.99%
SSA	2.85%	15.21%	20.88%	3.03%	13.25%	11.10%	6.15%	5.73%
FastICA [25]	36.78%	6.83%	20.45%	20.72%	17.01%	9.01%	22.13%	12.16%
PearsonICA [26]	32.06%	34.24%	21.73%	3.28%	7.75%	9.14%	14.37%	61.32%
Sobi [27]	20.92%	31.97%	21.19%	6.13%	9.74%	11.97%	30.36%	44.83%

that of the 12 datasets and 10 datasets, the first stage was replaced by RLS + SSA, only by RLS-based nonlinear adaptive filtering, only by SSA, FastICA [24], PearsonICA [25]) and one BSS algorithm (SOBI [26]). Clearly, we can conclude that our proposed MA removal method is superior to the other MA removal method listed here even when the is extremely strong.

IV. CONCLUSIONS

Heart rate estimation via wrist-type PPG signals is a popular research topic in the field of wearable sensing. In this paper, a hybrid motion artifact removal method is proposed, which combines nonlinear adaptive filtering and signal decomposition in wearable PPG during exercise, getting the best of the two approaches while overcoming their drawbacks. The novel MA removal method makes the proposed algorithm have high accuracy for HR estimation in the presence of strong MA, as confirmed by experiments on the PPG datasets in the 2015 IEEE Signal Processing Cup and other PPG datasets.

REFERENCES

[1] M. Pace and V.-A. Bricout, "Low heart rate response of children with autism spectrum disorders in comparison to controls during physical exercise," *Physiol. Behavior*, vol. 141, pp. 63–68, Mar. 2015.

[2] L. A. Sebastian, S. Reeder, and M. Williams, "Determining target heart rate for exercising in a cardiac rehabilitation program: A retrospective study," *J. Cardiovascular Nursing*, vol. 30, no. 2, pp. 164–171, 2015.

[3] E. A. P. J. Prawiro, C.-C. Hu, Y.-S. Chan, C.-H. Chang, and Y.-H. Lin, "A heart rate detection method for low power exercise intensity monitoring device," in *Proc. IEEE Int. Symp. Bioelectron. Bioinformat. (ISBB)*, Apr. 2014, pp. 1–4.

[4] H.-K. Jung and D.-U. Jeong, "Development of wearable ECG measurement system using EMD for motion artifact removal," in *Proc. 7th Int. Conf. Comput. Converg. Technol. (ICCCT)*, Dec. 2012, pp. 299–304.

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

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

[7] A. Schäfer and J. Vagedes, "How accurate is pulse rate variability as an estimate of heart rate variability?: A review on studies comparing photoplethysmographic technology with an electrocardiogram," *Int. J. Cardiol.*, vol. 166, no. 1, pp. 15–29, Jun. 2013.

[8] Y. Maeda, M. Sekine, and T. Tamura, "Relationship between measurement site and motion artifacts in wearable reflected photoplethysmography," *J. Med. Syst.*, vol. 35, no. 5, pp. 969–976, Oct. 2011.

[9] J. Y. A. Foo and S. J. Wilson, "A computational system to optimise noise rejection in photoplethysmography signals during motion or poor perfusion states," *Med. Biol. Eng. Comput.*, vol. 44, no. 1, pp. 140–145, Mar. 2006.

[10] 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.

[11] J. Yao and S. Warren, "A short study to assess the potential of independent component analysis for motion artifact separation in wearable pulse oximeter signals," in *Proc. 27th Annu. Conf IEEE Eng. Med. Biol.*, Sep. 2005, pp. 3585–3588.

[12] G. Comtois, Y. Mendelson, and P. Ramuka, "A comparative evaluation of adaptive noise cancellation algorithms for minimizing motion artifacts in a forehead-mounted wearable pulse oximeter," in *Proc. 29th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2007, pp. 1528–1531.

[13] A. B. Barreto, L. M. Vicente, and I. K. Persad, "Adaptive cancelation of motion artifact in photoplethysmyographic blood volume pulse measurements for exercise evaluation," in *Proc. IEEE 17th Annu. Conf. Eng. Med. Biol. Soc.*, vol. 2, Sep. 1995, pp. 983–984.

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

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

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

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

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

[19] P. S. R. Diniz, *Adaptive Filtering*, 4th ed. Boston, MA, USA: Kluwer, 2013.

[20] C. Mert and A. Milnikov, "Singular Spectrum Analysis Method as a universal filter," in *Proc. 5th Int. Conf. Appl. Inf. Commun. Technol. (AICT)*, Oct. 2011, pp. 1–5.

[21] N. Golyandina and A. Zhigljavsky, *Singular Spectrum Analysis for Time Series*. Berlin, Germany: Springer-Verlag, 2013.

[22] 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.

[23] B. Sun and Z. Zhang, "Photoplethysmography-based heart rate monitoring using asymmetric least squares spectrum subtraction and Bayesian decision theory," *IEEE Sensors J.*, vol. 15, no. 12, pp. 7161–7168, Dec. 2015.

[24] A. Hyvärinen, "Fast and robust fixed-point algorithms for independent component analysis," *IEEE Trans. Neural Netw.*, vol. 10, no. 3, pp. 626–634, May 1999.
[25] J. Karvanen, J. Eriksson, and V. Koivunen, "Pearson system based method for blind separation," in *Proc. 2nd Int. Workshop Independ. Compon. Anal. Blind Signal Separat.*, Helsinki, Finland, 2000, pp. 585–590.
[26] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "Second order blind separation of temporally correlated sources," in *Proc. Int. Conf. Digit. Signal Process.*, 1993, pp. 346–351.

Yalan Ye, photographs and biographies not available at the time of publication.

Yunfei Cheng, photographs and biographies not available at the time of publication.

Wenwen He, photographs and biographies not available at the time of publication.

Mengshu Hou, photographs and biographies not available at the time of publication.

Zhilin Zhang, photographs and biographies not available at the time of publication.