Annals of Biomedical Engineering, Vol. 42, No. 11, November 2014 (© 2014) pp. 2238–2250
DOI: 10.1007/s10439-014-1080-y

Photoplethysmograph Signal Reconstruction Based on a Novel Hybrid Motion Artifact Detection–Reduction Approach. Part I: Motion and Noise Artifact Detection

JO WOON CHONG,[1] DUY K. DAO,[1] S. M. A. SALEHIZADEH,[1] DAVID D. MCMANUS,[2] CHAD E. DARLING,[3] KI H. CHON,[1] and YITZHAK MENDELSON[1]

[1]Department of Biomedical Engineering, Worcester Polytechnic Institute, Worcester, MA 01609-2280, USA; [2]Cardiology Division, Departments of Medicine and Quantitative Health Sciences, University of Massachusetts Medical Center, Worcester, MA 01655, USA; and [3]Department of Emergency Medicine, University of Massachusetts Medical School, Worcester, MA 01655, USA

(*Received 28 February 2014; accepted 25 July 2014; published online 5 August 2014*)

Associate Editor Tingrui Pan oversaw the review of this article.

Abstract—Motion and noise artifacts (MNA) are a serious obstacle in utilizing photoplethysmogram (PPG) signals for real-time monitoring of vital signs. We present a MNA detection method which can provide a clean vs. corrupted decision on each successive PPG segment. For motion artifact detection, we compute four time-domain parameters: (1) standard deviation of peak-to-peak intervals (2) standard deviation of peak-to-peak amplitudes (3) standard deviation of systolic and diastolic interval ratios, and (4) mean standard deviation of pulse shape. We have adopted a support vector machine (SVM) which takes these parameters from clean and corrupted PPG signals and builds a decision boundary to classify them. We apply several distinct features of the PPG data to enhance classification performance. The algorithm we developed was verified on PPG data segments recorded by simulation, laboratory-controlled and walking/stair-climbing experiments, respectively, and we compared several well-established MNA detection methods to our proposed algorithm. All compared detection algorithms were evaluated in terms of motion artifact detection accuracy, heart rate (HR) error, and oxygen saturation (SpO_2) error. For laboratory controlled finger, forehead recorded PPG data and daily-activity movement data, our proposed algorithm gives 94.4, 93.4, and 93.7% accuracies, respectively. Significant reductions in HR and SpO_2 errors (2.3 bpm and 2.7%) were noted when the artifacts that were identified by SVM-MNA were removed from the original signal than without (17.3 bpm and 5.4%). The accuracy and error values of our proposed method were significantly higher and lower, respectively, than all other detection methods. Another advantage of our method is its ability to provide highly accurate onset and offset detection times of MNAs. This capability is important for an automated approach to signal reconstruction of only those data points that need to be reconstructed, which is the subject of the companion paper to this article. Finally, our MNA detection algorithm is real-time realizable as the computational speed on the 7-s PPG data segment was found to be only 7 ms with a Matlab code.

Keywords—Motion and noise artifacts, Photoplethysmography, Support vector machine.

INTRODUCTION

PPG is a non-invasive and low cost device to continuously monitor blood volume changes in peripheral tissues.[28] PPG is a useful technique since it is widely used to monitor HR, SpO_2, and can also be used to measure respiratory rates.[24] However, MNA can distort PPG recordings, causing erroneous estimation of HR and SpO_2.[28,32] It is because of the MNA that PPG has not yet been widely adopted as a possible sensor for mobile health applications. There are three distinct sources of MNA that can distort PPG recordings: environmental, physiological and experimental artifacts, which can be attributed to (1) electromagnetic and power interference around the body; (2) cross talk pickup of other physiological signals; and (3) instrumental noise, respectively.[18,36,39] MNA, which are comprised of all of the aforementioned noise sources, are difficult to filter since they do not have a predetermined frequency band and their spectrum often overlaps with that of the desired PPG signal. Despite

Address correspondence to Ki H. Chon, Department of Biomedical Engineering, Worcester Polytechnic Institute, Worcester, MA 01609-2280, USA. Electronic mail: jw9607@gmail.com, dkdao2013@gmail.com, smasalehizadeh@gmail.com, mcmanusd@ummhc.org, Chad.Darling@umassmed.edu, kichon@wpi.edu, ym@wpi.edu

these challenging scenarios, in our companion paper, we describe a method that can reconstruct the MNA contaminated data segments so that accurate heart rates and SpO_2 values can be estimated.

MNA in PPG readings are caused by (1) the movement of venous blood as well as other non-pulsatile components along with pulsatile arterial blood and (2) variations in the optical coupling between the sensor and the skin.[2,28,37,38] Various approaches to mitigate motion artifacts by improving sensor attachment have been proposed.[20,27] However, these design improvements do not provide a significant reduction of motion artifacts. Algorithm-based MNA reduction methods were also proposed. These include time and frequency domain filtering, power spectrum analysis, and blind source separation techniques.[12,15,19,29–31,40] However, these have high computational complexity and more importantly, they operate even on clean PPG portions where MNA reduction is not needed and consequently may distort the clean signal.[15–17,29] Hence, accurate MNA detection, which identifies clean PPG recordings from corrupted portions, is essential for the subsequent MNA reduction algorithm so that it does not distort the non-corrupted data segments.[16] Moreover, more computationally efficient MNA algorithms can be designed since they can be tailored only to the MNA contaminated data segments.

MNA detection methods are mostly based on a signal quality index (SQI) which quantifies the severity of the artifacts. Some approaches quantify SQI using waveform morphology[21,22,35] or filtered output,[14] while other derive SQI with the help of additional hardware such as accelerometer and electrocardiogram sensing.[7,17] Statistical measures, such as skewness, kurtosis, Shannon entropy, and Renyi's entropy, have been shown to be helpful in determining a SQI.[5,41] However, these techniques require manual threshold settings for each parameter to classify if the PPG signal is clean or corrupted. Although a support vector machine (SVM)-based classification method addresses the need of threshold setting,[41] this approach considers limited and controlled types of motions. The authors are not aware of any detailed studies providing representative and comprehensive features distinguishing clean from corrupted PPG signals under various types of motions.

In this paper, we propose an accurate and comprehensive MNA detection algorithm which detects MNA in PPG under various types of motion. We first introduce time-domain parameters quantifying MNA in the recorded PPG signal. We then consider their statistical measures as input variables for a machine learning-based MNA detection algorithm. Our MNA detection algorithm is self-trained by the SVM with clean and corrupted PPG data sets, and then the trained SVM tests the unknown PPG data. We tested the efficacy of our proposed algorithm on PPG data sets recorded from the finger and forehead pulse oximeters in simulations, laboratory-controlled and walking/stair-climbing experiments, respectively.

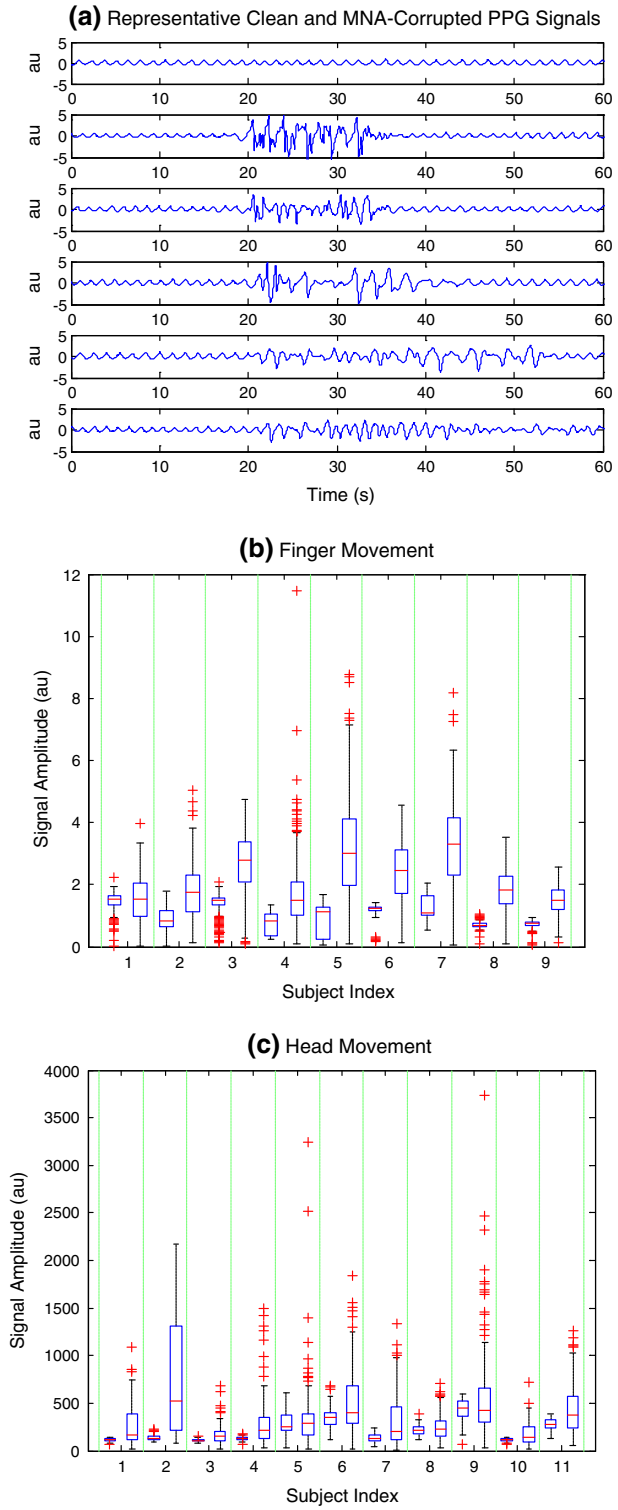
MATERIALS AND METHOD

Experimental Protocol and Preprocessing

PPG signals were obtained from custom reflectance-mode prototype pulse oximeters. PPG data with laboratory-controlled head and finger movement, daily-activity movement, or simulated movement were collected respectively from healthy subjects recruited from the student community of Worcester Polytechnic Institute (WPI). This study was approved by WPI's IRB and all subjects were given informed consent prior to data recording.

In *laboratory-controlled head movement* data, motion artifacts were induced by head movements for specific time intervals in both horizontal and vertical directions. Eleven healthy volunteers were asked to wear a forehead reflectance pulse oximeter along with a reference Masimo Radical (Masimo SET®, Masimo Corporation, CA, USA) finger type transmittance pulse oximeter. After baseline recording for 5 min without any movement, subjects were instructed to introduce motion artifacts for specific time intervals varying from 10 to 50% within a 1 min segment as shown in Fig. 1a. For example, if a subject was instructed to perform left–right movements for 6 s, a 1 min segment of data would contain 10% noise. MNA amplitudes varied for each subject as shown in Fig. 1b. Specifically, mean amplitude ratios and the deviation ratios (between 75th mean and 25th percentile values) between corrupted and clean signals are in the ranges of [0.9935 3.0092] and [1.3375 11.1250], respectively. The right middle finger with the sensor attached to the Masimo pulse oximeter was kept stationary. HR and SpO_2 signals were acquired by the Masimo pulse oximeter at 80 and 1 Hz, respectively, and were acquired synchronously with the PPG signals recorded from the forehead sensor.

In *laboratory-controlled finger movement* data, motion artifacts were induced by left–right movements of the index finger. Nine healthy volunteers were asked to sit and wear two reflection type PPG pulse oximeters (TSD200, BIOPAC Systems Inc., CA, USA) on their index and middle fingers, respectively. After baseline recording for 5 min without any movement to acquire clean data, motion artifacts were induced by left–right movements of the index finger while the middle finger was kept stationary as a reference. Similar to the head



◄FIGURE 1. PPG signals recorded during voluntary motion artifact conducted in a laboratory setting. (a) A representative clean forehead-PPG signal recorded during voluntary motion artifact conducted in a laboratory setting (1st row). The mixed (up–down and left–right) movement of the forehead to which the PPG probe was attached for predetermined time interval induced 10–50% noise (2nd–6th row) within a 60 s PPG segment (b) amplitudes of clean (left) and MNA-corrupted (right) finger-PPG signals of 9 subjects, and (c) amplitudes of clean (left) and MNA-corrupted (right) forehead-PPG signals of 11 subjects. The central line on each box corresponds to the median; the edges of the box correspond to the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered as outliers, and outliers are plotted individually.

movement data, motion was induced at specific time intervals corresponding to 10–50% duration in an 1 min segment. Such controlled movement was repeated five times per subject. MNA amplitudes varied for each subject as shown in Fig. 1c. The mean amplitude ratios and deviation ratios (between 75th mean and 25th percentile values) between corrupted and clean signals are in the range of [0.9360 4.0226] and [1.3892 32.2647], respectively. The pulse oximeters were connected to a biopotential amplifier (PPG100, BIOPAC Systems Inc., CA, USA) having a gain of 100 and cut-off frequencies of 0.05–10 Hz. The MP1000 (BIOPAC Systems Inc., CA, USA) was used to acquire finger PPG signals at 100 Hz. The *daily-activity movement PPG data* were recorded while subjects were walking straight or climbing stairs for 45 min. Nine subjects were asked to walk or climb stairs after wearing a forehead reflectance pulse oximeter along with a Holter electrocardiogram (ECG) monitor (Rozinn RZ153+, Rozinn Electronics, Inc., NY, USA) at 180 Hz and a Masimo Rad-57 pulse oximeter (Masimo Corporation, CA, USA) at 0.5 Hz. The reference ECG was obtained from the Holter ECG monitor while HR and SpO_2 readings were measured from the Masimo pulse oximeter connected to the subject's right index finger, which was held against the chest to minimize motion artifacts. Finally, the *simulation movement PPG data* were generated by the addition of white noise to the clean PPG data. PPG data were preprocessed by a 6th order infinite impulse response (IIR) band pass filter with cut-off frequencies of 0.5 and 12 Hz. Zero-phase forward and reverse filtering was applied to account for the non-linear phase of the IIR filter. After these preprocessing, the following parameters for classifying clean and corruption were derived.

Parameters from PPG Signals

Motion-corrupted PPG signals are observed to have noticeably different pulse amplitudes, pulse widths and trough depths when compared to clean PPG signals.[35] Moreover, morphology and amplitude ratios of corrupted PPG signals differ from those of clean signals. It was found that most PPG signals show strong similarity among noise-free pulses, but large variation among successive poor and bad quality pulses.[35] Based on our observations on measured clean and corrupted

PPG signals shown in Fig. 1 as well as previous work,[35] we considered standard deviation values of the pulse amplitude, pulse width and pulse shape to quantify differences among these features. We also considered systolic and diastolic ratios since they are observed to be within a well-defined range for clean PPG pulses.[23] The following four parameters were selected since they represent the variability present in corrupted PPG signals as shown in Fig. 1.

Standard Deviation of Peak-to-Peak Interval (STD_{HR})

The $STD_{HR,n}$ of the nth segment is defined by:

$$STD_{HR,n} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(D_{n,i} - \overline{D_n}\right)^2} \quad (1)$$

where $D_{n,i}$ is peak-to-peak interval at the ith pulse of the nth segment and $\overline{D_n}$ is mean peak-to-peak interval of the nth segment. The $D_{n,i}$ is calculated by the difference $T_{peak,n,i} - T_{peak,n,i-1}$ between two successive peak times.

Standard Deviation of Peak-to-Peak Amplitude (STD_{AMP})

The $STD_{AMP,n}$ is calculated by substituting $D_{n,i}$ and $\overline{D_n}$ in (Eq. 1) with $A_{n,i}$ and $\overline{A_n}$, respectively, where $A_{n,i}$ is peak amplitude at the ith pulse of the nth segment and $\overline{A_n}$ is mean peak-to-peak interval of the nth segment. The $A_{n,i}$ is defined by the difference between the ith peak and the forthcoming $(i+1)^{th}$ trough amplitude values.

Standard Deviation of Systolic and Diastolic Ratio (STD_{SD})

The $STD_{SD,n}$ of the nth segment is derived by replacing $D_{n,i}$ and $\overline{D_n}$ in (Eq. 1) by $R_{SD,n,i}$ and $\overline{R_{SD,n}}$, respectively, where $R_{SD,n,i}$ is systolic and diastolic time interval ratio at the ith pulse of the nth segment and $\overline{R_{SD,n}}$ is the mean systolic and diastolic time interval ratio of the nth segment. The $R_{SD,n,i}$ is calculated by

$$R_{SD,n,i} = \left(T_{trough,n,i} - T_{peak,n,i}\right)/\left(T_{peak,n,i} - T_{trough,n-1,i}\right) \quad (2)$$

where $T_{trough,n,i}$ denotes the trough times (or lowest point) at the ith pulse of the nth segment.

Mean-Standard Deviation of Pulse Shape (STD_{WAV})

To derive pulse shape, we take N_{samp} sample points of a pulse. The $STD_{WAV,n}$ of the nth segment is derived by taking average of the standard deviation at each sample point as follows:

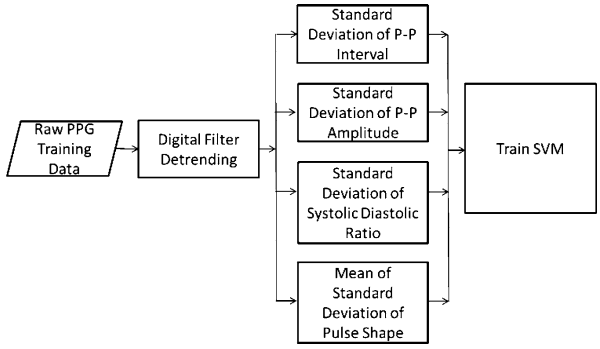


FIGURE 2. Training phase of the proposed SVM-based motion detection algorithm. Four time-domain features corresponding to (1) standard deviation of peak-to-peak intervals (2) standard deviation of peak-to-peak amplitudes (3) standard deviation of systolic and diastolic interval ratio, and (4) mean standard deviation of pulse shape, are candidate input variables to the SVM.

$$STD_{WAV,n} = \mathrm{E}[STD_{WAV,n,m}] \quad (3)$$

where $STD_{WAV,n,m}$ is calculated by substituting $D_{n,i}$ and $\overline{D_n}$ in (Eq. 1) with $q_{n,i}(m)$ and $\overline{q_n(m)}$, respectively, where $q_{n,i}(m)$ is the mth sample at the ith pulse of the nth segment and $\overline{q_n(m)}$ the mean at the mth sample of the nth segment.

SVM-BASED DETECTION OF MOTION/NOISE ARTIFACTS

Classification by Support Vector Machine (SVM)

SVM was applied to build a decision boundary classifying motion corruption from clean PPG signals. SVM is widely used in classification and regression due to its accuracy and robustness to noise.[25] The SVM consists of training and test phases described in the following sections.

Training Phase

A flow chart of the training phase in the SVM-based MNA detection algorithm is shown in Fig. 2. The SVM first derives the parameter values from clean and corrupted PPG training segments which are labeled separately (clean: 0, corrupted: 1). The SVM then trains itself with the labeled parameter values and finds the support vectors among the parameter values which maximize the margin (or the distance) between different classes. Finally, the SVM builds a decision boundary from the support vectors. If the estimated decision from the decision boundary is different from its known label, the decision is regarded as a *training error*. We consider a soft-margin SVM which can set the boundary even when the data sets are mixed and cannot be separated. In the soft-margin SVM

algorithm, slack variables are introduced to minimize the training error while maximizing the margin. Soft-margin SVM uses the following equation to find the support vectors.[13]

$$\text{Minimize } C\sum_{sv=1}^{N}\delta_{sv}+\frac{1}{2}\langle \mathbf{w}_s,\mathbf{w}_s\rangle,$$

$$\text{Subject to } T_{sv}(\langle \mathbf{w}_s,\mathbf{y}_{sv}\rangle+b_s)\geq 1=\delta_{sv} \text{ for } sv=1,2,\ldots,N \ =1,2,\ldots,N, \text{ and } \delta_{sv}\geq 0 \quad (4)$$

where C is regulation parameter, N is the number of vectors, δ_{sv} is the slack variable, $\mathbf{w}_s$is the weight vector and $<\cdot,\cdot>$ is the inner product operation. The $T_{sv}$is the svth target variable, $\mathbf{y}_{sv}$ are the svth input vector data, and b_s is the bias. The SVM decision boundary F_{sv} is derived as

$$F_{sv}=\langle \mathbf{w}_s^*,\mathbf{y}\rangle+b_s^*=0 \quad (5)$$

where $\mathbf{w}_s^*$ and b_s^* are weight factor and bias, respectively, obtained from Eq. (4), and $\mathbf{y}$ is the input point.

By transforming the $\mathbf{y}_{sv}$ and $\mathbf{y}$ term to $\mathbf{y}_{sv}\rightarrow(\mathbf{y}_{sv})$ and $\mathbf{y}\rightarrow(\mathbf{y})$, the non-linear SVM can be transformed to a linear SVM. For nonlinear SVM, Eq. (4) is modified as

$$T_{sv}(\langle \mathbf{w}_s,(\mathbf{y}_{sv})\rangle+b_s)\geq 1 \quad (6)$$

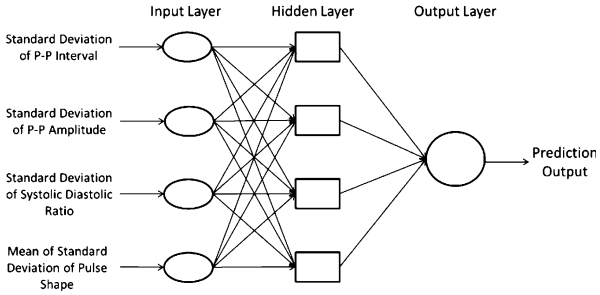


FIGURE 3. Test phase of the proposed SVM-based motion detection algorithm. The hidden layers correspond to kernel function of the SVM. The function between hidden layer and output layer is a linear operator.

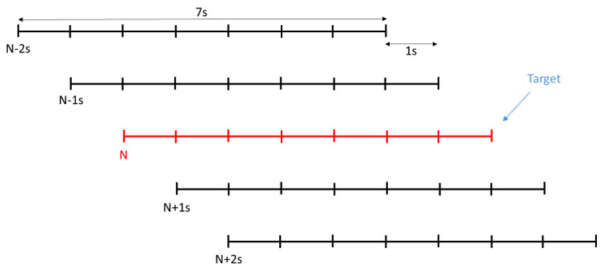


FIGURE 4. Enhancement of MNA detection by diversity. Neighbor segments are the segments surrounding a target segment within ±2 s. Decisions on the target segment are based on a majority vote from the decisions of neighbor segments as well as the one of the target segment (red).

To facilitate the operation on nonlinear SVM, a kernel function $K_s(\cdot,\cdot)$, which is a dot-product in the transformed feature space as follows, is used,

$$K_s(\mathbf{y}_{sv},\mathbf{y}_{sv'})=\langle(\mathbf{y}_{sv}),(\mathbf{y}_{sv'})\rangle \quad (7)$$

where $sv'=1,2,\ldots,N$.

Test Phase

Figure 3 shows a flow chart of the test phase in the SVM-based MNA detection algorithm. We partition PPG data into many 7-s segments, and derive parameters from each PPG portion to examine if it is corrupted by motion artifacts or not.

Enhancement of MNA Detection by Major Votes

To enhance MNA detection performance, the proposed algorithm incorporates multiple decisions on a set of neighbor segments in deciding whether a "target" segment is clean or corrupted. A neighbor segment is defined as a segment surrounding a target segment within $\pm T_{\text{neighbor}}$ seconds. A decision on a neighbor segment is highly likely to be the same as the decision on a target segment since PPG pulses in the neighbor segments are most likely to exhibit similar dynamics to the target segment.

The algorithm gathers the decisions of neighbor segments as well as target segment (see Fig. 4) and makes a final decision on the target segment based on a majority vote concept.

RESULTS

We evaluated the performance of the MNA detection algorithm for various types (simulated, laboratory controlled, and daily activities) of motion-corrupted PPGs to validate its performance in a wide range of scenarios. For all types of motions, the PPG recordings were divided into 7-s segments since this was determined to be the optimal size among the data length tested from 3 to 11 s (see Section IV-B). Another study[41] has also found the optimal segment size to be 7-s, hence, this allowed direct comparison between the two algorithms. We compared the proposed algorithm with four recently published MNA detection algorithms based on kurtosis (K), Shannon entropy (SE), Hjorth 1 (H1), and Hjorth 2 (H2) metrics,[9,34,41] respectively. As performance metrics, we considered classification accuracy, sensitivity and specificity. We also investigated mean HR and SpO_2 errors as well as detection error ratio.

Reference: Clean vs. corrupted

The following are criteria which we adopted to reference PPG segments (clean or corrupted) for each experiment. A visual reference was excluded to avoid subjective decisions by visual inspectors; for subtle MNA there were large disagreements among visual inspectors. Instead, we performed objective decisions based on controlled corruption start ($T_{corr,start}$) and end ($T_{corr,end}$) time points, ECG-derived heart rate (HR_{ECG}), PPG-derived heart rate (HR_{PPG}), and SpO_2 (SpO_{2PPG}) from PPG signals. HRs from ECG signals and pulse rates (PRs) from PPG signals are shown to have high agreement when the subjects do not move.[3] Specifically, the difference between HRs and PRs for 10 subjects is within approximately 5 beats/min on average. Even during heavy exercise, sudden increases or decreases in HR between successive beats are within 20 beats/min[6] while sudden SpO_2 increases or decreases between successive beats are around 2%.[1] Since subjects' movements in this work are less severe than heavy exercise, we expect sudden HR and SpO_2 changes to be smaller than the numbers noted above. Based on these experimental observations, we set the clean vs. motion-corrupted data classification criteria for PPG segments as follows.

Laboratory controlled data (forehead and finger)

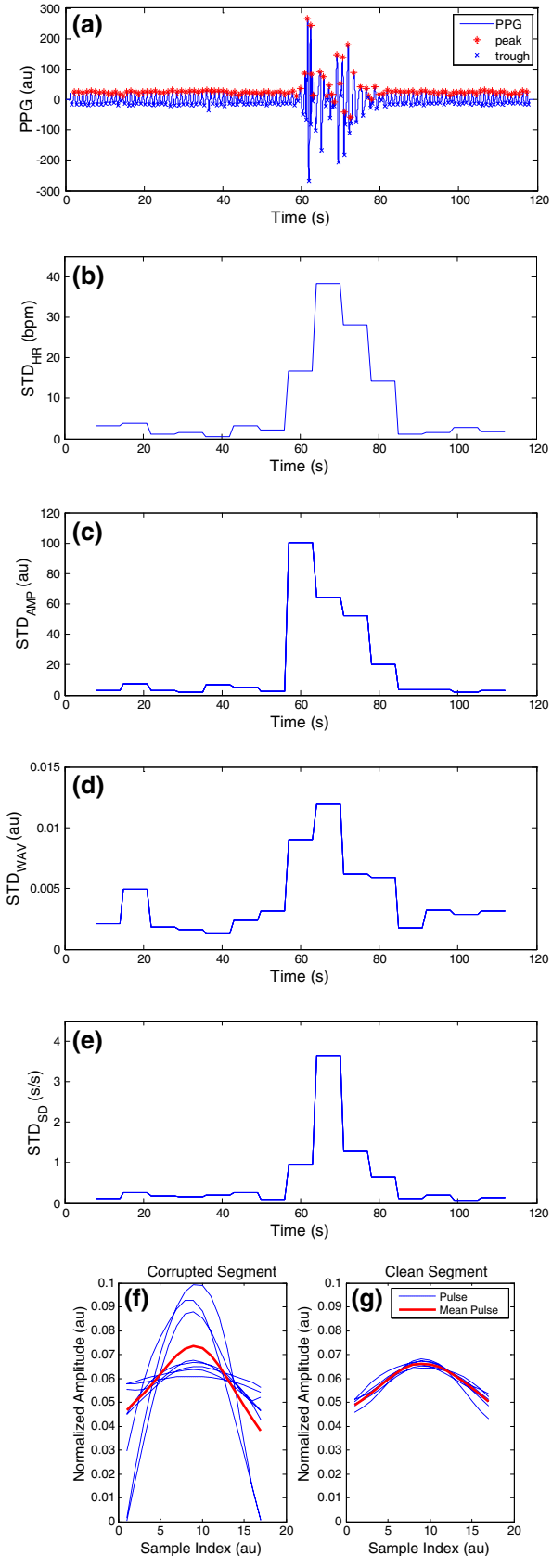
- If more than 85% of a segment is outside of [$T_{corr,start}$, $T_{corr,end}$], the segment was considered *clean*. Otherwise, the segment was referenced to be *corrupted*.
- If SpO_{2PPG} deviates by 10% from the mean of SpO_{2PPG} in a segment, then the segment was referenced to be corrupted.
- Successive difference, $|\text{diff}(HR_{PPG}(i+1) - HR_{PPG}(i))|$, from PPG signals is larger than 20 bpm for at least one pulse during a segment, then the segment was referenced to be corrupted.

Daily activity data (walking and stair-climbing)

- Successive difference, $|\text{diff}(HR_{ECG}(i+1) - HR_{ECG}(i))|$, from ECG signals is larger than

TABLE 1. Numbers of subjects and numbers of clean and corrupted segments per each motion artifact.

Type	Subtype	# of subjects	# of clean	# of corrupted
Simulation	Simulation	N/A	N/A	N/A
Laboratory controlled	Finger	13	195	105
	Forehead	11	190	110
Daily-activity	Walking/ stair-climbing	9	125	175



◀ **FIGURE 5. A sample forehead recorded PPG signal (a) along with the (b) standard deviation of *P–P* intervals (c) standard deviation of *P–P* amplitudes (d) standard deviation of systolic and diastolic time ratio, and (e) mean standard deviation of pulse shape, computed for each segment. The normalized sampled corrupt and clean PPGs for mean standard deviation of pulse shape (N_{samp} = 17) are given in (f).**

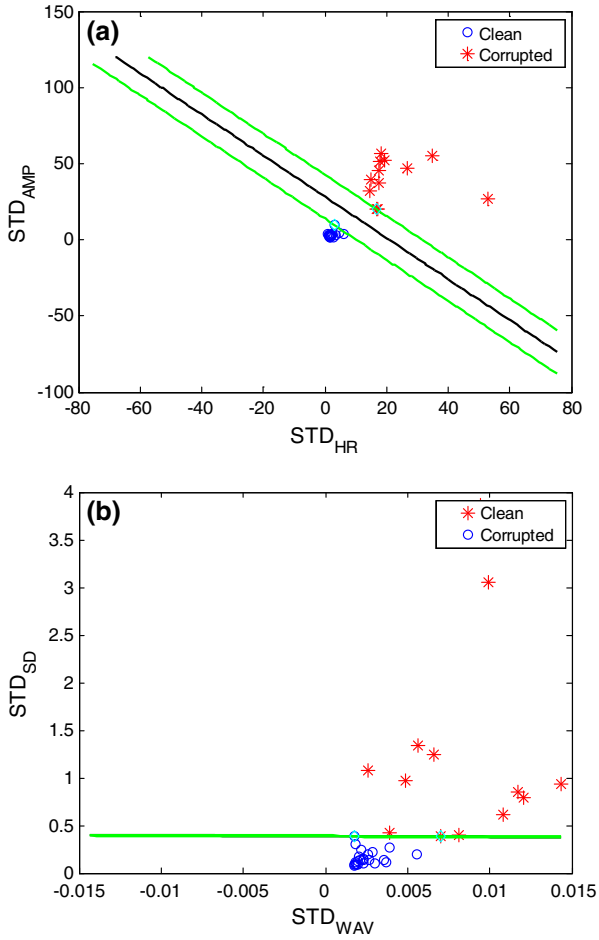


FIGURE 6. Trained SVM classification with a sample training finger recorded PPG signal is given with (a)–(b) pairs of two parameters. The SVM decision and margin boundaries are marked by black and green lines, respectively.

20 bpm for at least one pulse during a segment, then the segment was excluded.

- If SpO_{2PPG} deviates by 10% from the mean of SpO_{2PPG} in a segment, then the segment was referenced to be corrupted.
- If $|\text{diff}(HR_{PPG}(i+1) - HR_{PPG}(i))|$ is larger than 20 bpm for at least one pulse during a segment, then the segment was referenced to be corrupted.
- If $|HR_{ECG} - HR_{PPG}| < 5$ bpm during more than 85% of a segment, the segment was considered clean. Otherwise, the segment was referenced to be corrupted.

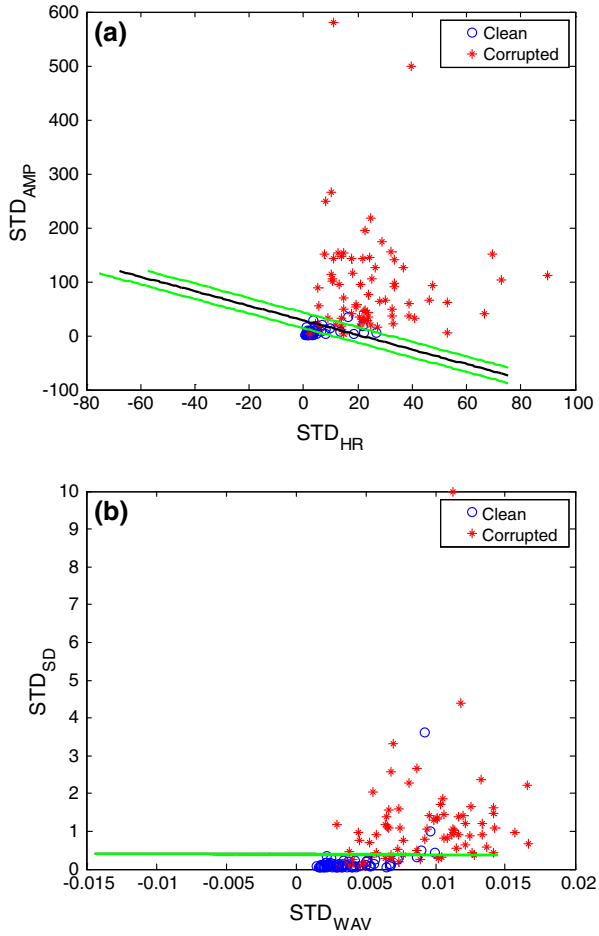


FIGURE 7. Validation: pairs of parameters for clean and corrupted PPG signals.

Table 1 describes the number of clean and corrupted PPG segments for each motion type used in the experiment as determined by the criteria defined above

Classification Accuracy

A sample forehead PPG signal and its corresponding 56–85 s have larger parameter values compared to clean segments between 1–56 s and 85–112 s. We sampled 17 points (N_{samp} = 17) of each pulse using the spline interpolation irrespective of test subjects and conditions to derive pulse shapes as shown in Fig. 5f. This enabled pulse shape comparisons within and among test subjects.

Figures 6a and 6b show $(STD_{HR}, STD_{AMP})$and (STD_{SD}, STD_{WAV}) of clean (circle) and corrupted (star) forehead signals, respectively, with corresponding SVM boundaries (black line). To lower computational complexity, a linear kernel was considered for the SVM in the experiment. We adopted inverse *k-fold* cross-validation which selects one-fold for training and the remaining k – 1-folds for testing.[33] We trained

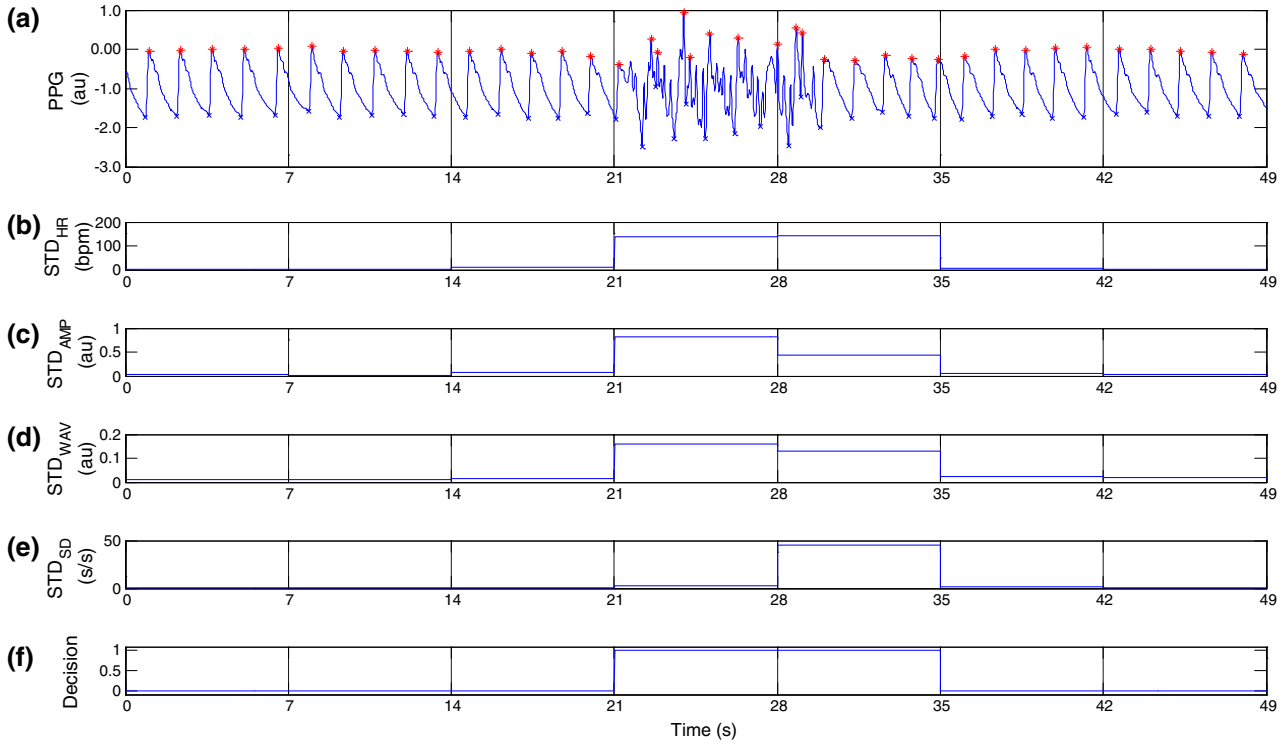


FIGURE 8. A representative PPG signal with detected peaks (red) (a) along with the (b) standard deviation of *P–P* intervals (c) standard deviation of *P–P* amplitudes (d) mean standard deviation of pulse shape and (e) standard deviation of systolic and diastolic time ratio, computed for each segment.

SVM with data from one subject after removing outliers and tested with data from 10 subjects. While there were small differences in the test errors among different k-fold cross-validation cases, we selected the least test-error case as the training data for each subject. Retraining was not performed since each subject's data consisted of a sufficient number of clean and corrupted PPG signals and the signal-to-noise ratios of the clean PPG signals were usually higher than the algorithm's sensitivity level. We optimized regularization parameter value (C) of the linear kernel SVM in terms of minimizing the training error rate. We adopted a 11-fold cross-validation and grid search ($C = \{10^{-3}, 10^{-2}, 10^{-1}, 1, 10^{1}, 10^{2}, 10^{3}\}$) which is widely used to determine C.[4] Figure 7 shows classification results by the SVM boundaries obtained from Fig. 6. Figure 8 shows a representative PPG signal with detected peaks (red) along with the corresponding statistical parameter values. Note the corrupted PPG signal interval between 21 to 31 s. The discrepancy between corrupted and clean portions is reflected by parameters STD_{HR}, STD_{AMP}, STD_{SD} and STD_{WAV}. The parameter values from the corrupted PPG segments exhibit larger variability and consequently have higher standard deviation value compared to those from clean data segments. The STD_{HR}, STD_{AMP} and STD_{WAV} have large values between 21 and 35 s (see Figs. 8b, 8c, and 8d), while STD_{SD} has large value only between 21 and 28 s (see Fig. 8e). Using SVM with these parameter values, the proposed algorithm correctly discriminated MNA corrupted segment between 21 and 35 s (see Fig. 8f). Table 2 presents C for finger, forehead, and walking/stair-climbing data. We tested our algorithm to different segment lengths varying from 3 to 11 s and calculated their mean classification accuracies, which are provided in Table 3. Among the different data segment lengths tested, the 7-s segment provided the highest classification accuracies for all data: finger, forehead and walking/stair-climbing PPG signals. Accuracy, specificity, and sensitivity for each dataset are presented in Table 4. On average, the SVM performance using the 7-s segment showed a 93.9% accuracy, 92.4% specificity, and 94.3% sensitivity. To derive the lower bound of our algorithm's

TABLE 2. *C* obtained by ninefold cross-validation and grid search method.

Type	Subtype	*C*
Simulation	Simulation	100
Laboratory controlled	Finger	1000
	Forehead	1
Daily-activity	Walking/stair-climbing	0.01

TABLE 3. Detection accuracy (mean ± SD) for varying window length.

	Window length				
Type	3	5	7	9	11
Finger	0.883 ± 0.042	0.906 ± 0.035	0.944 ± 0.033	0.944 ± 0.033	0.875 ± 0.042
Forehead	0.883 ± 0.023	0.880 ± 0.027	0.934 ± 0.035	0.856 ± 0.045	0.805 ± 0.044
Walk-stair climbing	0.813 ± 0.033	0.871 ± 0.039	0.937 ± 0.026	0.867 ± 0.052	0.856 ± 0.056

TABLE 4. Detection accuracy, specificity and sensitivity (mean ± SD) for 7-s segment.

Type	Accuracy	Specificity	Sensitivity
Finger	94.4 ± 3.3	94.7 ± 4.5	94.7 ± 3.4
Forehead	93.4 ± 3.5	96.7 ± 3.0	88.8 ± 7.9
Walking/stair-climbing	93.7 ± 2.7	91.4 ± 2.0	93.9 ± 5.0

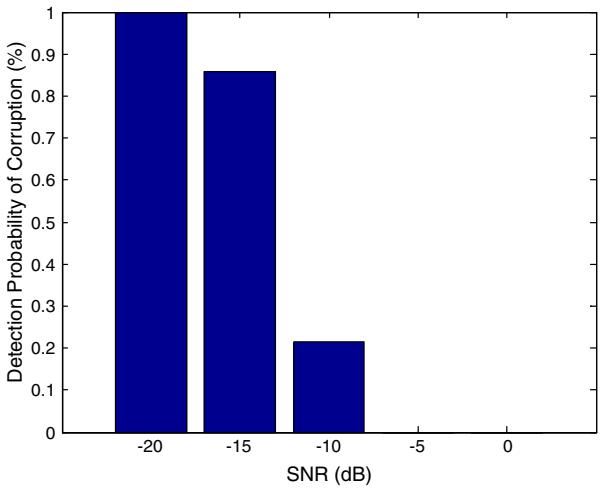


FIGURE 9. Detection probability of corruption by additive white Gaussian noise (AWGN) for varying SNR from −20 to 0 dB. 50 AWGN realizations for each SNR level are separately added to a non-MNA corrupted PPG. Each realization is tested by the proposed MNA detection algorithm to compute the detection probability of corruption.

performance[8] as well as to evaluate the sensitivity of our MNA detection algorithm, we added Gaussian white noise (GWN) of varying signal-to-noise (SNR) levels to a representative non-MNA corrupted PPG signal. For each SNR, 50 independent realizations of clean PPG signal with GWN are generated. As shown in Fig. 9, the PPG signals with a SNR below −10 dB are detected as corrupted data with our algorithm. For a SNR of −20 dB, every segment was detected as corrupted.

Performance Comparison of MNA detection Algorithms

Our algorithm was compared with other artifact detection methods based on H1, H2, K and SE since these methods have been shown to provide good detection accuracies.[9–11] The H1 and H2 parameters represent the central frequency and half of bandwidth, respectively, and are defined as follows:

$$H_1 = \sqrt{\frac{\overline{\nu}_2(n)}{\overline{\nu}_0(n)}} \text{ and } H_2 = \sqrt{\frac{\overline{\nu}_4(n)}{\overline{\nu}_2(n)} - \frac{\overline{\nu}_2(n)}{\overline{\nu}_0(n)}} \quad (8)$$

where $\overline{\nu}_i(n) = \int_{-\pi}^{\pi} \nu^i S_{ypDC}(e^{J\nu})d\nu$. Here, $S_{ypDC}(e^{J\nu})$ is the power spectrum of signal $y_{pDC}(n)$.

For a fair comparison, all detection methods used 7 s data segments. Figures 10a, 10b, and 10c compare the medians and 25th and 75th percentiles of detection accuracy, sensitivity, and specificity, which are obtained from each subject, for all five detection methods for the finger, head and walking/stair-climbing data sets. In general, our SVM method consistently yielded higher performance with a mean accuracy of 94%, sensitivity of 97%, and a specificity of 92%; whereas other methods showed fluctuations depending on which datasets were used. In the finger recorded data, H1 yielded a slightly higher accuracy than all other methods due to higher specificity, but the detection sensitivity was lower.

HR and SpO_2 Estimation

Figure 11a compares the mean HR error and detection error fraction from five MNA detection methods for walking/stair-climbing data. The HR errors were defined by the difference between the estimated HR derived from the PPG and the reference HR readings; the detection error fraction is defined as a ratio of the number of erroneous detection events to that of total detection events. Low mean HR error and low detection error fraction would reflect an effective artifact detection algorithm. Our algorithm yielded the lowest HR error and detection error fraction among

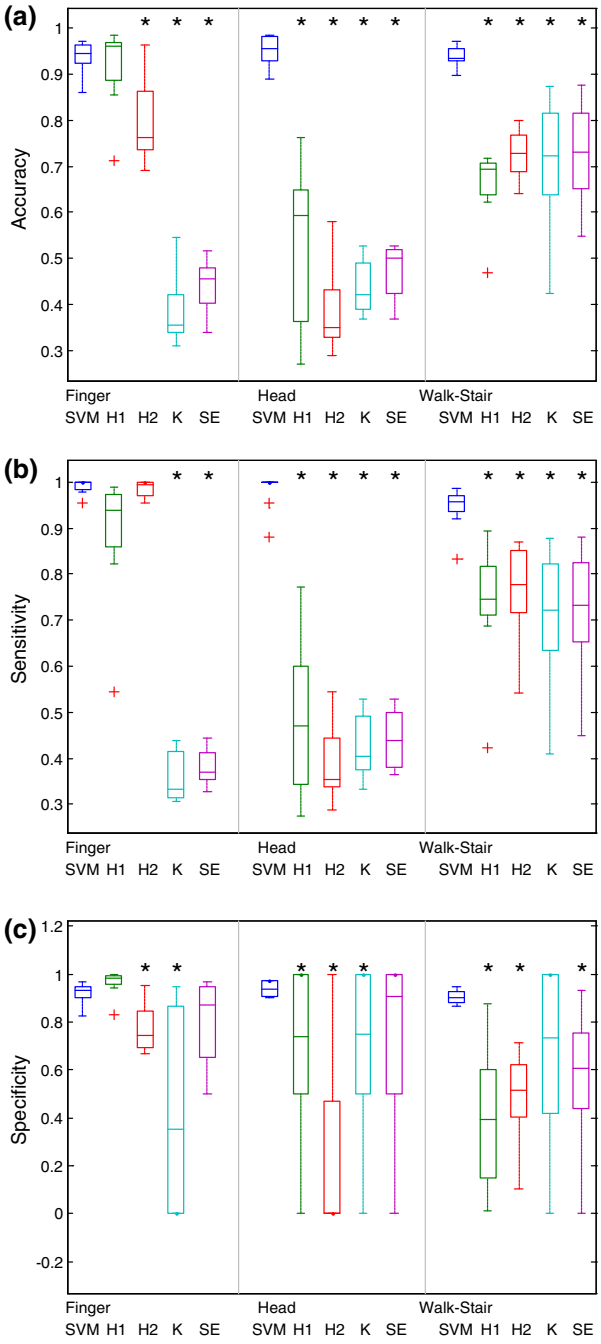


FIGURE 10. Classification performance comparison between our SVM algorithm, Hjorth (H1, H2), Kurtorsis and Shanon Entropy (K, SE) parameters. (a) Accuracy; (b) Sensitivity; (c) Specificity. The central mark on each box corresponds to the median; the edges of the box correspond to the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually. (*) indicate the mean is significantly different ($p < 0.05$ at 95% CI) between SVM and other methods used for comparison.

the other MNA methods we compared. Figure 11b compares mean SpO_2 error and detection error fraction from five MNA detection methods. The SE based detection method showed a lower mean SpO_2 error than our algorithm, but its detection error fraction was very high (>70%), indicating that the error was computed based on only 30% of clean data. On the other hand, the proposed SVM algorithm resulted in a mean SpO_2 error of 2.7 with a detection error of only 6.3%. Figure 12 compares five MNA detection methods in terms of paired t test results of HR and SpO_2 estimation and detection accuracy. On average, the SVM algorithm outperformed the K, SE, H1 and H2 methods with HR errors of 2.3 bpm, SpO_2 errors of 2.7% and detection error fraction of 6.3%.

DISCUSSION

Robust real-time MNA detection algorithms for raw PPG signals have been elusive to date. In this study, an SVM-based method is introduced to detect MNA-corrupted PPG data. Reconstruction of MNA-corrupted PPG segments is described in the companion paper. The aim of the current paper is to detect the MNA-corrupted PPG segments as accurately as possible. The question is how to detect MNA in an adaptive way to maximize detection accuracy so that PPG signal distortions is minimized. To address this question, we used four parameters derived from the PPG data: (a) standard deviation of peak-to-peak intervals (b) standard deviation of peak-to-peak amplitudes (c) standard deviation of systolic and diastolic time ratios, and (d) mean-standard deviation of pulse shapes. These four parameters are then used as inputs to an SVM-based MNA detection method with a sliding window with a major vote concept. We use these parameters since motion corrupted PPG signals are observed to have noticeably different amplitude values and have large variations for successive pulses when compared to the clean PPG signals.[35] We are currently working on finding more PPG relevant parameters to enhance detection performance on diverse types of MNA-corrupted PPG signals.

Many detection algorithms have been proposed to detect MNAs or quantify signal quality of PPG pulses. Various approaches use a set of several PPG-derived parameters to detect MNA, but the test data was confined to limited types of motions.[7,14,17,21,22,35,41] Given that most methods provide adequate MNA detection performance, we evaluated our proposed algorithm by comparing it to them.

The results demonstrated that our proposed SVM-based MNA detection algorithm provided higher classification accuracy as well as lower HR and SpO_2 errors compared to the conventional detection methods. The paired t test was performed to determine if there is significant difference between classification

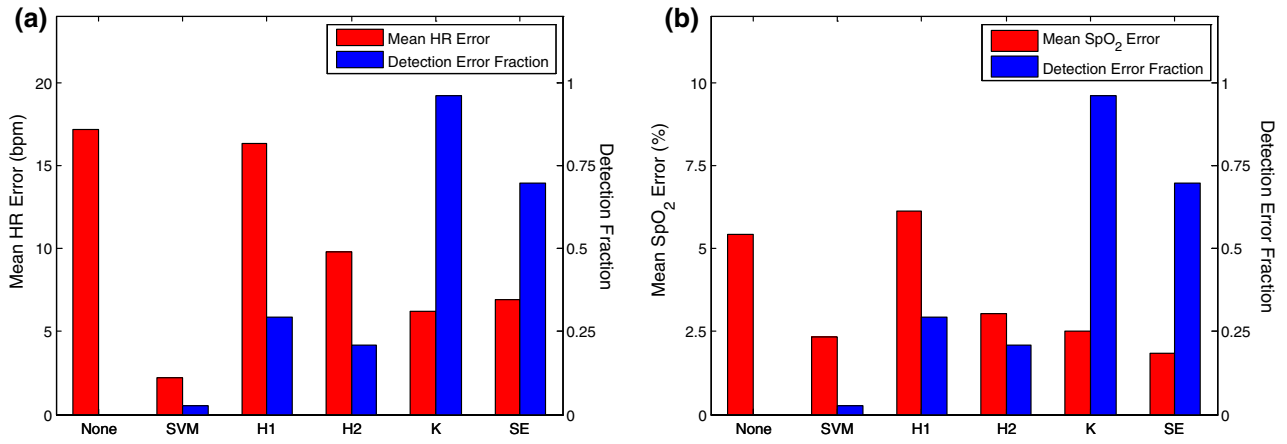


FIGURE 11. Comparison of mean errors and detection error fraction between original signal (labeled "None") and artifact removed signal from five detection methods (SVM, H1, H2, K, and SE). (a) HR error; (b) SpO_2 error.

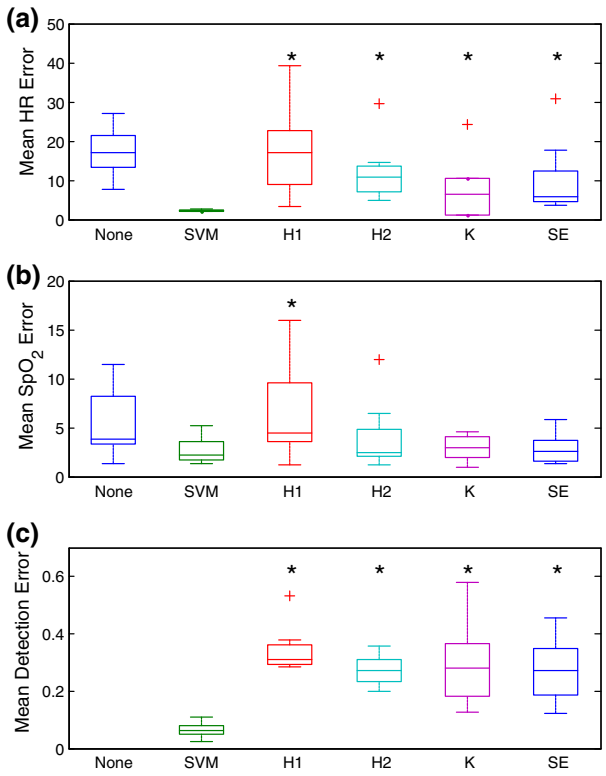


FIGURE 12. Mean error comparison between our SVM algorithm, Hjorth (H1, H2), Kurtorsis and Shanon Entropy (K, SE) parameters. (a) heart rate; (b) SpO_2; (c) detection error for walk/stair-climbing data. The central mark on each box corresponds to the median; the edges of the box correspond to the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually. (*) indicate the mean is significantly different ($p < 0.05$ at 95% CI) between SVM and other methods used for comparison. The *X*-axis labelled "None" in all panels refers to the mean errors when compared to the reference signals without removing the MNA detected segments as identified by any of the five computational methods.

errors obtained from our SVM approach compared with other published methods. For the finger recorded PPG segments, Fig. 10a indicates that the mean classification accuracy is significantly different ($p < 0.05$ at 95% CI) between our SVM and K, SE and H2 methods (except for H1). K, SE, H1 and H2 methods were significantly different from our SVM method for forehead and walking/stair-climbing PPG data. Kurtosis and Shannon entropy based detection methods' low performance may be caused by clean PPG signals' variations in amplitude and pulse. These variations can be induced unintentionally during the measurement due to minimal movement or physiological artifacts which are not severe but vary between test subjects or test conditions. Since H1 is based on estimating central frequency, the H1 method is expected to have high performance as the clean PPG signal gets more stable in the frequency domain. Hence, H1 shows high performance for finger movement when clean PPG signals are measured in the most stable conditions for finger, head and walking/stair-climbing PPG measurements as shown in Fig. 10a, 10b, and 10c. H1 performs the worst for head and walking/stair-climbing PPG data when slight movement and physiological artifacts become more prominent. Similarly, the H2 method which estimates the half of the bandwidth of H1 shows comparable performance for finger PPG signals as shown in Fig. 10a, 10b, and 10c. However, H2 performs worse for head and walking/stair-climbing PPGs. Figure 12a, 12b, and 12c summarizes paired-*t* test results of HR and SpO_2 estimations as well as detection accuracy for walking/stair-climbing data. As shown in Fig. 12a, 12b, and 12c, SVM is significantly different from H1, H2, K, and SE in terms of HR estimation and detection accuracy (see Figs. 11a and

11c), while SpO_2 derived from the SVM method is significantly different from only H1 (see Fig. 12b).

To enhance the decision accuracy of our MNA detection algorithm, we adopted a major vote concept which is widely used in other engineering fields to fuse the decisions from multiple entities.[26] The major vote concept in our algorithm took a role in providing a decision on a target PPG segment after fusing the decisions of neighbor PPG segments as well as that of the target segment. This is based on the observation that the decision on the target segment (clean or corrupted) is highly correlated to those of neighbor segments.

The advantage of our MNA detection algorithm is that it can classify MNA-corrupted PPG from clean PPG in an adaptive manner. Hence, it can be applied to either controlled or daily-activity moving scenarios. The MNA detection algorithm coded with Matlab (2012a) takes only 7 ms on an Intel Xeon 3.6 GHz computer for the 7-s data segment. Hence, the algorithm is real-time realizable especially when it is coded in either C or C++. Our proposed MNA algorithm, when incorporated with the reconstruction algorithm as detailed in our companion paper (ref), can also provide information on whether the signal is clean, relatively noise-free after reconstruction (detected as corrupted but succeeded in reconstruction) or MNA-corrupted (detected as noisy and failed in reconstruction). In conclusion, the potential for the method proposed in this work to have practical applications is high, and the integration of the algorithm described with a pulse oximeter device may have significant implications for real-time clinical applications and especially for ambulatory or smart phone monitoring of vital signs.

ACKNOWLEDGMENTS

This work was supported in part by the US Army Medical Research and Materiel Command (USAMRMC) under Grant No. W81XWH-12-1-0541.

REFERENCES

[1]Andersen, P., and B. Saltin. Maximal perfusion of skeletal muscle in man. *J. Physiol.* 366:233–249, 1985.

[2]Barker, S. J., and N. K. Shah. The effects of motion on the performance of pulse oximeters in volunteers (revised publication). *Anesthesiology* 86:101–108, 1997.

[3]Chang, K.-M., and K.-M. Chang. Pulse rate derivation and its correlation with heart rate. *J. Med. Biol. Eng.* 29:132–137, 2009.

[4]Comtois, G., Y. Mendelson, and P. Ramuka. A comparative evaluation of adaptive noise cancellation algorithms for minimizing motion artifacts in a forehead-mounted wearable pulse oximeter. Engineering in Medicine and Biology Society, 2007 EMBS 2007 29th Annual International Conference of the IEEE, 2007, pp. 1528–1531.

[5]Dash, S., K. H. Chon, S. Lu, *et al.* Automatic real time detection of atrial fibrillation. *Ann. Biomed. Eng.* 37:1701–1709, 2009.

[6]Engelen, M., J. Porszasz, M. Riley, *et al.* Effects of hypoxic hypoxia on O_2 uptake and heart rate kinetics during heavy exercise. *J. Appl. Physiol.* 81:2500–2508, 1996.

[7]Foo, J. Y., and S. J. Wilson. A computational system to optimise noise rejection in photoplethysmography signals during motion or poor perfusion states. *Med. Biol. Eng. Comput.* 44:140–145, 2006.

[8]Ganeshapillai, G., and J. Guttag. Real time reconstruction of quasiperiodic multi parameter physiological signals. *EURASIP J. Adv. Signal Process.* 1–15:2012, 2012.

[9]Gil, E., J. María Vergara, and P. Laguna. Detection of decreases in the amplitude fluctuation of pulse photoplethysmography signal as indication of obstructive sleep apnea syndrome in children. *Biomed. Signal Process. Control* 3:267–277, 2008.

[10]Hjorth, Bo. EEG analysis based on time domain properties. *Electroencephalogr. Clin. Neurophysiol.* 29:306–310, 1970.

[11]Hjorth, Bo. The physical significance of time domain descriptors in EEG analysis. *Electroencephalogr. Clin. Neurophysiol.* 34:321–325, 1973.

[12]Hong Enriquez, R., M. Sautie Castellanos, J. Falcon Rodriguez, *et al.* Analysis of the photoplethysmographic signal by means of the decomposition in principal components. *Physiol. Meas.* 23:N17–N29, 2002.

[13]Hsu, C.-W., C.-C. Chang, and C.-J. Lin. A practical guide to support vector classification. Tech. rep., Department of Computer Science, National Taiwan University, 2003.

[14]Karlen, W., K. Kobayashi, J. M. Ansermino, *et al.* Photoplethysmogram signal quality estimation using repeated Gaussian filters and cross-correlation. *Physiol. Meas.* 33:1617–1629, 2012.

[15]Kim, B. S., and S. K. Yoo. Motion artifact reduction in photoplethysmography using independent component analysis. *IEEE Trans. Biomed. Eng.* 53:566–568, 2006.

[16]Krishnan, R., B. Natarajan, and S. Warren. Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data. *IEEE Trans. Biomed. Eng.* 57:1867–1876, 2010.

[17]Lee, B., J. Han, H. J. Baek, *et al.* Improved elimination of motion artifacts from a photoplethysmographic signal using a Kalman smoother with simultaneous accelerometry. *Physiol. Meas.* 31:1585–1603, 2010.

[18]Lee, H. The periodic moving average filter for removing motion artifacts from PPG signals. *Int. J. Control Autom. Syst.* 5:701–706, 2007.

[19]Lee, J., W. Jung, I. Kang, *et al.* Design of filter to reject motion artifact of pulse oximetry. *Comput. Stand. Interfaces* 26:241–249, 2004.

[20]Li, K., and S. Warren. A wireless reflectance pulse oximeter with digital baseline control for unfiltered photoplethysmograms. *IEEE Trans. Biomed. Circuits Syst.* 6:269–278, 2012.

[21]Li, K., S. Warren, and B. Natarajan. Onboard tagging for real-time quality assessment of photoplethysmograms

acquired by a wireless reflectance pulse oximeter. *IEEE Trans. Biomed. Circuits Syst.* 6:54–63, 2012.
[22]Li, Q., R. G. Mark, and G. D. Clifford. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. *Physiol. Meas.* 29:15–32, 2008.
[23]Mannacio, V., L. Di Tommaso, V. De Amicis, *et al.* Coronary perfusion: impact of flow dynamics and geometric design of 2 different aortic prostheses of similar size. *J. Thoracic Cardiovasc. Surg.* 143:1030–1035, 2012.
[24]Nakajima, K., T. Tamura, and H. Miike. Monitoring of heart and respiratory rates by photoplethysmography using a digital filtering technique. *Med. Eng. Phys.* 18:365–372, 1996.
[25]Naraharisetti, K. V. P., M. Bawa, and M. Tahernezhadi. Comparison of different signal processing methods for reducing artifacts from photoplethysmograph signal. 2011 IEEE International Conference on Electro/Information Technology (EIT), 2011, pp. 1–8.
[26]Olfati-Saber, R., J. S. Shamma. Consensus filters for sensor networks and distributed sensor fusion. 2005 and 2005 European Control Conference CDC-ECC'05 44th IEEE Conference on Decision and Control, 2005, pp. 6698–6703.
[27]Patterson, J. A. C., and Y. Guang-Zhong. Ratiometric artifact reduction in low power reflective photoplethysmography. *IEEE Trans. Biomed. Circuits Syst.* 5:330–338, 2011.
[28]Petterson, M. T., V. L. Begnoche, and J. M. Graybeal. The effect of motion on pulse oximetry and its clinical significance. *Anesth. Analg.* 105:S78–S84, 2007.
[29]Ram, M. R., K. V. Madhav, E. H. Krishna, *et al.* A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter. *IEEE Trans. Instrum. Meas.* 61:1445–1457, 2012.
[30]Ram, M. R., K. V. Madhav, E. H. Krishna *et al.* Use of spectral estimation methods for computation of SpO_2 from artifact reduced PPG signals. Recent Advances in Intelligent Computational Systems (RAICS), 2011 IEEE, 2011, pp. 431–436.
[31]Rusch, T. L., R. Sankar, and J. E. Scharf. Signal processing methods for pulse oximetry. *Comput. Biol. Med.* 26:143–159, 1996.
[32]Sahni, R., A. Gupta, K. Ohira-Kist, *et al.* Motion resistant pulse oximetry in neonates. *Arch. Dis. Child. Fetal Neonatal Ed.* 88:F505–F508, 2003.
[33]Samorodov, Av. Application of a fuzzy integral for weak classifiers boosting. *Pattern Recognit. Image Anal.* 21:206–210, 2011.
[34]Selvaraj, N., Y. Mendelson, K. H. Shelley *et al.* Statistical approach for the detection of motion/noise artifacts in Photoplethysmogram. 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, 2011, pp. 4972–4975.
[35]Sukor, J. A., S. J. Redmond, and N. H. Lovell. Signal quality measures for pulse oximetry through waveform morphology analysis. *Physiol. Meas.* 32:369–384, 2011.
[36]Sweeney, K. T., T. E. Ward, and S. F. Mcloone. Artifact removal in physiological signals—practices and possibilities. *IEEE Trans. Information Technol. Biomed.* 16:488–500, 2012.
[37]Tobin, R. M., J. A. Pologe, and P. B. Batchelder. A characterization of motion affecting pulse oximetry in 350 patients. *Anesth. Analg.* 94:S54–S61, 2002.
[38]Trivedi, N. S., A. F. Ghouri, N. K. Shah, *et al.* Effects of motion, ambient light, and hypoperfusion on pulse oximeter function. *J. Clin. Anesth.* 9:179–183, 1997.
[39]Wijshoff, R. W., M. Mischi, J. Veen, *et al.* Reducing motion artifacts in photoplethysmograms by using relative sensor motion: phantom study. *J. Biomed. Opt.* 17:117007, 2012.
[40]Yan, Y. S., C. C. Poon, and Y. T. Zhang. Reduction of motion artifact in pulse oximetry by smoothed pseudo Wigner–Ville distribution. *J. Neuroeng. Rehabil.* 2:3, 2005.
[41]Yu, C., Z. Liu, T. Mckenna, *et al.* A method for automatic identification of reliable heart rates calculated from ECG and PPG waveforms. *J. Am. Med. Inform. Assoc.* 13:309–320, 2006.