AD_____

Award Number: W81XWH-12-1-0541

TITLE: Wearable Wireless Sensor for Multi-Scale Physiological Monitoring

PRINCIPAL INVESTIGATOR: Ki Chon

CONTRACTING ORGANIZATION: Worcester Polytechnic Institute
Worcester, MA 01609

REPORT DATE: October 2013

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
October 2013	Annual	25September2012-24September2013

4. TITLE AND SUBTITLE

Wearable Wireless Sensor for Multi-Scale Physiological Monitoring

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-12-1-0541

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Ki Chon
Yitzhak Mendelson

E-Mail: kichon@wpi.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Worcester Polytechnic Institute
100 Institute Road
Worcester, MA 01609

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
One of the aims of Year 1 of the project was to develop a prototype multi-channel pulse oximeter that can be used to collect physiological data from multiple body locations to combat motion artifact contamination. Specifically, the aim was to investigate if a motion artifact-free signal can be obtained in at least one of the multi-channels at any given time. Towards this aim, we have developed a prototype 6-photodetector reflectance-based pulse oximeter and preliminary results show that good signals can be obtained in one of the multi-channels at any given time. A conference proceedings paper describing detailed results is provided with the annual report. The second major aim of the project was to develop a motion and noise detection algorithm and a separate algorithm for the reconstruction of the motion and noise contaminated portion of the data. For detection of motion and noise artifacts, we have successfully developed an accurate and real-time realizable algorithm. Moreover, our detection algorithm is able to discriminate between severely and moderately corrupted data. The sensitivity and specificity of detecting severely corrupted data was found to be 98.7% and 92.9%, respectively. The sensitivity and specificity of detecting moderately corrupted data was found to be 94.4% and 90.4%, respectively. Comparison of our detection algorithm to some of the gold standard algorithms showed that our algorithm is far superior to the latter methods. For reconstruction of the motion and noise corrupted data segments, we have successfully developed an algorithm which significantly outperforms a gold standard method. It was found that our reconstruction algorithm consistently provides accurate estimations of the reconstructed motion and noise artifact contaminated signal's heart rates and oxygen saturation values when verified with the reference signal. However, the gold standard method's derived heart rates and oxygen saturation values significantly differ from the clean reference signal. Drafts of two manuscripts describing the detection and reconstruction algorithms are provided with the annual report.

15. SUBJECT TERMS
Motion and noise artifact detection and reconstruction; multi-channel pulse oximeter sensor

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU		19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Introduction

US combat experience has demonstrated that acute hemorrhage and subsequent hemodynamic decompensation (*shock*) account for about 50% of the deaths on the battlefield. Realizing the limits of current pre-symptomatic diagnosis and treatment capabilities on the battlefield, a reliable non-invasive physiological sensor and diagnostic algorithms that provide *clinical decision support* for early hemorrhage diagnosis and facilitate *remote assessment (triage)* for medical evacuation of the highest-priority combat casualties remains one of the primary objectives for Combat Casualty Care. Moreover, a sensor that can monitor the status of uninjured soldiers suffering from physiologic stress such as dehydration, may help optimize performance in the field. To address this challenging deficiency and reduce the medical logistics burden in the field, we propose to significantly enhance the current capabilities of our prototype wearable, pulse oximeter-based, physiological status sensor so that when donned by military personnel it will acquire and wirelessly transmit in real-time ***seven*** algorithmically derived vital physiological indicators: heart rate, perfusion index, oxygen saturation, respiratory rate, autonomic nervous system dynamics, arrhythmia detection, and blood volume loss. This critical information will be captured, analyzed and displayed on a hand-held monitoring device carried by a medic. Any change in a soldier's physiological status including early warnings of impending hemorrhagic shock or severe dehydration will alert the individual responsible for monitoring soldiers' conditions so that appropriate timely intervention may be taken. Our sensor will be applicable in at least two different scenarios: remote combat triage and bedside (point of care) monitoring. For the latter scenario, our recently developed smart phone technology which uses images processed from a fingertip to derive seven physiological parameters using our algorithms is also applicable. Our ***single*** sensor (either wearable pulse oximeter itself or pulse oximeter-like information derived from a smart phone) combines significant advancements in both sensors and patent pending detection algorithms that are especially applicable for accurate and early detection of hemorrhage on spontaneously breathing subjects, a feat that has not been achieved to date.

Body

Task 1a: Develop multi-channel pulse oximeter (MCPO) sensor prototypes to acquire PPG signals and combat motion artifact contamination.

The following Fig. 1 shows the fabricated MCPO sensor mounted inside a wearable forehead-mounted housing configuration (left) and the fully populated and functional printed circuit boards for the 6 LEDs sensor version (right). The top PCB shows the micro on/off activation switch (left), rechargeable battery (center) and USB receptacle (right).

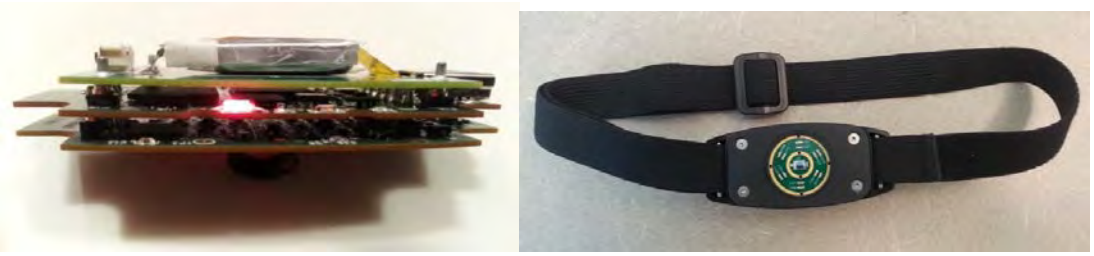


Fig. 1

The following items have been completed for the 6 LEDs sensor version:

- Printed circuits board (PCB) fabrication and population with SMT components
- PCB testing to ensure that design specifications were met

- Testing functionality of sensor housing
- Graphical user interface (GUI) for processing user input
- Integration of plotting library routines
- File I/O procedures for saving and plotting of log files
- Integration and speed testing of USB interface
- Integration of MSP430 USB stack
- Procedures for timer peripheral management
- Procedures for low power mode activation and wake-up
- Routines for start-up state detection
- Flash memory management routines
- ADC interface routines
- DAC interface routines

A parallel effort has begun to design the transmission based version of the proposed MCPO sensors that will enable us to acquire data from the earlobe and fingertip. Each sensor is based on a cluster of 3 adjacent PDs and a pair of R and IR LEDs, providing simultaneously acquired data from 3 independent pulse oximeter channels. To date, we have completed the design of the main PCB that will fit inside a wrist or upper arm mounted enclosure as well as the PCBs that will fit inside the finger and ear sensors. We expect to complete the PCB fabrication of these PCBs by the end of this month.

A separate effort involved the design of the enclosures to house the finger and earlobe versions of the MCPO sensors. Fig. 2 depicts the rendering of the earlobe (left) and finger (right) sensors. We expect to complete the fabrication of these enclosures by the end of this month.

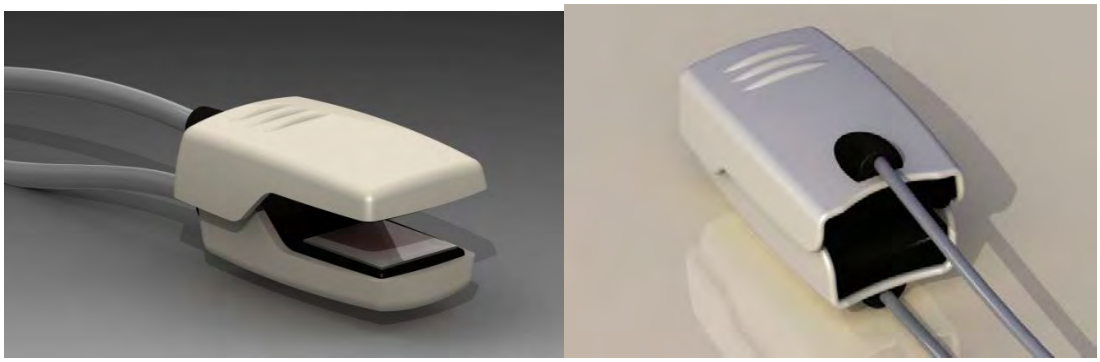


Fig. 2

The following items need to be completed for the 3PD earlobe and fingertip transmission sensors:

- Printed circuits board (PCB) fabrication and population with SMT components
- Firmware development
- PCB testing to ensure that design specifications were met
- Testing functionality of sensor housing
- Graphical user interface (GUI) for processing user input
- File I/O procedures for saving and plotting of log files

Task 1c: Design a utility program and GUI to display vital physiological information received from the MCPO

An updated version of the management software and GUI which allows the user to capture the data and view the signals from both the 6 photodetector (PD) or 6 LED sensors on the PC screen has been completed. In this software version, the real-time heart rate estimation has been implemented to control the status LED. The drop-down menu shows the unique device ID and firmware version with the given serial number, and firmware version. The status LED on the PCB above is flashing Red when the firmware boots properly. Once the PC software is run, and the "Connect" button is used to connect to the device, the status LED is flashing Green

instead of red. The plots on the "Monitor Sensors" tab show live sensor data captured, scrolling right-to-left. The "Auto Scale" button is used to zoom and pan the plots to make the sensor data fill the available plot space on the PC screen. A feature allowing the operator the ability to adjust the intensity levels of each LED manually or automatically has been added. The following Fig. 3 shows a screen-shot of the current appearance of the updated management software, which will continue to be refined as testing commences. Each trace depicts the PPG signals acquired simultaneously from 3 different sampled LED channels. Additional traces (not shown) can display in real-time the output of the on-board 3D-accelerometer. Laboratory test are currently being conducted to verify the performance of the 6 PD and 6 LED sensors to make sure that the sensors are ready for field studies.

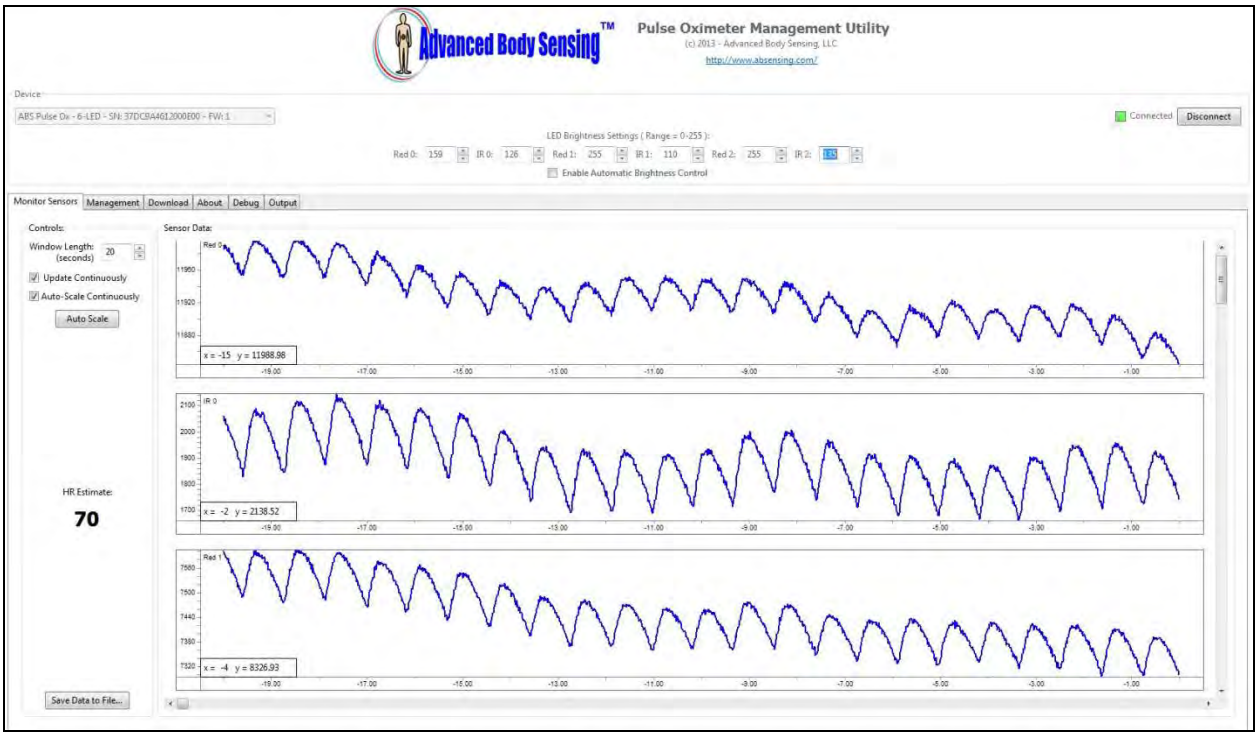


Fig. 3

Task 2a: Design and test motion & noise artifact detection and removal algorithms tailored to unwounded and wounded solider scenarios. Also develop smart phone application for extracting vital signs including blood loss.

We have attached a draft of two manuscripts: the first paper is for detection of motion and noise artifacts (MNA), and the second paper describes a method for reconstruction of the corrupted data segments. The detection of MNA algorithm uses the support vector machine algorithm for classification between clean and MNA-corrupted data segments. Our MNA detection algorithm not only detects severely corrupted data but it can also classify a data segment that is only moderately corrupted. Based on our experience studying many corrupted data segments including new data collected in this past year, we have learned that moderately-corrupted data can have either poor quality photoplethysmographic (PPG) signal morphology or large variability in heart rates and their amplitudes. The presence of both poor quality PPG waveforms and high heart rate & amplitude variability is most often associated with severely MNA-corrupted data. Distinguishing between severely corrupted data and marginally contaminated data is significant because this allows us to reconstruct only those data segments that are detected to be the latter. It is our opinion that the severely corrupted data segments should not be reconstructed since the data quality is so poor that no algorithm can reliably reconstruct the true dynamics of the signal. Hence, only those segments that are deemed to be moderately corrupted will be reconstructed using our singular spectrum analysis (SSA) algorithm. Our third

quarter's reported results were based on a limited dataset and less representative of realistic MNA scenarios. Hence, this past quarter, we have further examined the efficacy of our detection and reconstruction algorithms using more diverse sets of experimental data. The new PPG data from 10 healthy subjects was taken while the subjects walked and climbed stairs. With the new data, we found the sensitivity of the SVM detection algorithm to be 98.7% for detecting severely MNA corrupted data segments while the specificity was found to be 92.9%. Moreover, the sensitivity and specificity of the identification of moderately MNA corrupted data segments was found to be 94.4% and 90.4%, respectively. The reconstruction of the moderately MNA contaminated data segments remains accurate when the estimated HR and SpO2 values are compared to the reference (clean) data. These results are detailed in the attached manuscript of the Appendix section. These manuscripts will be submitted in early November of 2013.

Task 2b: Determine usable PPG data segments for blood loss and vital sign calculations if the data have been determined to contain insignificant motion & noise artifacts

We propose a quantitative approach based on time-frequency spectral analysis to determine usable PPG data (i.e., negligibly corrupted) among those data that have been designated to be corrupted. The method involves extraction of the amplitude values at the HR frequency band (AM_{HR}) using the time-frequency spectral technique on PPG data segments. Our criterion for usable data is to quantitatively determine if the artifacts are severe enough to compromise the fidelity of the extracted AM_{HR} of the PPG signal as this value is used to determine heart rate . If the data are severely corrupted, then we can discard them. We can do this because there exists a sufficient number of clean segments in each PPG recording that noise-contaminated segments can be discarded, thereby increasing the specificity of the results. We tested the efficacy of our computational approach on multi-site PPG data containing involuntary artifacts recorded under clinical settings (10 subjects) and on finger PPG data containing controlled voluntary movements recorded in a laboratory setting (14 subjects). For artifact detection, the combined metrics of SE and kurtosis were found to be 99.0%, 94.8% and 93.3% accurate in detecting artifacts simultaneously recorded from ear, finger and forehead PPGs, respectively, obtained in a clinical setting. The severity of noise was found to be negligible in 44.5%, 33.1% and 12.5% of the corrupted PPG segments obtained from forehead, finger, and ear sites so those segments are considered usable. Usable data was determined by applying the AM_{HR} median significance bounds (mean ± 2*SD). Thus, if a data segment was deemed to be contaminated, and if its AM_{HR} median values fell within statistical significance limits, then we classified it as a usable segment. The overall computation time for artifact detection and rejection stages was 0.742 seconds on average using Matlab® for one 60-second PPG data segment. This task is complete and further validation will be made using the newly-acquired data in years 2 and 3 of the grant.

Key Research Accomplishments

- For the sensor development, a prototype fabrication of a 6-photodetector reflectance-based sensor has been completed and preliminary results show good signals can be obtained.

- Developed a motion and noise artifact (MNA) detection algorithm
    - We can discriminate between severe and moderately MNA corrupted data.
    - For severely corrupted data: sensitivity and specificity are 98.7% and 92.9%, respectively.
    - For moderately corrupted data: sensitivity and specificity are 94.4% and 90.4%, respectively.

- Developed a signal reconstruction algorithm for the MNA corrupted data
    - Compared our algorithm to the independent component analysis (ICA) as the latter is the current gold standard method.
    - Our algorithm can reliably reconstruct MNA contaminated signals even for signal-to-noise ratios as low as -15 dB as verified via simulated data.
    - Our algorithm is far superior to the current gold-standard reconstruction algorithm, the ICA, as evidenced by:
        - Our algorithm consistently provides accurate estimation of the reconstructed MNA contaminated signal's heart rates and oxygen saturation values when verified with the

reference signal. The ICA's derived heart rates and oxygen saturation values significantly differ to the reference signal.

Reportable Outcomes

Journal Manuscripts

1: Chong et al., A real-time method for detection of motion and noise corruption in photoplethysmogram – Part I: Motion artifact detection, to be submitted.
2: Salehizadeh et al., Photoplethysmograph signal reconstruction based on a novel motion artifact detection-reduction approach – Part II: Motion artifact removal, to be submitted
3: Y. Nam, J. Lee and K.H. Chon, Respiratory rate estimation from the built-in cameras of smartphones and tablets, Accepted pending revision, Annals of Biomedical Engineering.

Conference Proceedings

1: Y. Mendelson, D.K. Dao and K.H. Chon, Multi-channel pulse oximetry for wearable physiological monitoring, Body Sensor Network, Boston, MA, 2013 (invited paper).
2: J. Chong, D.D. McManus and K.H. Chon, Arrhythmia discrimination using a smartphone, Body Sensor Network, Boston, MA, 2013 (selected as the 2nd best conference paper).

Conclusion

In Year 1 of our project, we have developed a prototype multi-channel pulse oximeter that can be used to collect physiological data from multiple body locations (forehead, finger and ear lobe) to combat motion-artifact contamination. This multi-channel pulse oximeter will be used to collect clinical data in Year 2 of the project at the Emergency Department of the University of Massachusetts Medical Center. Concurrent to the sensor development, we have successfully developed robust algorithms for the detection followed by reconstruction of the identified motion and noise corrupted data segments. In Year 2, we will enhance the reconstruction algorithm to include time-varying dynamics to account for possible sudden changes in heart rate and oxygen saturation values. Both the sensor and algorithms will be thoroughly tested and further refined, if needed, using the data collected in Year 2 of the project. We are currently on schedule for all milestone tasks originally proposed in our grant proposal.

References

None

Appendix

3 Manuscripts and 2 conference proceeding articles are provided.

Supporting data

Not applicable

A Real-Time Method for Detection of Motion/Noise Corruption in Photoplethysmograms – Part I: Motion Artifact Detection

***Abstract*—Motion and noise artifact have been a serious obstacle in utilizing photoplethysmogram (PPG) signals in real-time monitoring of vital signs. We present an accurate and comprehensive motion artifact detection method providing quantitative severity indication of motion artifacts. 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 ratio, and (4) mean standard deviation of pulse shape. We adopt a diversity technique to enhance the discrimination between corrupted and clean PPG signals. The algorithm was verified on PPG data segments recorded during laboratory controlled experiments and motion artifacts encountered during typical daily activities. To quantify the severity of motion artifacts, the proposed algorithm was applied to 7-seconds successive PPG segments. The proposed detection algorithm was evaluated in terms of motion artifact detection accuracy and accuracy of heart rate (HR), and oxygen saturation (SpO_2) estimations. For both laboratory controlled and daily activity data, the detection accuracy, sensitivity, and specificity were 94%, 97%, and 92%, respectively.**

***Index Terms*—motion and noise artifacts, photoplethysmography, support vector machine.**

I. Introduction

Photoplethysmography (PPG) is a non-invasive and low cost tool to continuously monitor blood volume changes in peripheral tissues. The PPG is useful since it is widely used to monitor heart rate (HR), arterial oxygen saturation (SpO_2), and can be used to measure respiratory rate [1]. However, motion artifacts are known to distort PPG recordings, causing erroneous estimation of HR and SpO_2. There are three distinct sources of artifacts that can impair PPG recordings: environmental, physiological and experimental artifacts which are due to power interference surrounding the body, other physiological signals, and experimental conditions [3-7]. Motion artifacts (MAs) are difficult to filter out compared to environmental and physiological artifacts since they do not have a predetermined frequency band and their spectrum often overlaps with that of the desired PPG signal. Therefore, it is crucial to develop a reliable PPG processing algorithm that is resilient to motion artifacts.

Motion artifacts 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) variation in the optical coupling between the sensor and the skin [8-11]. Various approaches to mitigate motion artifacts by improving sensor attachment have been proposed [12-13]. However, these design improvements do not guarantee a significant reduction of motion artifacts. Algorithm-based MA reduction methods were also proposed. These include time and frequency domain filtering, power spectrum analysis, and blind source separation techniques [18-28]. However, these have high computational complexity since they operate even on clean portions of the PPG signal where MA reduction is not needed. Hence, accurate MA detection, which precedes MA reduction and divides PPG recordings into clean and corrupted portions, is essential to enhance computational efficiency [29].

Motion artifact detection methods are mostly based on a signal quality index (SQI) which quantifies the severity of the MA. Some approaches quantify SQI using waveform morphology [30] or filtered output [31], while other derive SQI with the help of additional hardware such as accelerometer and electrocardiogram sensing [14-17]. Statistical measures, such as skewness, kurtosis, Shannon entropy, and Renyi's entropy, have been shown helpful in determining SQI [34, 35]. 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 [32], this approach considers limited and controlled types of motions. Moreover, previous studies provide no information about the start and end time of MA corrupted portions, which is important for efficient PPG removal and possible reconstruction. The authors are not aware of any detailed studies providing representative and comprehensive features distinguishing clean from corrupted PPG signals, including precise demarcations of the start and end points of motion corrupted portions in a PPG signal.

In this paper, we propose an accurate and comprehensive MA detection algorithm which (1) specifies the start and end points of MA in the PPG data and (2) quantifies the severity of the MA. We first introduce time and frequency-domain parameters

quantifying MA in the recorded PPG signal. We then take their statistical measures into consideration as input variables for a machine learning-based MA detection algorithm. A SVM is used as a machine learning algorithm for training and testing of MA. Our MA detection algorithm is self-trained by the SVM with clean and corrupted PPG data sets, and then the trained detection algorithm tests the unknown PPG data. We tested the efficacy of our proposed algorithm on PPG data sets obtained from the finger and forehead in both laboratory controlled studies and daily-life activities with various types of motions. This paper is organized as follows: In Section II, we describe the PPG recording protocol and preprocessing. Time- domain parameters quantifying motion artifacts are also introduced in Section II. In Section III, our proposed SVM-based motion artifact detection methods are described. Section IV presents the performance of the proposed artifact detection method in terms of classification accuracy. Section IV concludes this paper.

II. Materials and Method

A. Experimental Protocol and Preprocessing

PPG signals were obtained from a custom reflectance-mode prototype forehead sensor or a commercial finger pulse oximeter. The forehead PPG signals were sampled at 80 Hz while the finger signals were sampled at 100 Hz. Four sets of data were collected 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 the first experiment, 11 healthy volunteers were asked to wear the forehead reflectance pulse oximeter along with a reference Masimo Radical (Masimo SET®) finger type transmittance pulse oximeter. The PPG signals recorded from the forehead sensor and reference HR readings were acquired simultaneously. HR and SpO_2 signals were acquired by a PC at 80Hz and 1Hz, respectively. After baseline recording for 5 minutes without any movement, motion artifacts were induced by spontaneous movements of the subject's head in both horizontal and vertical directions while the right middle finger with the sensor attached to the Masimo pulse oximeter was kept stationary. Subjects were instructed to introduce motion artifacts for specific time intervals varying from 10 to 50% within a 1 minute segment. For example, if a subject was instructed to perform left-right movements for 6 seconds, a 1 min segment of data would contain 10% noise.

The second dataset consisted of finger recorded PPG signals obtained from the same 9 healthy volunteers in an upright sitting position using a reflection type PPG transducer (TSD200) and a biopotential amplifier (PPG100) with a gain of 100 and cut-off frequencies of 0.05-10Hz. The MP1000 (BIOPAC Systems Inc., CA, USA) was used to acquire finger PPG signals at 100 Hz. Two pulse oximeters were placed on the index and middle finger, simultaneously. After baseline recording for 5 minutes 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 first dataset, motion was induced at specific time intervals corresponding to 10 – 50% corruption duration in 1 minute segment. Such controlled movement was repeated five times per subject.

The third dataset consisted of data recorded from 9 subjects with PPG signals recorded simultaneously from the subjects' forehead using our custom sensor and reference ECG, HR and SpO_2 readings obtained from a Holter ECG monitor (Rozinn RZ153+) at 180Hz and a Masimo Rad-57 pulse oximeter at 0.5Hz, respectively. The reference pulse oximeter provided HR and SpO_2 data measured from the subject right index finger, which were held against the subject's chest to minimize motion artifacts. PPG and ECG signals were recorded while subjects were walking straight and climbing stairs for 45 min.

The fourth dataset simulated motion was generated by the introduction of white noise. PPG data were filtered by a 6th order infinite impulse response (IIR) band pass filter with cut-off frequencies of 0.5 Hz and 12Hz. Zero-phase forward and reverse filtering was applied to account for the non-linear phase of the IIR filter. After these preprocessing, our MA detection algorithm was used to compute the four time-domain parameters as described in Sections II.B and II.C, respectively.

B. Parameters from PPG Signals

The following parameters were selected since they represent the variability present in corrupted PPG signals as shown in Fig. 1.

1) Standard deviation of peak-to-peak interval ($STD_{\text{PPint,}}$): The $STD_{\text{PP-int, }n}$ at the n^{th} segment is defined by:

$$STD_{\text{PP-intv, }n} = \frac{1}{N-1}\sum_{i=1}^{N}\left(D_{n,i} - \overline{D_n}\right) \tag{1}$$

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

2) Standard deviation of peak-to-peak amplitude (STD_{PPamp}): The $STD_{\text{PP-amp, }n}$ at the n^{th} segment is defined by:

$$STD_{\text{PP-intv, }n} = \frac{1}{N-1}\sum_{i=1}^{N}\left(A_{n,i} - \overline{A_n}\right) \tag{2}$$

where $A_{n,i}$ is the peak amplitude at the i^{th} pulse of the n^{th} segment and $\overline{A_n}$ is the mean peak-to-peak interval of the n^{th} segment. The $A_{n,i}$ is defined by the difference between the i^{th} peak and the forthcoming $(i+1)^{\text{th}}$ trough amplitudes.

3) Standard deviation of systolic and diastolic ratio (STD_{SD}):

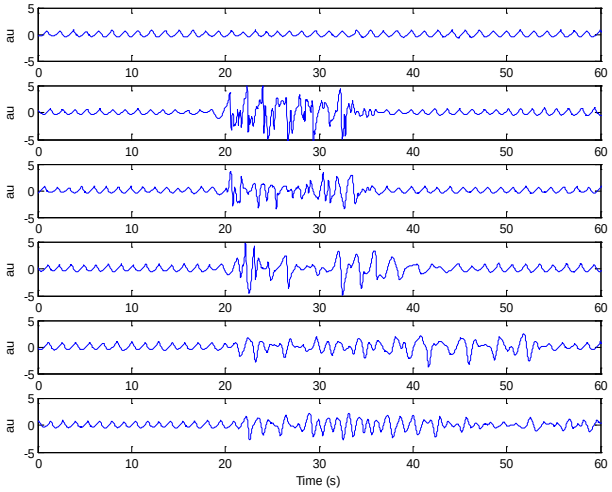


Figure 1. A representative clean forehead-PPG signal recorded during voluntary motion artifact conducted in a laboratory setting (1st row). The mixed (up-down and left-right) movement of the forehead to which the PPG probe was attached for predetermined time interval induced 10% to 50% noise (2nd – 6th row) within a 60s PPG segment.

The $STD_{\mathrm{SD},\,n}$ at the n^{th} segment is defined by:

$$STD_{\mathrm{SD},\,n} = \frac{1}{N-1}\sum_{i=1}^{N}\left(R_{SD,n,i} - \overline{R_{SD,n}}\right) \tag{3}$$

where $R_{\mathrm{SD},n,i}$ is the corresponding systolic and diastolic time interval ratio at the i^{th} pulse of the n^{th} segment and $\overline{R_n}$ is the mean systolic and diastolic time interval ratio of the n^{th} segment. The $R_{\mathrm{SD},\,n,i}$ is calculated by

$$R_{\mathrm{SD},\,n,i} = \left(T_{\mathrm{trough},\,n-1,i} - T_{peak,n,i}\right) / \left(T_{peak,\,n,i} - T_{\mathrm{trough},\,n-1,i}\right) \tag{4}$$

where $T_{trough,\,n,i}$ denotes the trough (or lowest point) at the i^{th} pulse of the n^{th} segment.

4) Mean-standard deviation of pulse shape ($STD_{\mathrm{wav},\,n}$): To derive pulse shape, we take N_{samp} sample points of a pulse. The $STD_{\mathrm{wav},\,n}$ at the n^{th} segment is defined by:

$$STD_{\mathrm{wav},\,n} = \mathbf{E}[STD_{\mathrm{wav},n,m}] \tag{5}$$

where the mean-standard deviation of pulse shapes is derived by calculating the standard deviation at each sample point and taking their average. The $STD_{\mathrm{wav},\,n,m}$ is calculated by:

$$STD_{\mathrm{wav},n,m} = \frac{1}{N-1}\sum_{i=1}^{N}\left(q_{n,i}(m) - \overline{q_n(m)}\right) \tag{6}$$

where $q_{n,i}(m)$ is the m^{th} pulse sample at the i^{th} pulse of the n^{th} segment and $\overline{q_n(m)}$ the mean at the m^{th} pulse sample of the n^{th} segment.

III. SVM-BASED DETECTION OF MOTION/NOISE ARTIFACTS

A. 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

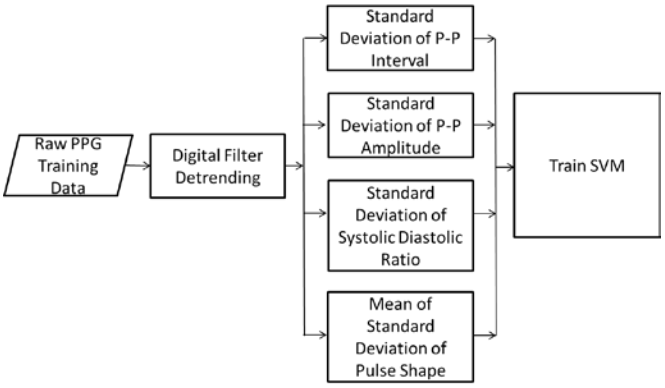


Figure 2. Training phase of the proposed SVM-based motion detection algorithm. Four time-domain features corresponding to (1) standard deviation of peak-to-peak intervals (2) standard deviation of peak-to-peak amplitudes (3) standard deviation of systolic and diastolic interval ratio, and (4) mean standard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

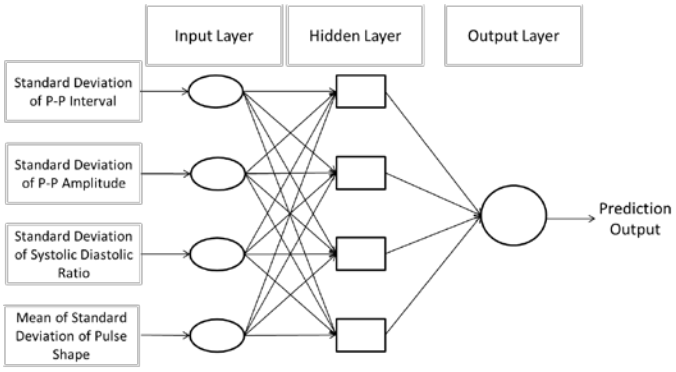


Figure 3. Test phase of the proposed SVM-based motion detection algorithm. The hidden layers correspond to kernel function of the SVM. The function between hidden layer and output layer is a linear operator.

robustness to noise [18]. There are two SVM operation phases: training and test phases. In the training phase, the SVM makes a decision boundary. In the test phase, the SVM classifies unknown input PPG segments into clean or corrupted using the decision boundary.

1) Training phase: A flow chart of the training phase in the SVM-based motion artifact detection algorithm is shown in Fig. 2. The SVM takes the parameter values of clean and corrupted PPG segments as a training data set and finds the support vectors among the training data set which maximize the margin (or the distance) between different classes, and finally builds a decision boundary. We consider a soft-margin SVM which can set the boundary even when the data sets are mixed and cannot be separated.

If the estimated decision is different from its known label, the decision is regarded as a *training error*. In the soft-margin SVM algorithm, slack variables are introduced to minimize the training error with maximizing the margin. Soft-margin SVM uses the following equation to find the support vectors.

$$\text{Minimize} \quad d(\mathbf{w}_s) = \frac{1}{2}\langle \mathbf{w}_s, \mathbf{w}_s \rangle,$$

$$\text{Subject to} \quad T_{sv}\left(\langle \mathbf{w}_s, \mathbf{y}_{sv} \rangle + b_s\right) \geq 1 = \delta_{sv} \quad \text{for } sv = 1,2,\ldots,N, \quad \text{and } \delta_{sv} \geq 0 \qquad (7)$$

where $\mathbf{w}_s$ and $\mathbf{y}_{sv}$ are the weight vector and the sv^{th} input vector data, respectively. T_{sv} is the sv^{th} target variable, b_s is the bias, δ_{sv} is the slack parameter, and C is the regulation parameter. $\langle \mathbf{w}_s, \mathbf{y}_{sv} \rangle$ is the inner product operation of $\mathbf{w}_s$ and $\mathbf{y}_{sv}$. The decision boundary F_{sv} is derived as

$$F_{sv} = \langle \mathbf{w}_s^*, \mathbf{y} \rangle + b_s^* = 0 \qquad (8)$$

where $\mathbf{w}_s^*$ is the weight factor and b_s^* is the bias obtained from Eq. (7) and $\mathbf{y}$ is the input point.

By transforming the $\mathbf{y}_{sv}$ term to $\mathbf{y}_{sv} \rightarrow \Phi(\mathbf{y}_{sv})$, the non-linear SVM can be transformed to a linear SVM. For nonlinear SVM, Eq. (7) is modified as

$$T_{sv}\left(\langle \mathbf{w}_s, \Phi(\mathbf{y}_{sv}) \rangle + b_s\right) \geq 1 \qquad (9)$$

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

$$K_s(\mathbf{y}_{sv}, \mathbf{y}_{sv'}) = \langle \Phi(\mathbf{y}_{sv}), \Phi(\mathbf{y}_{sv'}) \rangle \qquad (10)$$

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

2) Test phase: Fig. 3 shows a flow chart of the test phase in the SVM-based motion detection algorithm. We define by *segment* a unit window dividing an input PPG signal stream, and derive parameters in each PPG segment to examine if the segment is corrupted by motion artifact or not. The SVM takes unknown input PPG segments as input and determines whether they are clean or corrupted segments.

B. Enhancement of MA Detection by Diversity

To enhance MA detection probability, the proposed algorithm incorporates multiple decisions on a set of neighbor segments in deciding whether a "target" segment is clean or corrupted. Neighbor segment is defined as the segments surrounding a target segment within ± $T_{neighbor}$ seconds. Decision on the neighbor segments is highly probable to be the same as the decision on a target segment since the pulses in neighbor segments are to a great extent common to the target segment.

As shown in Fig. 4, the algorithm gathers the decisions of neighbor segments as well as target segment and makes a final decision regarding the target segment based on a majority vote concept.

IV. Results

We evaluated the performance of the MA 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-second segments. We compared the proposed algorithm with four recently published MA detection algorithms based on kurtosis (K), Shannon entropy

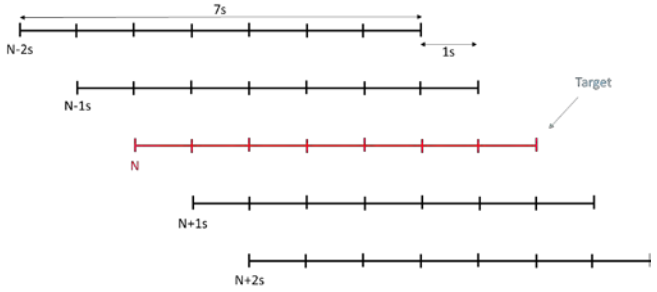


Figure 4. Enhancement of MA detection by diversity. Neighbor segments are the segments surrounding a target segment within ± *2* seconds. Decisions on the target segment are based on a majority vote from the decisions of neighbor segments as well as the one of the target segment (red).

(SE), Hjorth 1 (H1), and Hjorth 2 (H2) metrics [49, 50]. As performance metrics, we considered classification accuracy, sensitivity and specificity. We also investigated mean HR and SpO_2 errors as well as detection error fraction.

A. Reference: Clean vs. corrupted

The following are criteria which we adopted to reference PPG segments (clean or corrupted) in each experiment. A visual reference was excluded to avoid subjective decisions by

visual inspectors. Instead, we used objective information including 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.

- *Laboratory controlled data (Forehead and finger)*

- If 85% of a segment is outside of [$T_{corr,start}$, $T_{corr,end}$], the segment was considered *corrupted*. Otherwise, the segment was referenced as *clean*.

- If SpO_{2PPG} deviates by 10 % from the mean of SpO_{2PPG} in a segment, then the segment was referenced as corrupted.

- Successive difference, |diff($HR_{PPG}(i+1)$- HR_{PPG})|, from PPG signals is larger than 20 bpm for at least one pulse during a segment, then the segment was referenced as corrupted.

- *Daily activity data (Walking and stair-climbing)*

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

- Successive difference, |diff($HR_{PPG}(i+1)$- HR_{PPG})|, from PPG signals is larger than 20 bpm for at least one pulse during a segment, then the segment was referenced as corrupted.

- If $|HR_{ECG} - HR_{PPG}| < 5$ bpm during 85 % of a segment, the segment was considered clean. Otherwise, the segment was referenced as corrupted.

Table I describes the number of clean and corrupted PPG segments for each motion type used in the experiment.

B. Classification Accuracy

As input parameters for the SVM classifier, we considered the standard deviations of (1) the peak-to-peak interval and (2) peak-to-peak amplitude in a 7-second PPG segment. For example, a sample forehead PPG signal and its corresponding parameters are given in Fig. 5a and Figs. 5b-e, respectively. The sample signal is corrupted from t=56 to t=85 seconds.

TABLE I
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

Parameter values calculated segment-by-segment are presented in Figs. 5b-5e. Corrupted PPG segments between 56-84s have larger parameter values compared to clean segments between 1-56s and 84-112s.

The SVM decision boundary is obtained from the training data set. To lower computational complexity, a linear kernel was considered for the SVM in the experiment. We optimized regularization parameter value (C) of the linear kernel SVM in terms of minimizing the training error rate. We adopted a 9-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 [20].

Fig. 6a shows clean (*upper*) and corrupted (*lower*) forehead PPG segments and Figs. 6b-c gives their corresponding parameter values with SVM boundaries (green line). Fig. 7 shows classification results by the SVM boundaries obtained from Fig. 6. The SVM performance showed a 96.1% accuracy, 97.7% specificity, and 93.1% sensitivity. For finger PPG segments, we obtained a similar classification performance of 95.5% accuracy, 98.1% specificity, and 92.4% sensitivity. Table II presents C for finger, forehead, walking, and stair-climbing data. To evaluate the sensitivity of our MA detection algorithm, we added white noise to the measured PPG signals for varying the SNR levels. As shown in Fig. 8, the PPG signals with a SNR below -10 dB starts to become corrupted. For a SNR of -20 dB, every segment was detected as corrupted.

C. Performance Comparison of MA detection Algorithms

Our algorithm was compared with other artifact detection methods such as H1, H2 , K and SE. For a fair comparison, all the detection method were converted to 7s based decision which determined that a segment was clean if more than 85 % of that segment is clean. Similarly, K and SE detection output based on a 60s window segment and a 10s shift were converted to a 7s decision criterion. Figs. 9a-c compare the medians and 75th and 25th percentiles of the detection accuracy, sensitivity, and specificity for all five detection methods of the Finger, Head and Walking/Stair Climbing data sets. In general, our SVM method consistently yielded higher quality results with a mean accuracy of 94%, sensitivity of 97%, and a specificity of 92%; whereas other methods showed fluctuations depending on which datasets were utilized. In the Finger recorded data, H1 yielded a slightly higher accuracy due to higher sensitivity, but the detection specificity was lower.

D. HR and SpO_2 Estimation

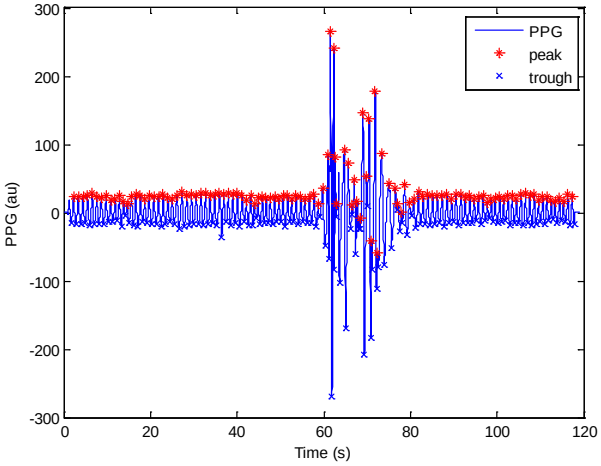
In Fig. 10, we compared mean HR and SpO_2 errors between the estimated values obtained from the PPG and reference signals. The errors were computed from the original signals and the signals where positive artifact detection segments were removed with speculation that errors in the artifact-free signals would be much lower. Our algorithm and the other four detection methods were used separately in the artifact removal process for comparison. However, the mean errors provide only quality assessment. In some instances, the errors were low but they were computed from a small subset of the clean segments. Thus, detection error fraction was needed as quantity assessment to accompany the mean HR and SpO_2 errors. Low values on both of the errors would be reflected from an effective artifact detection algorithm. Fig. 10a-b show the mean errors and detection error results from the Walking/Stair Climbing dataset. Our algorithm yielded lowest errors among the other methods we compared. The SE based detection method showed a lower mean SpO_2 error than our algorithm but its detection error was very high (>70%), indicating that its error was computed based on only 30% of clean data. Fig. 11 compares

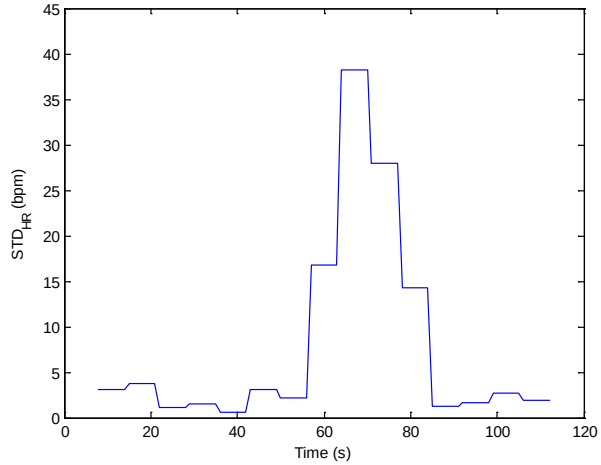
detection methods in terms of HR and SpO_2 estimation and detection accuracy. On average, the SVM algorithm outperforms K, SE, H1 and H2 methods with HR errors of 2.3 bpm, SpO_2 errors of 2.7% and detection errors of 6.1%.

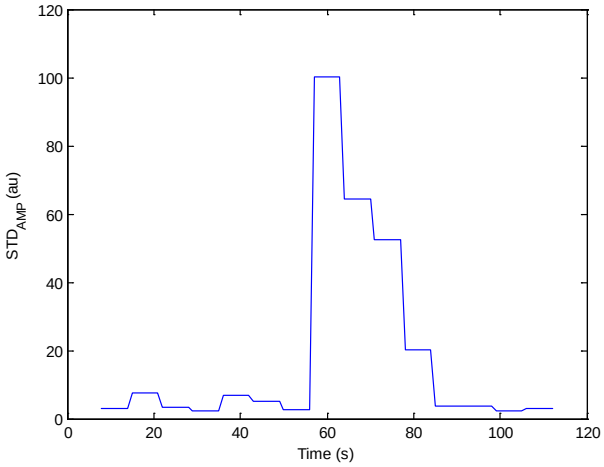
V. Discussion

Robust real-time artifact specification algorithms for raw PPG signals have been elusive to date. Our proposed computational algorithm has been designed based on four parameters: (a) standard deviation of peak-to-peak interval (b) standard deviation of peak-to-peak amplitude (c) standard deviation of systolic and diastolic time ratio, and (d) mean-standard deviation of pulse shape. The proposed MA detection algorithm has been examined for classification of artifacts in PPG data recorded during laboratory controlled experiments and motion artifacts encountered during typical daily activities.. Our results demonstrated that the parameters from neighbor segments as well as target segment have been able to enhance MA detection robustness. The algorithm may also be used to quantify the severity of artifacts based on these parameter values. Our SVM-based motion detection algorithm has offered higher accuracy. This is because the short burst of involuntary artifacts results in different characteristics from clean data and it exhibits larger variability which consequently results in higher time-domain parameters compared to clean data.

The paired-*t* test was performed to determine whether there is significant difference in classification errors obtained from SVM versus other published methods. Results reported in Fig. 9a-c indicate that the mean is significantly different ($p<0.05$ at 95% CI) between SVM and the corresponding methods we compared. For the finger PPG segments, conventional methods, except for H1, show significant difference compared to SVM. On the other hand, all other methods we compared were







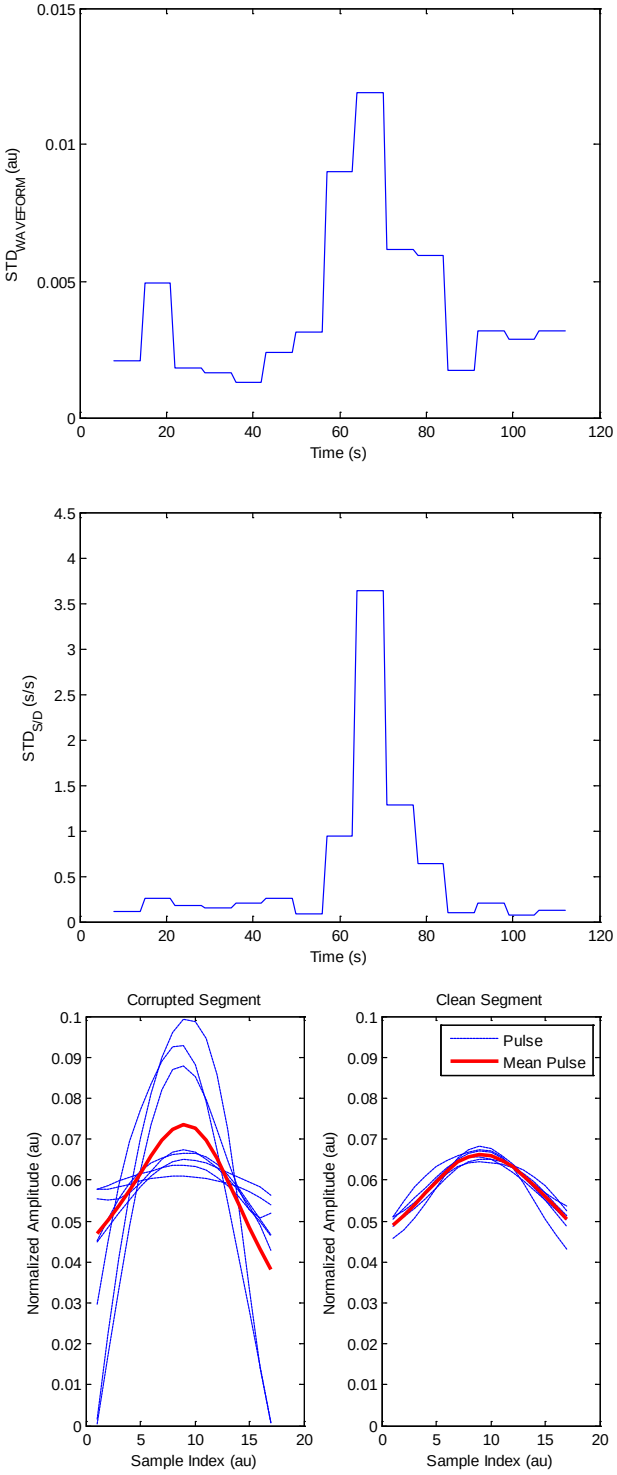


Figure 5. A sample forehead recorded PPG signal (a) along with the (b) standard deviation of P-P intervals (c) standard deviation of P-P amplitudes (d) standard deviation of systolic and diastolic time ratio, and (e) mean standard deviation of pulse shape successive KL-divergence ratio, computed for each segment. The normalized and sampled clean and corrupted PPGs for mean standard deviation of pulse shape is given in (f).

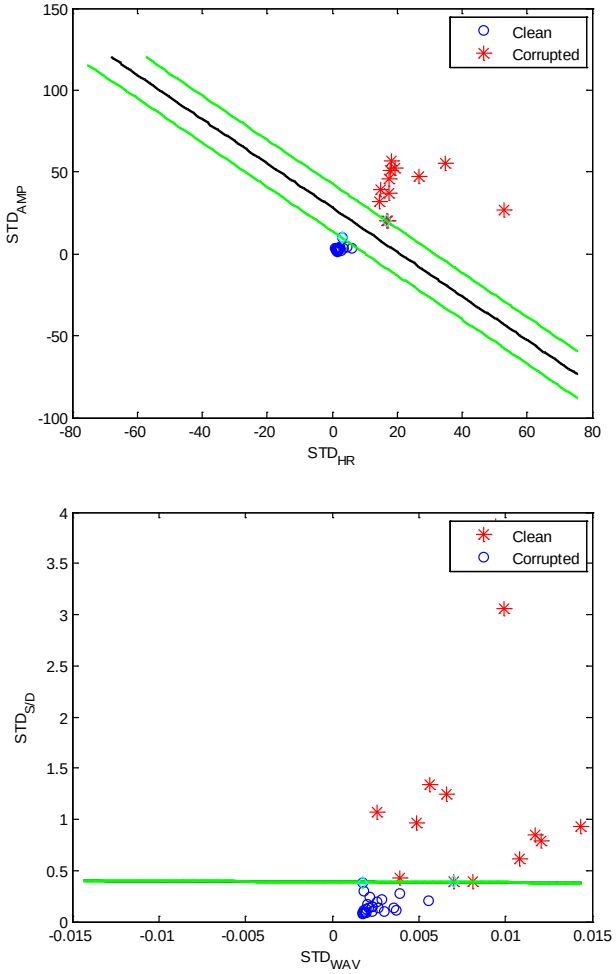


Figure 6. Trained SVM classification with (a) sample training finger recorded PPG signal is given with (b)-(c) pairs of two parameters. The SVM decision and margin boundaries are marked by black and green lines, respectively.

TABLE II
C obtained by 9 fold 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

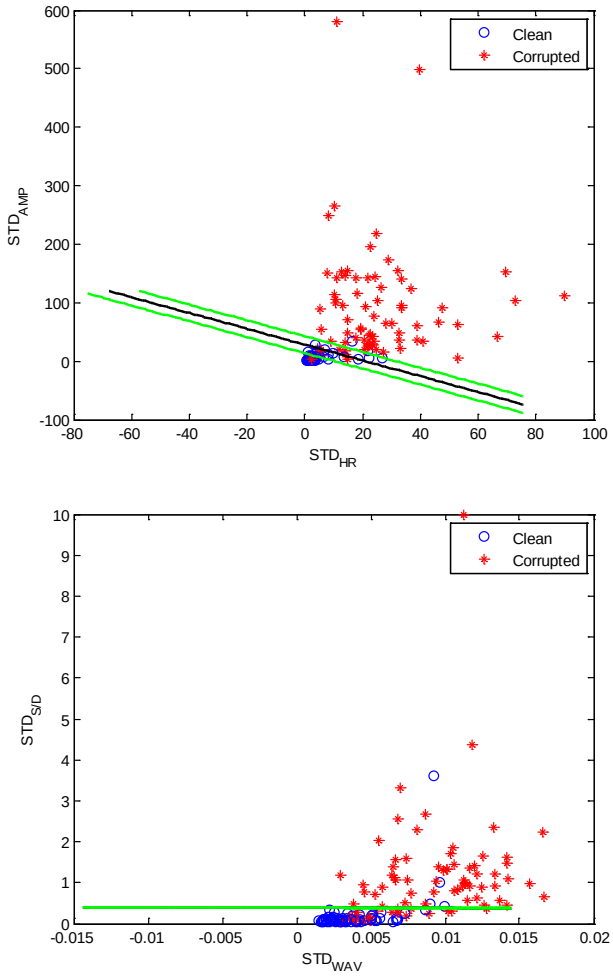


Figure 7. Validation: pairs of parameters for clean and corrupted PPG signals.

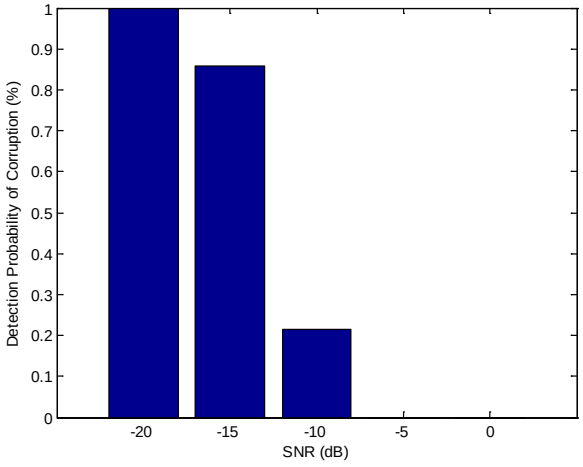


Figure 8. Validation: pairs of parameters of clean and corrupted PPG signals.

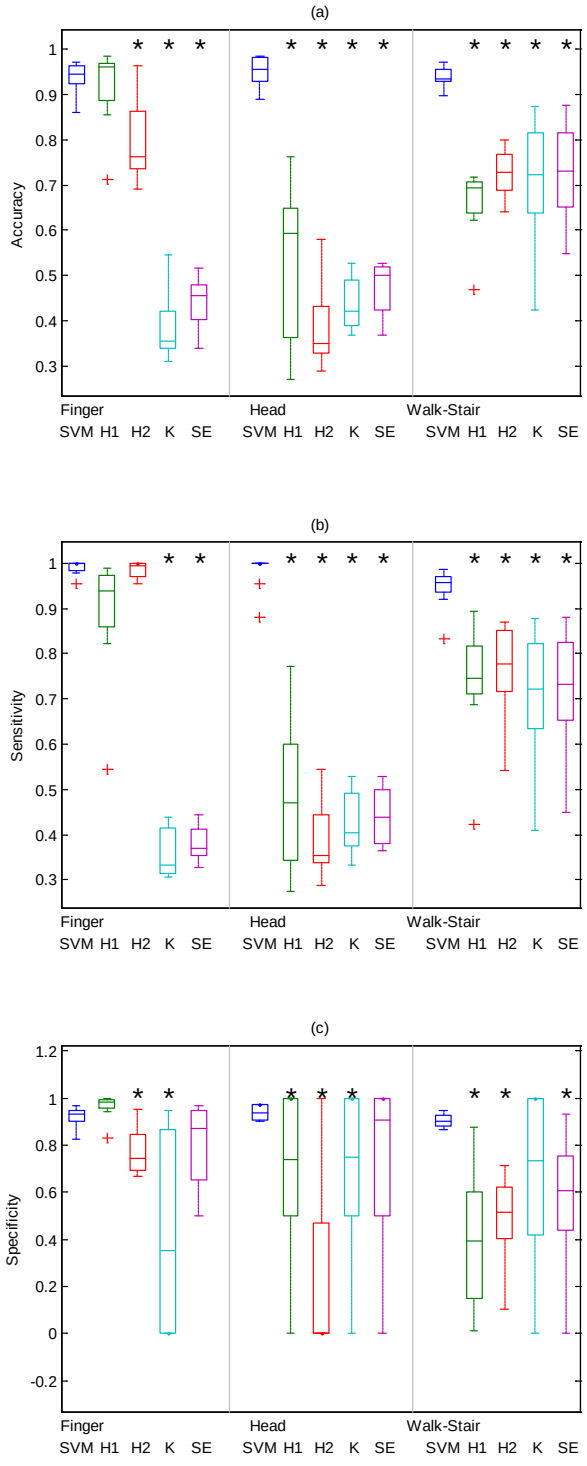


Figure 9. Classification performance comparison between our SVM algorithm, Hjorth (H1, H2), Kurtorsis and Shanon Entropy (K, SE) parameters. (a) Accuracy; (b) Sensitivity; (c) Specificity. The central mark on each box corresponds to the median; the edges of the box correspond to the 25th and 75th percentiles, 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.

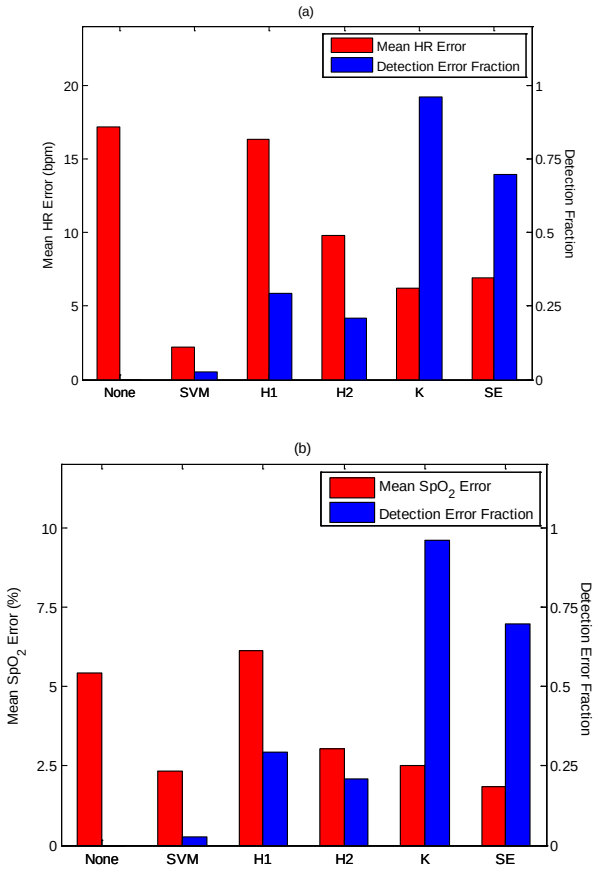


Figure 10. Comparison of Mean errors / detection error fraction between original signal (None) and artifact removed signal from five detection methods (SVM, H1, H2, K, and SE). (a) HR error; (b) SpO_2 error.

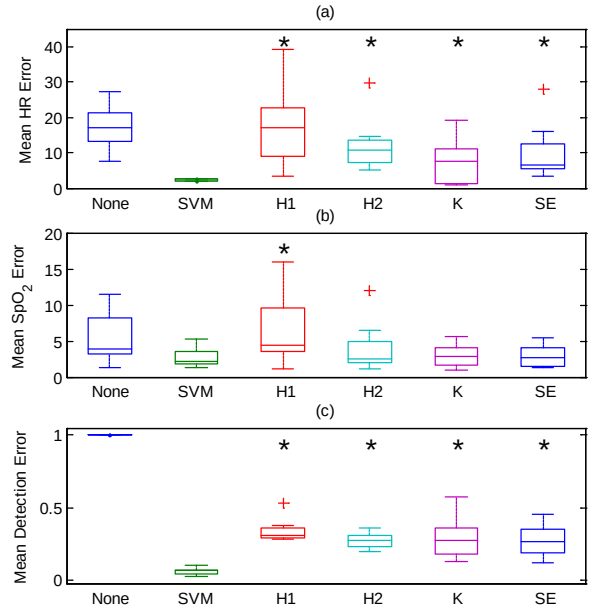


Figure 11. Mean error comparison between our SVM algorithm , Hjorth (H1, H2), Kurtorsis and Shanon Entropy (K, SE) parameters. (a) heart rate; (b) SpO_2; (c) detection error. The central mark on each box corresponds to the median; the edges of the box correspond to the 25th and 75th percentiles, the 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.

significantly different from SVM for forehead recorded PPG segments.

The paired-*t* test results on HR and SpO_2 estimations as well as detection accuracy are summarized in Fig. 11a-c. As shown, SVM is significantly different from H1, H2, K, and SE in terms of HR estimation and detection accuracy (see Figs. 11a and c), while SpO_2 derived from SVM is significantly different from only H1(see Fig. 11b). Hence, our SVM based approach outperforms four conventional methods based on HR and SpO_2 estimations as well as detection accuracy.

In conclusion, our real-time computational algorithm for motion artifact detection has provided high classification and timing estimation accuracy. The potential for the method proposed in this work to have practical applications is high, and integration of the algorithms into a pulse oximeter may have significant implications in real-time clinical monitoring of vital signs.

REFERENCES

[1] B. Widrow, et al., "Adaptive noise cancelling: principles and applications," *Proc. of the IEEE*, vol. 63, no. 12, pp. 1692-1716, 1975.

[2] J. Weng, Z. Ye, and J. Weng. "An improved pre-processing approach for photoplethysmographic signal." Engineering in Medicine and Biology Society, 2005. *IEEE-EMBS 2005. 27th Annual International Conference of the. IEEE*, 2006.

[3] K. T. Sweeney, T. E. Ward, and S. F. McLoone, "Artifact removal in physiological signals—Practices and possibilities." Information Technology in Biomedicine, IEEE Transactions on 16.3 (2012): 488-500.

[4] H. Lee, et al., "The periodic moving average filter for removing motion artifacts from PPG signals." *International Journal of Control Automation and Systems* 5.6 (2007): 701.

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

[6] J. E. Scharf, S. Athan, and D. Cain. "Pulse oximetry through spectral analysis." Biomedical Engineering Conference, 1993., Proceedings of the Twelfth Southern. IEEE, 1993.

[7] J. E. Scharf, S. Athan, and D. Cain, "Reducing motion artifacts in photo plethysmograms by using relative sensor motion: phantom study", *Proceedings of the 12th Southern Biomedical Engineering Conference*, 1993, New Orleans, LA, pp. 227-229, Apr. 1993.

[8] N. S. Trivedi, A. F. Ghouri, N. K. Shah, E. Lai, and S. J. Barker, "Effects of motion, ambient light, and hypoperfusion on pulse oximeter function," J. Clin. Anesth., vol. 9, no. 3, pp. 179–183, 1997.

[9] S. J. Barker and N. K. Shah, "The effects of motion on the performance of pulse oximeters in volunteers," *Anesthesiology*, vol. 86, no. 1, pp. 101–108, 1997.

[10] M. T. Petterson, V. L. Begnoche, and J. M. Graybeal, "The effect of motion on pulse oximetry and its clinical significance," *Anesth Analg*, vol. 105, no. 6 Suppl, pp. S78-84, Dec, 2007.

[11] R. M. Tobin, J. A. Pologe, and P. B. Batchelder, "A characterization of motion affecting pulse oximetry in 350 patients," Anesth Analg, vol. 94, no. 1 Suppl, pp. S54-61, Jan, 2002.

[12] J. A. C. Patterson and G.-Z. Yang. "Ratiometric artifact reduction in low power reflective photoplethysmography." *Biomedical Circuits and Systems, IEEE Transactions on* 5.4 (2011): 330-338.

[13] K. Li, and S. Warren. "A wireless reflectance pulse oximeter with digital baseline control for unfiltered photoplethysmograms." *Biomedical Circuits and Systems, IEEE Transactions on* 6.3 (2012): 269-278.

[14] L. B. Wood and H. H. Asada, "Noise cancellation model validation for reduced motion artifact wearable PPG sensors using MEMS accelerometers," *Conf Proc IEEE Eng Med Biol Soc*, vol. 1, pp. 3525-8, 2006.

[15] N. H. Lovell, S. J. Redmond, J. Basilakis and B. G. Celler, "Biosignal quality detection: an essential feature for unsupervised telehealth

applications," in Proc. 12th *IEEE Int. Conf. on e-Health Networking Applications and Services (Healthcom)*, Jul. 2010, Lyon, France, pp. 81–5

[16] B. Lee, J. Han, H. J. Baek, J. H. Shin, K. S. Park and W. J. Yi, "Improved elimination of motion artifacts from a photoplethysmographic signal using a Kalman smoother with simultaneous accelerometry," *Physiol. Meas*. vol. 31, pp. 1585–603

[17] Y.-D. Lee, S.-J. Jung, and Y.-S. Seo, and W.-Y. Chung, "Measurement of Motion Activity during Ambulatory Using Pulse Oximeter and Triaxial Accelerometer", *Proc. Intl Conf. on Convergence and Hybrid Information Tech.*, 2008.

[18] K. Naraharisetti, V. Prasad, M. Bawa, and M. Tahernezhadi, "Comparison of different signal processing methods for reducing artifacts from photoplethysmograph signal." Electro/Information Technology (EIT), 2011 IEEE International Conference on. IEEE, 2011.

[19] M. R. Ram, K. V. Madhav, 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 Tr. Instru. and Meas.*, vol. 61, no. 5, May 2012

[20] G. Comtois, Y. Mendelson, P. Ramuka, "A Comparative Evaluation of Adaptive Noise Cancellation Algorithms for Minimizing Motion Artifacts in a Forehead-Mounted Wearable Pulse Oximeter," *Proc. Intl. Conf. IEEE EMBS*, 2007.

[21] S. Kunchon, T. Desudchit, and C. Chinrungrueng, "Comparative Evaluation of Adaptive Filters in Motion Artifact Cancellation for Pulse Oximetry,", *Proc. Intl. Colloquium on Singal Proc. & Its Appl.*, 2009.

[22] T. Rusch, R. Sankar, J. E. Scharf, "Signal processing methods for pulse oximetry," Comput. Biol. Med. Ann Biomed Eng, vol. 26, pp. 143-59

[23] M. Raghuram, K. V. Madhav, E. H. Krishna, N. R. Komalla, K Sivani, K. A. Reddy, "Dual-tree complex wavelet transform for motion artifact reduction of PPG signals,", *Proc. IEEE Symposium on Medical Measurement and Applications*, 2012

[24] R. H. Enr´ıquez, M. S. Castellanos, J. F. Rodr´ıguez and J. H. H. C´aceres J L 2002 Analysis of the photoplethysmographic signal by means of the decomposition in principal components," *Physiol. Meas.* 23 N17–29

[25] K. V. Madhav, M. R. Ram, E. H. Krishna, K. N. Reddy and K. A. Reddy, "Extraction of respiratory activity from PPG and BP signals using principal component analysis," *Int. Conf. on Communications and Signal Processing (ICCSP)*, Feb 2011, Kerala, India, pp 452–6

[26] N. Mammone, and F. C. Morabito, "Independent component analysis and high-order statistics for automatic artifact rejection." pp. 2447-2452.

[27] 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. 566-8, Mar, 2006.

[28] Yao, Jianchu, and Steve Warren. "A short study to assess the potential of independent component analysis for motion artifact separation in wearable pulse oximeter signals." Engineering in Medicine and Biology Society, 2005. *IEEE-EMBS 2005. 27th Annual International Conference*, 2005.

[29] R. Krishnan, B. Natarajan, and S. Warren, "Two-Stage Approach for Detection and Reduction of Motion Artifacts in Photoplethysmographic Data," *IEEE Trans Biomed Eng*, Feb 17, 2010.

[30] J. A. Sukor, S. J. Redmond, and N. H. Lovell. "Signal quality measures for pulse oximetry through waveform morphology analysis." *Physiological measurement* 32.3 (2011): 369.

[31] W. Karlen, et al. "Photoplethysmogram signal quality estimation using repeated Gaussian filters and cross-correlation." *Physiological Measurement* 33.10 (2012): 1617.

[32] C. Yu, Z. Liu, T. McKenna, A. T. Reisner, J. Reifman, "A Method for Automatic Identification of Reliable Heart Rates Calculated from ECG and PPG Waveforms," *J Am Med Inform Assoc.* 2006 May-Jun;13(3):309-20

[33] S. Dash, E. Reader, S. Merchant, and K. Chon, "A statistical approach for accurate detection of atrial fibrillation and flutter." *Computers in Cardiology*, pp. 137-140, 2009.

[34] S. Dash, K. H. Chon, S. Lu et al., "Automatic real time detection of atrial fibrillation," Ann Biomed Eng, vol. 37, no. 9, pp. 1701-9, Sep, 2009.

[35] C. Yu, et al. "A method for automatic identification of reliable heart rates calculated from ECG and PPG waveforms." *Journal of the American Medical Informatics Association* 13.3 (2006): 309-320.

[36] S. Lee, B. L. Ibey, W. Xu, et al., "Processing of pulse oximeter data using discrete wavelet analysis," *IEEE Trans Biomed Eng*, vol. 52, no. 7, pp. 1350-2, Jul, 2005.

[37] C.-W. Hsu, C.-C. Chang and C.-J. Lin, "A practical guide to support vector classification." Technical Report Department of Computer Science, National Taiwan University, 2010.

[38] R. Sahni, A. Gupta, K. Ohira-Kist, et al., "Motion resistant pulse oximetry in neonates," *Arch Dis Child Fetal Neonatal Ed*, vol. 88, no. 6, pp. F505-8, Nov, 2003.

[39] Y. S. Yan, C. C. Poon, and Y. T. Zhang, "Reduction of motion artifact in pulse oximetry by smoothed pseudo Wigner-Ville distribution"

[40] J. Y. 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. 1-2, pp. 140-5, Mar, 2006.

[41] J. M. Goldman, M. T. Petterson, R. J. Kopotic, et al., "Masimo signal extraction pulse oximetry," *J Clin. Monit. Comput.*, vol. 16, no. 7, pp. 475-
83, 2000.

[42] C. O. Kamlin, J. A. Dawson, C. P. O'Donnell, et al., "Accuracy of pulsecomparative oximetry measurement of heart rate of newborn infants in the delivery room," *J Pediatr.*, vol. 152, no. 6, pp. 756-60, Jun, 2008.

[43] S. J. Barker, "Motion-resistant" pulse oximetry: a comparison of new and old models," Anesth Analg, vol. 95, no. 4, pp. 967-72, table of contents, Oct, 2002.

[44] S. Tong, Z. Li, Y. Zhu, et al., "Describing the nonstationarity level of neurological signals based on quantifications of time-frequency representation," *IEEE Trans Biomed Eng.*, vol. 54, no. 10, pp. 1780-5, Oct, 2007.

[45] K. H. Chon, S. Dash, and K. Ju, "Estimation of respiratory rate from photoplethysmogram data using time-frequency spectral estimation," *IEEE Trans Biomed Eng.*, vol. 56, no. 8, pp. 2054-63, Aug, 2009.

[46] H. Wang, K. Siu, K. Ju et al., "A high resolution approach to estimating time-frequency spectra and their amplitudes," *Ann Biomed Eng*., vol. 34, no. 2, pp. 326-38, Feb, 2006.

[47] N. Selvaraj, J. Lee, and K. H. Chon, "Time-varying Methods for Characterizing Nonstationary Dynamics of Physiological Systems," *Methods Inf Med*, vol. 49, no. 5, pp. 435-42, Oct 13, 2010.

[48] J. Lee, et al. "Design of filter to reject motion artifact of pulse oximetry." *Computer Standards & Interfaces* 26.3 (2004): 241-249.

[49] C. Yu, et al. "A method for automatic identification of reliable heart rates calculated from ECG and PPG waveforms." *Journal of the American Medical Informatics Association* 13.3 (2006): 309-320.

[50] N. Selvaraj, et al. "Statistical approach for the detection of motion/noise artifacts in Photoplethysmogram." *Engineering in Medicine and Biology Society, EMBC*, 2011 Annual International Conference of the IEEE. IEEE, 2011.

Photoplethysmograph Signal Reconstruction based on a Novel Motion Artifact Detection-Reduction Approach – Part II: Motion Artifact Removal

S. M. A. Salehizadeh[1], Duy K Dao[1], Jo Woon Chong[1], Yitzhak Mendelson[1], David McManus[2], Chad Darling[2] and Ki Chon[1] *Senior Member, IEEE*

[1]Department of Biomedical Engineering, Worcester Polytechnic Institute, Worcester, MA

[2]Department of Medicine, UMASS Medical Center, Worcester, MA

***Abstract*—We introduce a new method to reconstruct motion and noise (MNA) contaminated photoplethysmogram (PPG) data. A method to detect MNA corrupted data is provided in the companion paper. Our reconstruction algorithm is based on the iterative motion artifact removal (IMAR) approach, which utilizes the singular spectral analysis algorithm to remove MNA artifacts so that the most accurate estimates of uncorrupted heart rates (HR) and oxygen saturation values (SpO_2) can be derived. Using both computer simulations and three different experimental data sets, we show that the proposed IMAR approach can reliably reconstruct MNA corrupted data segments, as the estimated HR and SpO_2 values do not significantly deviate from the uncorrupted reference measurements. Comparison of the accuracy of reconstruction of the MNA corrupted data segments between our IMAR approach and the independent component analysis (ICA) is made for all data sets as the latter method has been shown to provide good performance. For simulated data, there were no significant differences in the reconstructed HR and SpO_2 values starting from 10 dB down to -15 dB for both white and colored noise contaminated PPG data using IMAR; for ICA, significant differences were observed starting at 10 dB. Two experimental PPG data sets were created with contrived MNA by having subjects perform random forehead and rapid side-to-side finger movements. The performance of the IMAR approach on these data sets was quite accurate as non-significant differences in the reconstructed HR and SpO_2 were found compared to non-contaminated reference values, in most subjects. However, the accuracy of the ICA was poor as there were significant differences in reconstructed HR and SpO_2 values in most subjects. For non-contrived MNA corrupted PPG data, which were collected with subjects performing walking and stair climbing tasks, the IMAR significantly outperformed ICA as the former method provided HR and SpO_2 values that were non-significantly different than MNA free reference values.**

***Index Terms*—**

I. Introduction

Tissue oxygen saturation reflects the amount of oxyhemoglobin in the blood circulation. The most common method to measure it is based on pulse oximetry, whereby oxidized hemoglobin and reduced hemoglobin have significantly different optical spectra. Specifically, at a wavelength of about 660 nm, and a second wavelength between 805 and 960, there is a large difference in light absorbance between reduced and oxidized hemoglobin. A measurement of the percent oxygen saturation of blood is defined as the ratio of oxyhemoglobin to the total concentration of hemoglobin present in the blood, or simply a comparison of the logarithm of the transmitted light power to emitted light power at the two wavelengths. Pulse oximetry assumes that the attenuation of light is due to both the venous blood and the bloodless tissue. Fluctuations of the pulse oximeter signal are caused by changes in arterial blood volume associated with each heart beat, where the magnitude of the fluctuations depends on the amount of blood rushing into the peripheral vascular bed, the optical absorption of the blood, skin, and tissue, and the wavelength used to illuminate the blood.

The pulse oximeter signal not only contains the blood oxygen saturation and heart rate data, but also other vital information regarding the state of health of a person. The fluctuations of PPG signals contain the influences of arterial, venous, autonomic and respiratory systems on the peripheral circulation. In the current environment where health care costs are ever increasing, a single sensor that has multiple functions is very attractive from a financial perspective. Moreover, wide acceptance of the pulse oximeter as a multi-purpose vital sign monitor is readily expected, since it is easy to use and comfortable for the patient. Knowledge of respiratory rate [1] and heart rate patterns would be clinically useful in many situations in which only the pulse oximeter is routinely monitored. Using the latter to determine the former is not only attractive from an economic perspective, but would eliminate additional sensors, wires and hardware devices a healthcare provider would have to configure and a patient would have to tolerate.

While there are many promising and attractive features of

using pulse oximeters for vital sign monitoring, currently they are mainly used in non-ambulatory settings. This is mainly because motion and noise artifacts (MNA) result in unreliable heart rate and especially the SpO_2 estimation. Certainly, this is why clinicians have cited motion artifacts in pulse oximetry as the most common cause of false alarms, loss of signal, and inaccurate readings [2]

In practice MNA are difficult to remove because they do not have a predefined narrow frequency band and their spectrum often overlaps that of the desired signal [3]. Consequently, development of algorithms capable of reconstructing the corrupted signal and removing corresponding artifacts is a challenging issue.

There are a number of general techniques used for artifact detection and removal. One of the methods used to remove motion artifacts is adaptive filtering [4-8]. The adaptive filter is easy to implement and it also can be used in real-time applications, though the requirement of additional sensors to provide reference inputs is the major drawback of such methods.

There are many MNA reduction techniques based on the concept of the blind source separation (BSS). The BSS is attractive and garnered significant interests since this approach does not require a reference signal. The aim of the BSS is to estimate a set of uncorrupted signals from a set of mixed signals which is assumed to contain both the clean and MNA sources [9]. Some of the popular BSS techniques are independent component analysis (ICA) [10], canonical correlation analysis (CCA) [11], principle component analysis (PCA) [12], and singular spectrum analysis (SSA) [13].

ICA is a technique which the recorded signals are decomposed into their independent components or sources [10]. Canonical correlation analysis (CCA) uses the second order statistics (SOS) to generate components derived from their uncorrelated nature, is another method for separating a number of mixed or contaminated signals [14]. Principle component analysis (PCA) is another noise reduction technique which aims to separate the clean signal dynamics from the MNA data. A multi-scale PCA has been also proposed to account for time-varying dynamics of the signal and motion artifact from PPG recordings [15].

A promising approach which can be applied for signal reconstruction is the singular spectrum analysis (SSA). The SSA is a model-free BSS technique, which decomposes the data into a number of components which may include trends, oscillatory components, and noise (see historical reviews in [16]). The main advantage of SSA over ICA is that SSA does not require a user input to choose the appropriate components for the reconstruction and MA removal. Comparing PCA to SSA, the SSA can be applied in the cases where the number of signal components is more than the rank of the PCA covariance matrix. Applications of the SSA include extraction of the amplitude and low frequency artifacts from single channel EEG recordings [17], and removing heart sound dynamics from respiratory signals [18].

In this Part II of the paper, we introduce a novel approach to reconstruct a PPG signal for those portion of data that have been identified to be corrupted using the algorithm detailed in Part I of the companion paper. The fidelity of the reconstructed signal was determined by comparing the estimated SpO_2 and HR to the reference values. In addition, we compare the reconstructed SpO_2 and HR values obtained via the ICA to our method. We have chosen to compare our method to the ICA since the latter has recently been shown to provide accurate reconstruction of the corrupted PPG signals [19].

Materials and Method

A. Experimental Protocol and Preprocessing

Three sets of data were collected from healthy subjects recruited from the student community of Worcester Polytechnic Institute (WPI). This study was approved by WPI's human ethics committee and all the subjects were given informed consent before data recording.

In the first experiment, 11 healthy volunteers were asked to wear a forehead reflectance pulse oximeter developed in our lab along with a reference Masimo Radical (Masimo SET®) finger transmittance pulse oximeter. The PPG signals from the forehead sensor and the reference heart rate (HR) were acquired simultaneously. The heart rates and arterial oxygen saturation (SpO_2) signals were acquired at 80Hz and 1Hz, respectively. After baseline recording for 5 minutes without any movement (i.e. clean data), motion artifacts were induced in the PPG data by the spontaneous movements in both horizontal and vertical directions of the subject's head while the right middle finger was kept stationary and attached to the Masimo pulse oximeter. The subjects were directed to introduce the motions for specific time intervals that determined the percentage of noise within each 1 minute segment, varying from 10 to 50%. For example, if a subject was instructed to make left-right movements for 6 seconds, 1 min segment of data would contain 10% noise.

The second dataset consisted of finger-PPG signals from the same 9 healthy volunteers in an upright sitting posture using an infrared reflection type PPG transducer (TSD200) and a biopotential amplifier (PPG100) with a gain of 100 and cut-off frequencies of 0.05-10Hz. The MP1000 (BIOPAC Systems Inc., CA, USA) was used to acquire finger PPG signals at 100Hz. Two pulse oximeters were placed on the same index and middle finger simultaneously. After baseline recording for 5 minutes without any movement (i.e. clean data), motion artifacts were induced in the PPG data by the left-right movements of the index finger while the middle finger was kept stationary and this was used as a reference. Similar to the first dataset, motion was induced at specific time intervals corresponding to 10 to 50% corruption duration in 1 minute segment. Such controlled movement was carried out for five times per subject.

The third dataset consisted of data measurements from 9 subjects with the PPG signal recorded from the subjects' forehead using our custom sensor simultaneously with the reference ECG, HR and SpO_2 from Holter Monitor at 180Hz and Masimo (Rad-57) pulse oximeter at 0.5Hz respectively. The reference pulse oximeter provided HR and SpO_2 measured

from the subject right index finger, which were held steadily around his chest. The signals were recorded while the subjects were going through sets of walking and climbing up-down flights of stairs for around 45 min.

Once data were acquired, PPG signals from all three experiments outlined above were preprocessed offline using Matlab (MathWorks,R2012a). The PPG signals were filtered using a zeros-phase forward-reverse, 4th order IIR band-pass filter with cutoff frequency 0.5-12Hz.

II. Motion Artifact Removal

To reconstruct the corrupted portion (i.e. with motion artifacts) of the PPG signal which has been detected using the support vector machine approach as provided in the accompanying paper, we propose a novel hybrid procedure using Iterative Singular Spectrum Analysis (ISSA) and a frequency matching algorithm. Henceforth, we will call these combined procedures as the iterative motion artifact removal (IMAR) algorithm.

A. Singular Spectrum Analysis

The SSA is composed of two stages: singular decomposition and spectral reconstruction. The singular is the spectral decomposition or eigen-decomposition of the data matrix whereas the term spectrum is for the reconstruction of the signal based on using only the significant eigenvectors and associated eigenvalues. The assumption is that given a relatively high signal-to-noise ratio of the data, the significant eigenvector and associated eigenvalues represent the signal dynamics and the less significant, the MNA components.

The calculation of the singular stage of the SSA consists of two steps: embedding followed by the singular value decomposition (SVD). In essence, these procedures decompose the data into signal dynamics consisting of trends, oscillatory components, and MNA. The spectrum stage of the SSA algorithm consists of two stages: grouping, and diagonal averaging. These two procedures are used to reconstruct the signal dynamics but without the MNA components. In the proceeding section, we detail all four stages of the SSA algorithm.

1) Signal Decomposition:

a) Embedding

Assume we have a nonzero real-value time series of length N samples, i.e., $X = \{x_1, x_2, ..., x_N\}$. In the embedding step, window length $f_s / f_l < L < N/2$ is chosen to embed the initial time series, where f_s is the sampling frequency and f_l is the lowest frequency in the signal. We map the time series X into the L lagged vectors, $X = \{x_i, x_{i+1}, ..., x_{i+L-1}\}$ for $i = 1, ..., K$, where $K = N - L + 1$ [16]. The resultant is the trajectory data matrix T_x with X_i which is each row of T_x for $i = 1, ..., K$.

$$T_x = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_K \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_L \\ x_2 & x_3 & \cdots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_K & x_{K+1} & \cdots & x_N \end{bmatrix} \quad (1)$$

From Eq.1, it is evident that the trajectory matrix, T_x, is a Hankel matrix.

b) Singular Value Decomposition

The next step is to apply the SVD to the trajectory matrix T_x which results in eigenvalues and eigenvectors of the matrix $T_x T_x^T$ so to obtain the decomposed eigentriple product trajectory matrices T_i for $i = 1, ..., L$ as $T = USV^T$ [16]. U_i for $1 < i < L$ is a $K \times L$ orthonormal matrix. S_i for $1 < i < L$ is a diagonal matrix and V_i for $1 < i < L$ is a $L \times L$ square orthonormal matrix, which is considered as the principle component [37]. In this step, T_x has L many singular values which are $\sqrt{\lambda_1} > \sqrt{\lambda_2} >, ..., \sqrt{\lambda_L}$. Thus the i^{th} eigentriple of T_i can be written as $U_i \times \sqrt{\lambda_i} \times V_i^T$ for $i = 1, 2, ..., d$, in which $d = \max(i : \sqrt{\lambda_i} > 0)$ is the number of nonzero singular values of T_x. Normally every harmonic component with a different frequency produces two eigentriples with similar singular values. So the trajectory matrix T_x can be denoted as [16]

$$\begin{aligned} T_x &= T_1 + T_2 + ... + T_d \\ &= U_1 \sqrt{\lambda_1} V_1^T + ... + U_d \sqrt{\lambda_d} V_d^T \\ &= \sum_{i=1}^{d} U_i \sqrt{\lambda_i} V_i^T \end{aligned} \quad (2)$$

Projecting the time series onto the direction of each eigenvector yields the corresponding temporal principal component (PC) [20].

2) Reconstruction:

The reconstruction stage has two steps: grouping and diagonal averaging. The reconstruction stage involves the grouping of the subgroups of the decomposed trajectory matrices and a diagonal averaging step is needed so that a new time series can be formed [13].

a) Grouping

The grouping step of the reconstruction stage is to decompose the $L \times K$ matrix T_i into subgroups according to the trend, oscillatory components, and MNA dynamics. The grouping step divides the set of indices $\{1, 2, ..., d\}$ into a collection of m disjoint subsets of $I = \{I_1, ..., I_m\}$ [21]. Thus, T_I corresponds to the group $I = \{I_1, ..., I_m\}$. T_{I_i} is a sum of T_j, where

$j \in I_i$. So T_x can be expanded as

$$T_x = \overbrace{T_1 + \ldots + T_L}^{SVD} = \overbrace{T_{I_1} + \ldots + T_{I_m}}^{Grouping} \quad (3)$$

b) Diagonal averaging

In the final stage of analysis, each resultant matrix , T_{Ii}, in Eq. (3) is transformed into a time series of length N . We obtain the time series $\tilde{T}^{(i)}$ by averaging the corresponding diagonals of the matrix T_{I_i} [20]. Let the Hankelization operator H be averaging of the corresponding diagonals of the matrix $\tilde{T}^{(i)}$ that is H transforms $\tilde{T}^{(i)}$ into $\tilde{X}^{(i)} = HT_{I_i}$ for $i = 1,\ldots,m$ [21], then under the assumption of weak separability the initial signal X can be reconstructed by

$$X = \tilde{X}^{(1)} + \tilde{X}^{(2)} + \ldots + \tilde{X}^{(M)} \quad (4)$$

We can assert $\tilde{X}^{(1)}$ can be related to the trend of the signal; however, harmonic and noisy components do not necessarily follow the order of $\sqrt{\lambda_1} > \sqrt{\lambda_2} > \ldots > \sqrt{\lambda_M}$.

B. Iterative Motion Artifact Removal based on SSA

In order to reconstruct the MNA corrupted segment of the signal, an iterative motion artifact removal approach based on SSA is introduced. The ultimate goodness of the reconstructed signal is determined by the accuracy of the estimated SpO_2 and HR values. The top and bottom panels of Fig.1 show clean and MNA corrupted signals, respectively.

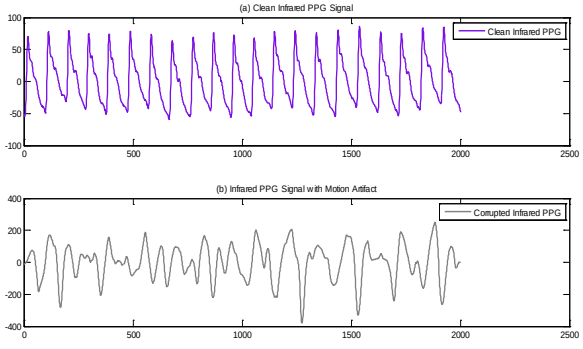
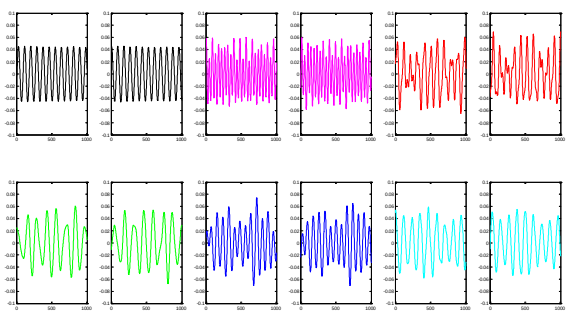
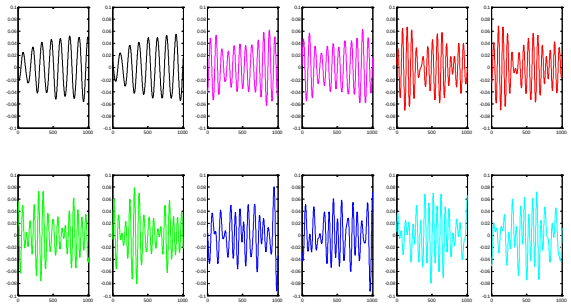


Figure 1. Typical infrared PPG signal; (a) clean, (b) corrupted with motion artifacts.



(a)



(b)

Figure 2. The first 12 eigenvector components of the PPG signal for: (a) Clean Infrared PPG, (b) Corrupted Infrared PPG.

Fig. 2(a) and (b) shows the first 12 eigenvectors of the clean and MNA corrupted data as shown in Fig. 1, respectively. The most important part of the SSA is to choose the proper eigenvector components for reconstruction of the signal. Under the assumption of high SNR, the normal practice is to select only the largest eigenvalues and associated eigenvectors for signal reconstruction. However, most often it is difficult to determine the demarcation of the significant from non-significant eigenvalues. Further, the MNA dynamics can overlap with the signal dynamics, hence, choosing the largest eigenvalues do not necessarily result in MNA free signal.

To overcome the above limitations, we have modified the SSA approach. The first step of our modified SSA involves computing singular value decomposition on both corrupted data segment and its most prior adjacent clean data segment. Under the assumption of a high SNR of the data, the second step is to retain only the top 5% of the eigenvalues and the associated eigenvectors. The third step is to replace the corrupted segment's top 5% eigenvalues with the clean segment's eigenvalues. The fourth step is to further limit the number of eigenvectors by choosing only those eigenvectors that have heart rates between $0.66^{Hz} < f_s < 3^{Hz}$ for both the clean and noise corrupted data segments. The two extrema heart rate ranges are chosen so that they account for possible scenarios that one may encounter with low and high heart rates. With the remaining candidate eigenvectors as a result of the step four, we further prune non-significant eigenvectors by performing frequency matching of the noise corrupted eigenvectors to those of the clean data segment's eigenvectors, in the fifth step. Only those eigenvectors' frequencies that match to those of the clean eigenvectors are retained from the pool of eigenvectors remaining from the step four. For the remaining eigenvector candidates, we perform iterative SSA to further reduce MNA and match the dynamics of the clean data segments eigenvectors for the final step. For each iteration we perform the standard SSA algorithm. It is our experience that this convergence is achieved within 4 iterations.

Fig. 3 shows an example of the iterative SSA procedure applied to candidate eigenvectors that have resulted from the procedure step four of the modified SSA algorithm. Note that

there may be several eigenvectors remaining after the fifth step, hence, this example shows an iterative SSA procedure performed on a particular set of candidate eigenvectors that may match most closely to an eigenvector of a clean data segment. The top panels of Fig. 3 represent one of the eigenvectors of the clean signal and the second panels represents MNA corrupted signal's candidate eigenvectors which have the same frequency as that of the clean signal's eigenvector. The remaining lower panels represent the candidate eigenvectors after they have gone through successive four iterations of the SSA algorithm. For this portion of the SSA algorithm, we perform SVD on the trajectory matrix of Eq. (1) created from the candidate eigenvector and then rreconstruct the eigenvectors based on SSA using only the first 3 largest eigenvalues obtained from the SVD. This process repeats iteratively until the shape of the reconstructed eigenvector closely resembles one of the clean eigenvectors with the same frequency. It can be seen from Fig. 3 that after 4 iterations the result shown in the panel (a) correspond most closely to the clean signal's eigenvector, hence, this eigenvector is selected rather than the eigenvectors shown in the panels b-c. We calculate the discarding metric (DM) at each iteration and compare this value to the DM value of the corresponding clean component. The DM is calculated according to:

$$DM = \sum |u| \Big/ L(u) \quad (5)$$

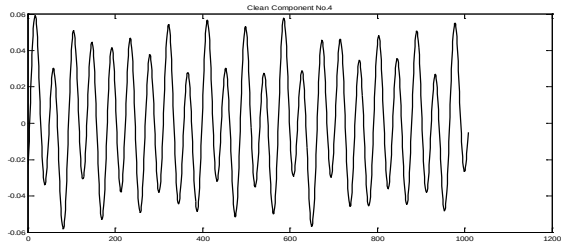
where u is the signal component, and $|.|$, $L(.)$ are absolute operator and component length, respectively. The entire procedure for the modified SSA algorithm is summarized in Table I.

TABLE I
Iterative Motion Artifact Removal (IMAR) Procedure

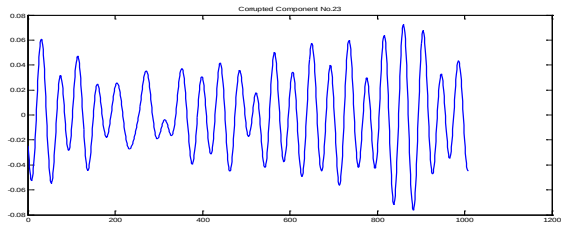
Assumption –Heart rate and SpO_2 do not change abruptly and it is stationary within the short data segment.
Application – Offline Motion Artifact Removal
Objective – Reconstruction of corrupted PPG segment for the purpose of estimating Heart Rates and SpO_2.

Routine

Step 1. First, compute SVD on both corrupted data segment and its most prior adjacent clean data segment
Step 2. Next, keep the top 5% of the clean and corrupted components, of which the eigenvalues are sorted from the largest to smallest.
Step 3. In this step replace the corrupted eigenvalues with corresponding clean eigenvalues.
Step 4. Among the clean and corrupted components, only choose those with frequency within the heart rate frequency range of $0.66<Fs<3^{Hz}$.
Step 5. Apply frequency matching to discard those corrupted components (from Step4) with different frequency comparing to clean components frequencies.
Step 6. Removing corruption from each component obtained from Step5 applying the basic SSA algorithm iteratively.
6.a. Calculate discarding metric for components achieved from SSA-iterative and counterpart clean components from the Eq. 5)
6.b. Select those processed components with the closest DM and frequency value to the corresponding clean component's DM and frequency value.
Step 7. Finally, reconstruct the corrupted PPG segment based on the components achieved from Step 6.



(a)



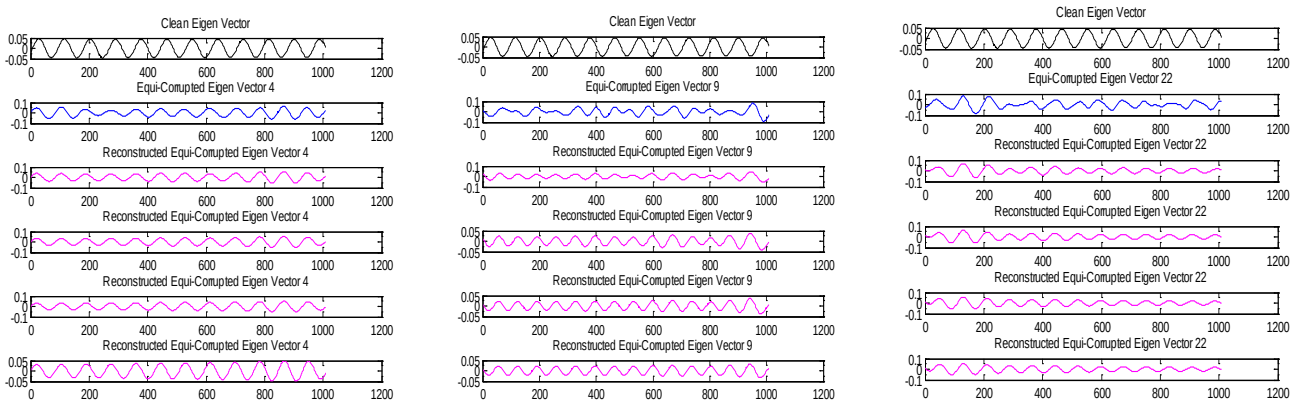
(b)

Figure 3. Sample optimal component with frequency of 1.875 Hz selected based on frequency matching algorithm; (a) 4th clean component, (b) 23rd corrupted component.

III. RESULTS

A. *Noise Sensitivity Analysis*

As a part of the validation of the proposed IMAR procedure, different SNR levels of Gaussian white noise (GWN) and colored noise were added to an experimentally collected clean segment of PPG signal. The purpose of the simulations was to quantitatively determine the level of noise that can be tolerated by the algorithm. Seven different SNR levels ranging from 10 dB to -25 dB was considered. For each SNR level, 50 independent realizations of GWN and colored noise were added separately to a clean PPG signal. Euler-Maruyama method was used to generate colored noise [22].



(a) (b) (c)

Figure 4 Iterative reconstruction of a corrupted eigenvector with frequency of 0.967 Hz. Black font signals (top panels) represent the clean component with frequency of 0.976 Hz; Blue font signals (2nd rows) indicate the corrupted component with the same frequency; Pink font signals are related to iterative evolution of corrupted component to a clean oscillatory signal. (a) Reconstruction of 4th corrupted eigenvector comparing to the counterpart clean component. The final pattern after 4 iterations resembles the black font clean component in the top panel. This component is chosen among the components with the same frequency, since it shows the most similarity to the black font clean component. (b) Reconstruction of 9th corrupted eigenvector comparing to counterpart clean component. (c) Reconstruction of 22nd corrupted eigenvector comparing to counterpart clean component.

Fig. 5 shows the results of these simulations with additive GWN. The left panels show pre- and post-reconstructed HR comparison to the reference HR; the right panels show the same comparison for SpO_2. Tables II and III show the mean and the standard deviation values of the pre- (2nd column) and post-reconstructed (4th column), and the reference (3rd column) HR and SpO_2 values, respectively for all SNR. The last columns of Table II and III also shows comparison of the estimated HR and SpO_2 values obtained by the ICA method [19]. As shown in Fig. 4 and Tables II and III, the reconstructed HR and SpO2 values using our IMAR approach was found to be not statistically different when compared to the reference values for all SNR except for -25 dB. However, the ICA method fails and we obtain significantly different values to those of the reference HR and SpO_2 values when the SNR is lower than -10 dB.

Fig. 5 and Tables IV and V show corresponding results to that of Fig. 4 and Tables II and III but with additive colored noise. Similar to the GWN case, the reconstructed HR and SpO2 values using the proposed IMAR approach is found to be not significantly different than the reference values for all SNR except for -25 dB. Moreover, the ICA compares poorly compared to our IMAR as the HR and SpO_2 values from the former method are found to be significantly different to the reference values for all SNR.

TABLE II
Comparison Statistical Analysis of HR Estimations from IMAR reconstructed PPG for Different Levels of Additive White Noise

SNR (dB)	Head HR (mean ± std)	Finger HR (Reference) (mean ± std)	IMAR Reconstructed HR (mean ± std)	ICA Reconstructed HR (mean ± std)
10	54.80 ± 2.08	54.81 ± 1.81	55.05 ± 0.15	52.86 ± 0.44*
0	54.80 ± 2.72	54.81 ± 1.81	55.05 ± 0.14	50.58 ± 0.62*
-5	56.37 ± 8.18	54.81 ± 1.81	55.05 ± 0.15	48.64 ± 0.51*
-10	46.02 ± 22.93	54.81 ± 1.81	55.09 ± 0.15	46.85 ± 0.45*
-15	121.62 ± 69.33	54.81 ± 1.81	54.73 ± 0.62	45.17 ± 0.28*
-20	80.08 ± 37.69	54.81 ± 1.81	56.49 ± 2.69	43.08 ± 0.32*
-25	103.62 ± 52.49	54.81 ± 1.81	76.45 ± 7.52*	41.11 ± 0.30*

TABLE III
Comparison Statistical Analysis of Estimations from IMAR reconstructed PPG for Different Levels of Additive White Noise

SNR (dB)	Head SpO2 (mean ± std)	Finger SpO2 (Reference) (mean ± std)	IMAR Reconstructed SpO2 (mean ± std)	ICA Reconstructed SpO2 (mean ± std)
10	106.88 ± 0.51	94.23 ± 0.80	94.83 ± 0.38	90.92 ± 0.38*
0	108.98 ± 0.14	94.23 ± 0.80	94.81 ± 0.42	86.88 ± 0.16*
-5	109.42 ± 0.06	94.23 ± 0.80	94.77 ± 0.26	82.86 ± 0.27*
-10	109.69 ± 0.04	94.23 ± 0.80	94.68 ± 0.30	78.81 ± 0.29*
-15	109.82 ± 0.02	94.23 ± 0.80	94.90 ± 0.41	74.88 ± 0.23*
-20	109.89 ± 0.01	94.23 ± 0.80	107.38 ± 1.06*	70.87 ± 0.22*
-25	109.94 ± 0.00	94.23 ± 0.80	97.38 ± 7.39*	66.91 ± 0.26*

B. *Heart Rate and SpO_2 Estimation from Forehead sensor*

As described in section II, we collected PPG data under three different experimental settings so that our proposed approach can be more thoroughly tested and validated. For all three experimental settings, the efficacy of our IMAR approach for the reconstruction of the MNA affected portion of the signal will be compared with the reference HR and SpO_2 values for all experimental datasets.

TABLE IV
Comparison Statistical Analysis of HR Estimations from IMAR reconstructed PPG for Different Levels of Additive Colored Noise

SNR (dB)	Head HR (mean ± std)	Finger HR (Reference) (mean ± std)	IMAR Reconstructed HR (mean ± std)	ICA Reconstructed HR (mean ± std)
10	54.75 ± 1.73	54.81 ± 1.81	55.05 ± 0.26	53.36 ± 0.79

0	55.64 ± 2.72	54.81 ± 1.81	55.06 ± 0.27	$50.83 \pm 0.54^{*}$
-5	55.67 ± 2.88	54.81 ± 1.81	55.06 ± 0.15	$48.90 \pm 0.32^{*}$
-10	51.05 ± 8.24	54.81 ± 1.81	55.07 ± 0.13	$46.79 \pm 0.30^{*}$
-15	61.65 ± 32.08	54.81 ± 1.81	55.17 ± 0.08	$45.15 \pm 0.30^{*}$
-20	73.41 ± 47.73	54.81 ± 1.81	$45.96 \pm 5.59^{*}$	$42.96 \pm 0.41^{*}$
-25	66.37 ± 40.80	54.81 ± 1.81	$61.86 \pm 2.12^{*}$	$41.04 \pm 0.37^{*}$

TABLE V
Comparison Statistical Analysis of SpO2 Estimations from IMAR reconstructed PPG for Different Levels of Additive Colored Noise

SNR (dB)	Head SpO2 (mean ± std)	Finger SpO2 (Reference) (mean ± std)	IMAR Reconstructed SpO2 (mean ± std)	ICA Reconstructed SpO2 (mean ± std)
10	94.14 ± 0.99	94.23 ± 0.80	94.85 ± 0.41	$90.95 \pm 0.18^{*}$
0	94.71 ± 1.20	94.23 ± 0.80	94.85 ± 0.53	$86.84 \pm 0.24^{*}$
-5	96.19 ± 1.41	94.23 ± 0.80	93.92 ± 0.83	$82.86 \pm 0.34^{*}$
-10	99.27 ± 1.46	94.23 ± 0.80	94.88 ± 0.96	$78.89 \pm 0.18^{*}$
-15	103.00 ± 0.88	94.23 ± 0.80	94.42 ± 1.71	$74.87 \pm 0.25^{*}$
-20	107.63 ± 0.26	94.23 ± 0.80	$74.74 \pm 7.92^{*}$	$70.89 \pm 0.17^{*}$
-25	105.91 ± 0.49	94.23 ± 0.80	$70.75 \pm 15.08^{*}$	$66.89 \pm 0.26^{*}$

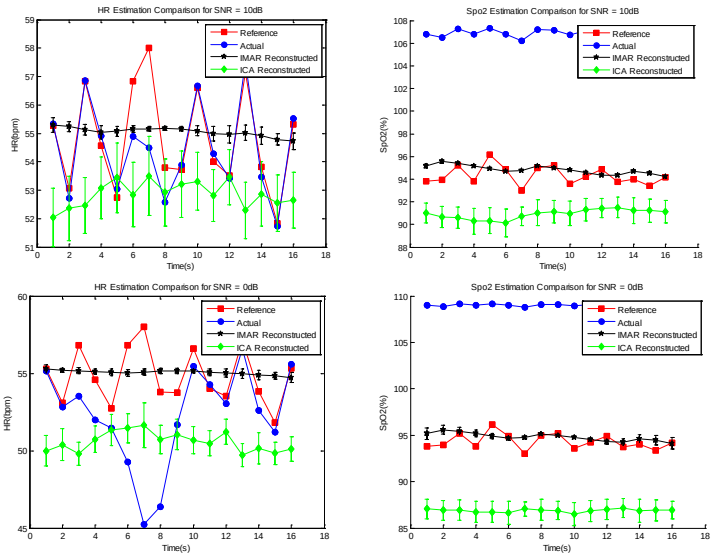
In order to estimate the HR from the PPG signal, a custom designed peak detection algorithm was applied in this study. For the error-free SpO_2 estimation, Red and IR PPG signals with clearly separable DC and AC parts are required. The pulsatile components of the Red and IR PPG signals are denoted as AC_{Red} and DC_{Red}, respectively, the "ratio-of-ratio" R is estimated [23, 24] as

$$R = \frac{AC_{\mathrm{Red}}/DC_{\mathrm{Red}}}{AC_{IR}/DC_{IR}} \tag{6}$$

Accordingly, SpO_2 is computed by substituting the R value in an empirical linear approximate relation given by

$$SpO_2(\%) = (110 - 25R)(\%) \tag{7}$$

After applying the proposed IMAR procedure on the identified MNA segment of the PPG signal, we estimate the SpO_2 (using Eqs. 6-7) and HR, and compare it to the corresponding reference and MNA contaminated segment values. As was the case with the noise sensitivity analysis section, we compare the performance of the IMAR algorithm to the ICA method. The top and bottom panels of Fig. 6 represent a representative HR and SpO_2 comparison results, respectively. We can see from these figures that the estimated values for both HR (left panels) and SpO_2 (right panels) from the IMAR (black font) track closely to the reference values recorded by the Masimo transmittance type finger pulse oximeter (red square-line), while the estimated HR and SpO_2 obtained from the ICA method (green font) deviate significantly to the reference signal. Tables VI and VII show comparison of the IMAR and the ICA reconstructed HR and SpO_2 values, respectively, for all 10 subjects. As shown in Table VI, there was no significant difference between the finger reference HR and the IMAR reconstructed HR in 6 out of 10 subjects. However, there was significant difference between the finger reference HR and the ICA reconstructed HR in all 10 subjects. Similarly, the reconstructed SpO_2 values from the IMAR was found to be not significantly different than the finger reference values in 6 out of 10 subjects, but ICA method was found to be significantly different for all 10 subjects.



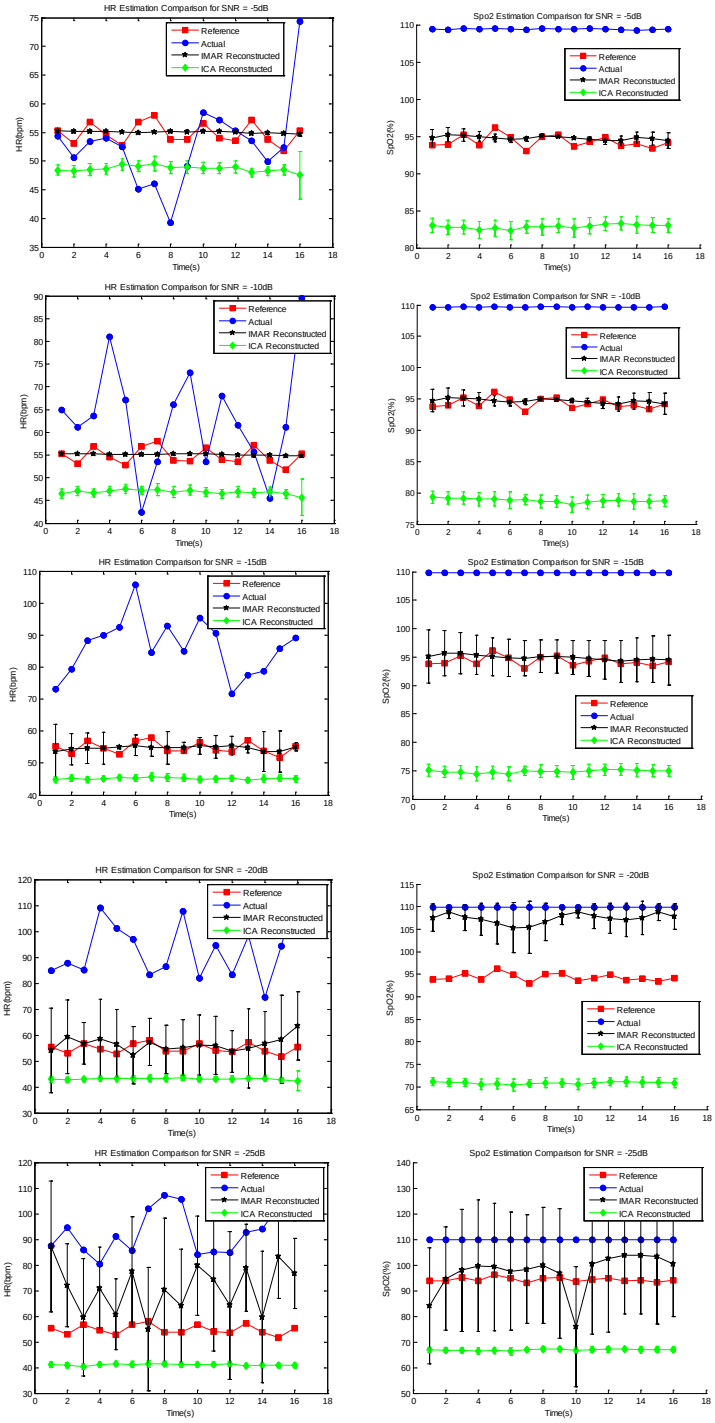
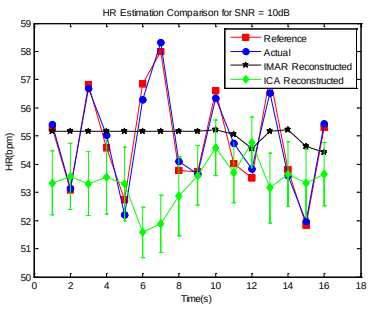
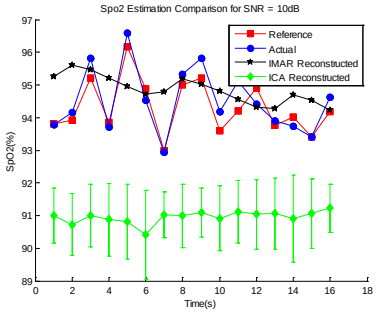
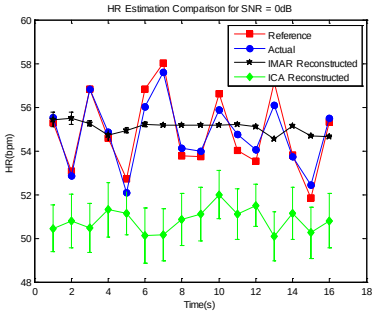
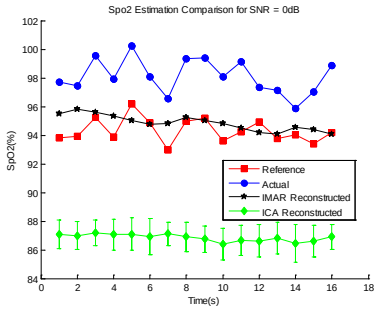


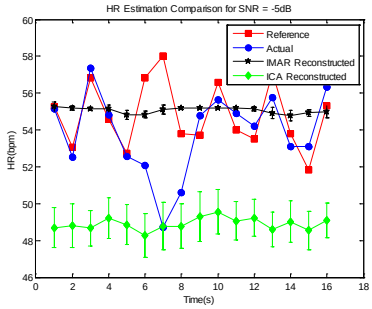
Figure 5 (Left) Heart Rate estimated from reconstructed PPG for different additive white noise level; (Right) SpO2 estimated from reconstructed PPG for different levels of additive white noise.

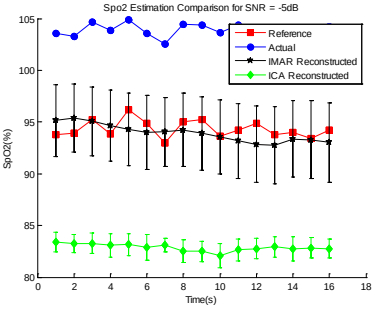

HR Estimation Comparison for SNR = 10dB
Reference
Actual
IMAR Reconstructed
ICA Reconstructed
HR(bpm)
Time(s)

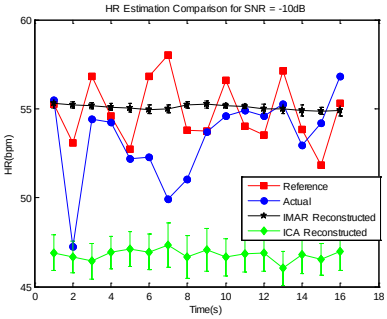

Spo2 Estimation Comparison for SNR = 10dB
Reference
Actual
IMAR Reconstructed
ICA Reconstructed
SpO2(%)
Time(s)

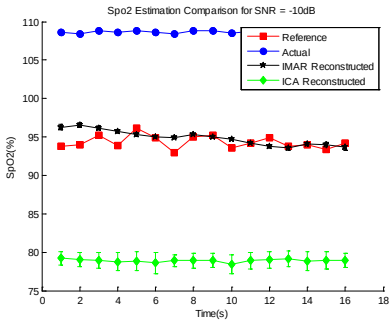

HR Estimation Comparison for SNR = 0dB
Reference
Actual
IMAR Reconstructed
ICA Reconstructed
HR(bpm)
Time(s)

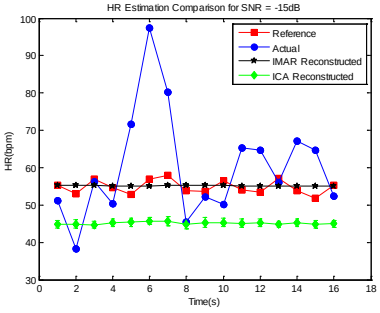

Spo2 Estimation Comparison for SNR = 0dB
Reference
Actual
IMAR Reconstructed
ICA Reconstructed
SpO2(%)
Time(s)

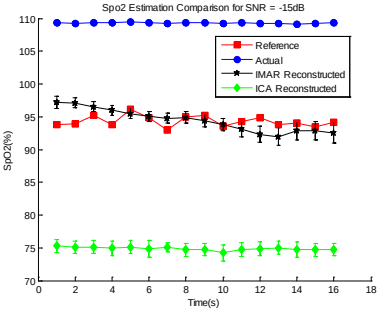

HR Estimation Comparison for SNR = -5dB
Reference
Actual
IMAR Reconstructed
ICA Reconstructed
HR(bpm)
Time(s)


Spo2 Estimation Comparison for SNR = -5dB
Reference
Actual
IMAR Reconstructed
ICA Reconstructed
SpO2(%)
Time(s)


HR Estimation Comparison for SNR = -10dB
Reference
Actual
IMAR Reconstructed
ICA Reconstructed
HR(bpm)
Time(s)


Spo2 Estimation Comparison for SNR = -10dB
Reference
Actual
IMAR Reconstructed
ICA Reconstructed
SpO2(%)
Time(s)


HR Estimation Comparison for SNR = -15dB
Reference
Actual
IMAR Reconstructed
ICA Reconstructed
HR(bpm)
Time(s)


Spo2 Estimation Comparison for SNR = -15dB
Reference
Actual
IMAR Reconstructed
ICA Reconstructed
SpO2(%)
Time(s)

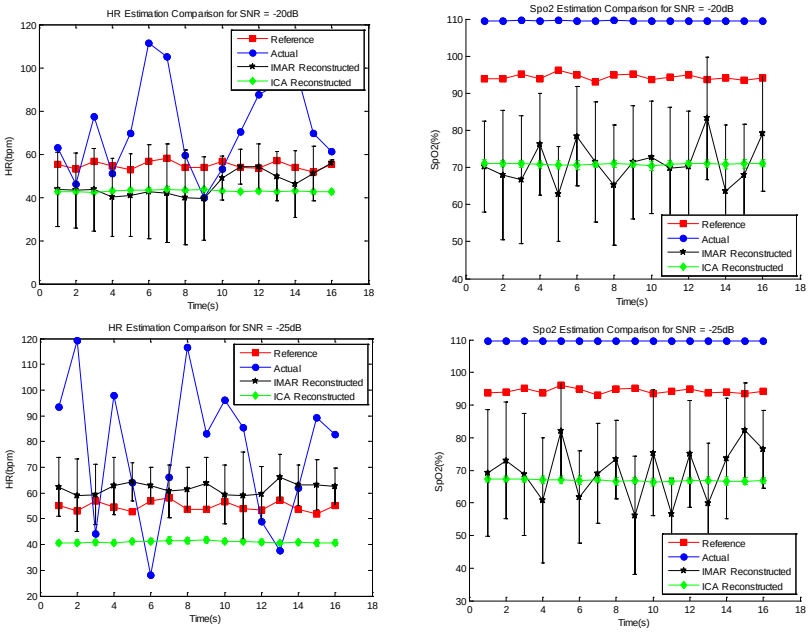


Figure 6. (Left) Heart Rate estimated from reconstructed PPG for different additive colored noise level; (Right) SpO2 estimated from reconstructed PPG for different levels of additive colored noise.

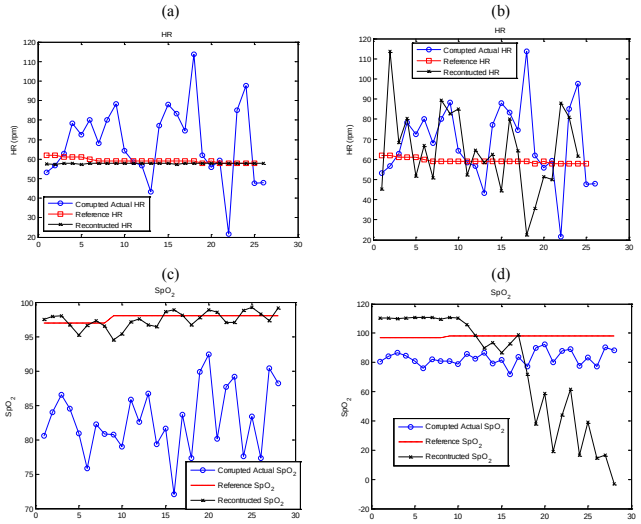


Figure 7. (a) HR estimated from IMAR reconstructed PPG comparing to reference and corrupted PPG; (b) HR estimated from ICA reconstructed PPG comparing to reference and corrupted PPG; (c) SpO2 estimated from IMAR reconstructed PPG comparing to reference and corrupted PPG; (d) SpO2 estimated from ICA reconstructed PPG comparing to reference and corrupted PPG;

TABLE VI
Comparison Statistical Analysis of HR Estimations from IMAR reconstructed PPG for 10 Different Subjects (Head Experiment)

Subject	Head HR (mean ± std)	Finger HR (Reference) (mean ± std)	IMAR Reconstructed HR (mean ± std)	ICA Reconstructed HR (mean ± std)
1	68.31 ± 19.25	59.23 ± 1.49	59.76 ± 0.22*	65.68 ± 20.98*
2	85.39 ± 34.53	71.55 ± 3.037	73.72 ± 0.31*	91.02 ± 35.48*
3	76.19 ± 8.88	77.39 ± 1.360	78.705 ± 0.33	68.06 ± 14.14*
4	94.47 ± 39.05	70.55 ± 3.686	73.66 ± 0.38*	75.32 ± 13.42*
5	72.33 ± 29.82	67.88 ± 4.643	66.83 ± 0.39	69.97 ± 20.20*
6	45.09 ± 10.06	51.44 ± 1.481	49.00 ± 0.09*	59.43 ± 22.97*
7	44.82 ± 24.47	59.82 ± 1.486	57.56 ± 0.21	64.49 ± 35.63*
8	63.46 ± 13.35	62.08 ± 0.865	62.23 ± 0.25	60.68 ± 10.70*
9	59.37 ± 30.85	49.05 ± 1.555	49.19 ± 0.20	60.27 ± 13.24*
10	46.89 ± 32.25	79.35 ± 1.323	78.93 ± 0.45	64.80 ± 25.60*

TABLE VII
Comparison Statistical Analysis of SpO2 Estimations from IMAR reconstructed PPG for 10 Different Subjects (Head Experiment)

Subject	Head SpO2 (mean ± std)	Finger SpO2 (Reference) (mean ± std)	IMAR Reconstructed SpO2 (mean ± std)	ICA Reconstructed SpO2
1	82.86 ± 4.86	97.70 ± 0.46	97.94 ± 0.93	76.721 ± 38.132*
2	80.33 ± 2.82	97.67 ± 0.47	97.972 ± 4.048*	111.097 ± 1.496*
3	87.20 ± 4.54	95.41 ± 0.49	98.53 ± 0.727*	74.081 ± 21.678*
4	87.36 ± 2.64	97 ± 0	97,13 ± 0.23	81.391 ± 11.81*
5	84.25 ± 3.76	98 ± 0	96.82 ± 5.25*	77.593 ± 22.16*
6	92.38 ± 2.64	98 ± 0	97.47 ± 0.97	84.069 ± 14.84*
7	85.18 ± 3.06	98.41 ± 0.49	96.68 ± 0.38	75.632 ± 17.24*
8	90.94 ± 2.38	99.82 ± 0.06	97.99 ± 0.38	89.322 ± 17.77*
9	83.93 ± 4.54	98 ± 0	99.61 ± 3.87*	100.15 ± 16.96*
10	84.94 ± 4.24	95.97 ± 0.67	96.53 ± 4.62	86.731 ± 19.305*

C. PPG Signal Reconstruction Performance in Finger Experiment

The performance of the signal reconstruction of the proposed IMAR approach is compared to ICA for the PPG data with an index finger moving left-to-right patterns. The pulse oximeter on the middle finger of the right hand, which was stationary, was used as the reference signal. Since the subjects were directed to produce the motions for 30 seconds within each 1-minute segment, corresponding to 50% corruption by duration, the window length of both clean and corrupted segments were both set as half length of the signal. Table VIII compares the HR reconstruction results between the IMAR and ICA methods for all 10 subjects. As shown in Table VIII, the IMAR reconstructed HR values are not significantly different from the reference HR in 7 out of 10 subjects. However, the ICA's reconstructed HR is significantly different from the reference HR in 8 out of 10 subjects indicating poor reconstruction fidelity.

TABLE VIII
Comparison Statistical Analysis of HR Estimations from IMAR reconstructed PPG for 10 Different Subjects (Finger Experiment)

Subject	Head HR (mean ± std)	Finger HR (Reference) (mean ± std)	IMAR Reconstructed HR (mean ± std)	ICA Reconstructed HR (mean ± std)
1	77.43 ± 1.91	70.61 ± 0.73	70.42 ± 0.42	77.32 ± 8.34*
2	63.60 ± 2.42	78.80 ± 0.41	78.36 ± 0.35	79.57 ± 9.68
3	70.82 ± 15.01	66.18 ± 0.76	67.21 ± 0.26	62.96 ± 22.53*
4	87.70 ± 20.53	72.59 ± 0.26	70.85 ± 0.34	73.58 ± 11.34*
5	84.34 ± 4.86	74.43 ± 0.29	73.51 ± 0.29*	77.62 ± 18.55*
6	81.75 ± 6.34	67.78 ± 0.36	69.07 ± 0.26*	67.75 ± 18.01
7	63.75 ± 3.05	57.57 ± 0.54	58.32 ± 2.49	52.51 ± 24.06*
8	66.75 ± 5.03	58.27 ± 0.75	60.34 ± 0.44*	61.64 ± 28.83*
9	97.27 ± 22.74	74.39 ± 0.46	74.25 ± 0.68	63.60 ± 14.96*
10	73.76 ± 2.85	61.58 ± 0.50	61.40 ± 0.35	50.80 ± 13.72*

D. PPG Signal Reconstruction Performance for the Walk and Stairs Climbing Experimental Data

The signal reconstruction of the MNA identified data segments of the walking and stairs climbing experiments using our proposed IMAR and its comparison to ICA are provided in this section. Detection of the MNA data segments was performed using the algorithm described in Part I of the companion paper. The reconstructed HR and SpO_2 values using our proposed algorithm and the ICA are provided in Tables IX and X, respectively. For both HR and SpO_2 reconstruction, the measurements were carried out using PPG data recorded from the head pulse oximeter. The right hand index finger's PPG data was used as HR and SpO_2 references. As shown in Table IX, 7 out of 9 subjects' reconstructed HR values were found to be not significantly different from the reference HR values using our algorithm. While 2 subjects' reconstructed HR values were found to be significantly different than the reference, the differences in the actual HR values are minimal. For ICA's reconstructed HR values, they all deviate significantly from the reference values.

For the reconstructed SpO_2 values, our algorithm again significantly outperforms ICA. All but one subject is not significantly different than the SpO_2 reference values for ICA. For our IMAR algorithm only 4 out of 9 subjects do not show significant difference from the reference values. Note the zero standard deviation reference SpO_2 values from Massimo's pulse oximeter in 7 out of 9 subjects. This is because Massimo uses a proprietary averaging scheme based on several past values. Hence, it is possible that the significant difference seen with our algorithm in some of the subjects may turn out to be not significant if the averaging scheme is not used. While some of the SpO_2 values from our algorithm are significantly different from the reference, the actual deviations are minimal and they are far less than ICA.

TABLE IX
Comparison Statistical Analysis of HR Estimations from IMAR reconstructed PPG for 9 Different Subjects (Walk & Stairs Climbing Experiment)

Subject	Head HR (mean ± std)	Finger HR (Reference) (mean ± std)	IMAR Reconstructed HR (mean ± std)	ICA Reconstructed HR (mean ± std)
1	62.16 ± 18.96	70.73 ± 5.80	70.55 ± 0.56	77.39 ± 11.90*
2	94.30 ± 20.37	94.40 ± 1.69	95.54 ± 0.86	92.94 ± 9.99*
3	105.53 ± 17.23	120.64 ± 2.98	122.00 ± 1.05	95.67 ± 13.10*
4	95.48 ± 8.37	101.61 ± 3.06	99.89 ± 0.44*	90.89 ± 8.28*
5	82.20 ± 13.07	86.99 ± 3.71	87.71 ± 1.07	82.84 ± 17.96*
6	77.40 ± 12.69	82.48 ± 1.68	81.93 ± 0.48	86.81 ± 12.54*
7	121.02 ± 19.26	107.58 ± 1.51	109.15 ± 0.07	138.62 ± 6.18*
8	86.57 ± 9.85	91.95 ± 6.07	91.73 ± 0.57	80.44 ± 4.61*
9	87.09 ± 6.56	82.55 ± 5.24	84.22 ± 1.93*	104.30 ± 21.43*

TABLE X
Comparison Statistical Analysis of SpO_2 Estimations from IMAR reconstructed PPG for 9 Different Subjects (Walk & Stairs Climbing Experiment)

Subject	Head SpO2 (mean ± std)	Finger SpO2 (Reference) (mean ± std)	IMAR Reconstructed SpO2 (mean ± std)	ICA Reconstructed SpO2
1	95.70 ± 7.62	99.00 ± 0	97.64 ± 2.50	84.21 ± 1.34*
2	94.55 ± 5.51	95.37 ± 0	96.37 ± 0.99	95.53 ± 1.59
3	91.00 ± 15.58	96.75 ± 0	94.51 ± 0.42*	84.64 ± 4.63*
4	89.61 ± 3.36	99.62 ± 0	102.25 ± 0.65*	87.33 ± 2.67*
5	94.27 ± 8.12	98.00 ± 0.50	97.34 ± 1.45	76.50 ± 1.53*
6	88.50 ± 13.95	96.00 ± 0.31	94.97 ± 4.07*	82.94 ± 1.05*
7	94.92 ± 16.77	98.00 ± 0	100.37 ± 3.15	90.69 ± 8.11*
8	96.11 ± 6.60	97.00 ± 0	98.70 ± 4.16*	96.11 ± 0.39

9	93.78 ± 6.60	98.62 ± 0	95.99 ± 2.39*	89.11 ± 5.03*

IV. DISCUSSION

In this study, a novel methodology IMAR is introduced to study motion artifact detection and reconstruction of PPG-based measurements. The proposed methodology is composed of two distinct steps, first using the SVM based approach for detection of corrupted segments in the PPG signal of interest and, in the second part (this paper), applying the introduced IMAR approach to reconstruct the corrupted segments of the original PPG signal. Comparing HR and SpO_2 estimations of the final reconstructed signal to the reference measurements have shown that the proposed IMAR method is a promising tool as the reconstructed values were found to be accurate. It has been also shown that the proposed IMAR approach far outperforms the ICA in motion artifact removal in all three experimental data. Finally, the simulation results from noise sensitivity analysis showed that SNR level down to -20dB and -15dB from additive white and colored noise, respectively, can be tolerated well by the application of the proposed IMAR procedure, compared to the SNR values of -10dB and -15dB for the ICA method.

REFERENCES

1. Chon, K.H., S. Dash, and K. Ju, *Estimation of respiratory rate from photoplethysmogram data using time-frequency spectral estimation.* IEEE Trans Biomed Eng, 2009. **56**(8): p. 2054-63.
2. Jubran, A., *Pulse oximetry*, in *Applied Physiology in Intensive Care Medicine*, G. Hedenstierna, et al., Editors. 2009, Springer Berlin Heidelberg. p. 45-48.
3. Thakor, N.V. and Z. Yi-Sheng, *Applications of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection.* Biomedical Engineering, IEEE Transactions on, 1991. **38**(8): p. 785-794.
4. Diniz, P., *Adaptive filtering: algorithms and practical implementation*. 2008: . Springer Science, Business Media L.L.C., .
5. Kalman, R.E., *A New Approach to Linear Filtering and Prediction Problems* Transactions of the ASME – Journal of Basic Engineering, 1960. **Series D**(82).
6. Bishop, G.W.a.G., *An introduction to the Kalman filter*. 2006: Tech. Rep.
7. Seyedtabaii, S.S.a.L., *Kalman filter based adaptive reduction of motion artifact from photoplethysmographic signal.* World Academy of Science, Engineering and Technology, 2008. **37**.
8. Morbidi, F., et al., *Application of Kalman Filter to Remove TMS-Induced Artifacts from EEG Recordings.* Control Systems Technology, IEEE Transactions on, 2008. **16**(6): p. 1360-1366.
9. S. Choi, A.C., H. min Park, and S. young Lee, *Blind source separation and independent component analysis: A review.* Neural Information Processing, Letters and Reviews, 2004. **6**(1): p. 1–57, .
10. Comon, P., *Independent component analysis, a new concept?* Signal Process., 1994. **36**(3): p. 287-314.
11. Thompson, B., *Canonical Correlation Analysis: Uses and Interpretation*. 1984: SAGE Publications.
12. Jolliffe, I.T., *Principal component analysis*. 1986: Springer-Verlang.
13. Elsner, J.B. and A.A. Tsonis, *Singular Spectrum Analysis: A New Tool in Time Series Analysis*. 1996: Springer.
14. Gao, J.Z., Chongxun; Wang, Pei., *Online Removal of Muscle Artifact from Electroencephalogram Signals Based on Canonical Correlation Analysis.* Clinical EEG and Neuroscience, EEG and Clinical Neuroscience Society ECNS, 2010.
15. Ram, M.R., et al. *Use of Multi-Scale Principal Component Analysis for motion artifact reduction of PPG signals*. in *Recent Advances in Intelligent Computational Systems (RAICS), 2011 IEEE*. 2011.
16. Golyandina, N., V. Nekrutkin, and A.A. Zhigljavsky, *Analysis of Time Series Structure: SSA and Related Techniques*. 2001: Taylor & Francis.
17. Teixeira, A.R., et al., *Automatic removal of high-amplitude artefacts from single-channel electroencephalograms.* Computer Methods and Programs in Biomedicine, 2006. **83**(2): p. 125-138.
18. Ghaderi, F., H.R. Mohseni, and S. Sanei, *Localizing Heart Sounds in Respiratory Signals Using Singular Spectrum Analysis.* Biomedical Engineering, IEEE Transactions on, 2011. **58**(12): p. 3360-3367.
19. Krishnan, R., B. Natarajan, and S. Warren, *Two-Stage Approach for Detection and Reduction of Motion Artifacts in Photoplethysmographic Data.* Biomedical Engineering, IEEE Transactions on, 2010. **57**(8): p. 1867-1876.
20. Feleppa, J.M.a.E.J., *Singular spectrum analysis applied to ultrasonic detection and imaging of brachytherapy seeds.* Acoust. Soc. Amer., 2007. **21,** (3): p. 1790–1801.
21. Hassani, H., *Singular Spectrum Analysis: Methodology and Comparison.* Journal of Data Science 2007. **5**(1): p. 239-257.
22. Kloeden, P.E., & Platen, E. , *Numerical Solution of Stochastic Differential Equations*. 1992: Springer.
23. Reddy, K.A., et al. *A Novel Method of Measurement of Oxygen Saturation in Arterial Blood*. in *Instrumentation and Measurement Technology Conference Proceedings, 2008. IMTC 2008. IEEE*. 2008.
24. Reddy, K.A., et al., *A Novel Calibration-Free Method of Measurement of Oxygen Saturation in Arterial Blood.* Instrumentation and Measurement, IEEE Transactions on, 2009. **58**(5): p. 1699-1705.

Respiratory Rate Estimation from the Built-in Cameras of Smartphones and Tablets

[1]Yunyoung Nam, [2]Jinseok Lee, and [1]Ki H Chon*

[1]Department of Biomedical Engineering, Worcester Polytechnic Institute, MA 01609 USA

[2]Department of Biomedical Engineering, Wonkwang University School of Medicine, Iksan, Jeonbuk, Republic of Korea

* Corresponding Author:

Ki H. Chon, Ph.D.

Email: kichon@wpi.edu,

Tel: +15088314114,

Fax: +15088314121

Abstract—This paper presents a method for respiratory rate estimation using the camera of a smartphone, an MP3 player or a tablet. The iPhone 4S, iPad 2, iPod 5, and Galaxy S3 were used to estimate respiratory rates from the pulse signal derived from a finger placed on the camera lens of these devices. Prior to estimation of respiratory rates, we systematically investigated the optimal signal quality of these 4 devices by dividing the video camera's resolution into 12 different pixel regions. We also investigated the optimal signal quality among the red, green and blue color bands for each of these 12 pixel regions for all four devices. It was found that the green color band provided the best signal quality for all 4 devices and that the left half VGA pixel region was found to be the best choice only for iPhone 4S. For the other 3 devices, smaller 50x50 pixel regions were found to provide better or equally good signal quality than the larger pixel regions. Using the green signal and the optimal pixel regions derived from the 4 devices, we then investigated the suitability of the smartphones, the iPod 5 and the tablet for respiratory rate estimation using three different computational methods: the autoregressive (AR) model, variable-frequency complex demodulation (VFCDM), and continuous wavelet transform (CWT) approaches. Specifically, these time-varying spectral techniques were used to identify the frequency and amplitude modulations as they contain respiratory rate information. To evaluate the performance of the three computational methods and the pixel regions for the optimal signal quality, data were collected from 10 healthy subjects. It was found that the VFCDM method provided good estimates of breathing rates that were in the normal range (12-24 breaths/min). Both CWT and VFCDM methods provided reasonably good estimates for breathing rates that were higher than 26 breaths/min but their accuracy degraded concomitantly with increased respiratory rates. Overall, the VFCDM method provided the best results for accuracy (smaller median error), consistency (smaller interquartile range of the median value), and computational efficiency (less than 0.5 s on 1 min of data using a MATLAB implementation) to extract breathing rates that varied from 12-36 breaths/min. The AR method provided the least accurate

respiratory rate estimation among the three methods. This work illustrates that both heart rates and normal breathing rates can be accurately derived from a video signal obtained from smartphones, an MP3 player and tablets with or without a flashlight.

***Keywords*- Respiratory rate estimation, Autoregressive model, Continuous wavelet transform, Variable frequency complex demodulation method, Smartphone, Tablet.**

I. INTRODUCTION

RESPIRATORY rate is an important indicator for early detection and diagnosis of potentially dangerous conditions such as sleep apnea[24], sudden infant death syndrome[18], cardiac arrest[3] and chronic obstructive pulmonary disease[5]. In addition, for some patients who undergo surgery, relative changes in respiratory rates are much greater than changes in heart rate or systolic blood pressure, thus, respiratory rates can be an important vital sign indicator[21]. Respiratory rate is most accurately measured using transthoracic impedance plethysmography[1], nasal thermocouples[20] or capnography[16]. However, these methods all require expensive external sensors which may require donning a mask, nasal cannula or chest band sensors. More importantly, since these devices may disturb natural breathing and sleep positions, they are mostly applicable in constrained environments such as operating rooms and intensive care units.

Recently, photoplethysmography (PPG) has been widely considered for respiratory rate extraction due to its simplicity and non-invasive measurement capability[11–13]. The PPG signal contains components that are synchronous with respiratory and cardiac rhythms. Indeed, the respiratory rhythm is modulated by frequency and/or amplitude of the cardiac rhythm. The occurrence of temporal variations of frequency and amplitude is characteristic of the respiratory sinus arrhythmia[6]. Thus, the respiratory rate can be obtained by detecting the presence of either amplitude modulation (AM) or frequency modulation (FM) in the PPG signal[2].

Numerous advanced signal processing algorithms (both parametric and nonparametric approaches) have been applied to extract respiratory rates by looking for AM or FM signatures from a PPG signal[2,19]. For a parametric approach, the autoregressive (AR) model approach has been shown to provide relatively good respiratory rate estimation[7–10]. For nonparametric approaches, time-frequency spectrum (TFS) methods such as continuous wavelet transform (CWT) and variable frequency complex demodulation method (VFCDM) have also been shown to provide accurate respiratory rate estimation[2,11–13].

To our knowledge, respiratory rate estimation using the camera of either a smartphone or a tablet has never been demonstrated nor discussed in the literature. We have recently demonstrated that a pulsatile signal that has similar dynamics to that of a PPG signal can be obtained from a smartphone's camera when a fingertip is pressed onto it[4,19]. Utilizing these pulsatile signals derived from an iPhone, we have also shown that accurate detection of atrial fibrillation can be made[17]. Given these advances, the aims of this work were: 1) a systematic examination of the pulsatile signal quality derived from a video camera from several measurement modalities including iPhone 4S, iPad 2, iPod 5, and Galaxy S3; and 2) to determine if accurate respiratory rates can be estimated directly from the pulsatile signals of the different measurement modalities. The challenge here is that PPG signals are often sampled at greater than 100 Hz whereas most smartphones' video sampling rates are no more than 30 Hz. Since previous studies have shown good estimation of respiratory rates using the AR model, CWT, and VFCDM from a PPG signal, we also use these methods to compare the accuracy of breathing rates from pulsatile signals obtained from various models of a smartphone, MP3 player (iPod5) and a tablet.

II. METHODS

A. Data Collection

Data were collected on 10 healthy subjects on 2 separate occasions using 4 different devices: iPhone 4S, iPad 2, iPod 5, and Galaxy S3. Only two devices were used simultaneously for data collection in a given experimental setting. For the pulsatile signal acquisition, we used the Objective-C programming language and the Xcode platform for iPhone 4S, iPad 2, and iPod 5; Java was used for the Galaxy S3 on the mobile platform Android 4.1 (Jelly Bean). Specifically, we used Eclipse IDE Indigo R2 for the development environment and Samsung Galaxy S3 for the development and debugging purposes. For the video recordings of iPhone, iPad, and iPod, we examined four different sizes of pixel regions: 50×50, 320×240 (QVGA), 640×240 (vertical HVGA), and 640×480 (VGA) for determining the optimal signal quality. For all five different pixel sizes, the pulsatile signal was

obtained by averaging the entire pixel size for each of the three color bands (red, green and blue) for every frame. All four devices provided sampling rate close to 30 frames per second. However, when the video sampling rate was lower than 30 Hz, a cubic spline algorithm was used to interpolate the signal to 30 Hz.

No subject had cardiorespiratory pathologies. Data were collected in the sitting upright position, and the sensor was placed in proximity to the subject's left index or middle finger as shown in Fig. 1. All subjects were instructed to breathe at a metronome rate according to a timed beeping sound, i.e., to start inspiring when a beep sound starts and to expire before the next beep sound occurs. The data were collected for breathing frequencies ranging from 0.2 to 0.9 Hz at an increment of 0.1 Hz. Prior to data collection, all subjects were acclimated to the breathing frequency rate being measured. Three minutes of data were collected for each frequency for each subject. Electrocardiogram (ECG) recordings were collected with an HP 78354A acquisition system using a standard 5-lead electrode configuration. A respiration belt was placed around a subject's chest and abdomen to monitor the true breathing rate (Respitrace Systems, Ambulatory Monitoring Inc.). Respiratory and ECG recordings were obtained using the LabChart software (ADInstruments) at a sampling rate of 400 Hz. Fig. 1 shows data collection on the four devices by placing a fingertip on the video camera.

B. Extraction of Respiratory Rates

1) VFCDM: Detection of AM and FM from a pulsatile signal using the power spectral density (PSD) is difficult since the dynamics are time-varying, hence, require high resolution time-frequency spectral (TFS) methods to resolve them. We have recently shown that because the VFCDM method provides one of the highest time-frequency spectral (TFS) resolutions, it can identify AM and FM dynamics. Consequently, Fourier transform of either the AM or FM time series extracted from the heart rate frequency band can lead to accurate estimation of respiratory rates when the acquired signal is PPG data[23].

Details concerning the VFCDM algorithm are described in [23]. Hence, we will only briefly describe

the main essence of the algorithm. The VFCDM method thus involves a two-step procedure. The first step is to use complex demodulation (CDM) or what we termed the fixed frequency CDM (FFCDM) to obtain an estimate of the TFS, and the second step is to select only the dominant frequencies of interest for further refinement of the time-frequency resolution using the VFCDM approach. In the first step of the VFCDM method, a bank of LPFs is used to decompose the signal into a series of band-limited signals. The analytic signals that are obtained from these, through use of the Hilbert transform, then provide estimates of the instantaneous amplitude, frequency, and phase within each frequency band.

2) CWT: As described in the Introduction section, numerous studies[11–13] showed relatively good results using the continuous wavelet transform for extraction of respiratory rates directly from a pulse oximeter. The Morlet wavelet was also applied to a half-length of five samples at the coarsest scale for estimating the scalogram of the pulsatile signal[22]. The procedures of the CWT for extracting respiratory rates is nearly identical to the VFCDM in that identified AM and FM series are Fourier transformed to estimate respiratory rates.

3) AR Modeling: This approach involves estimation of autoregressive (AR) model parameters using the optimal parameter search (OPS) criteria[15]. The AR parameters are formulated as the transfer function followed by factorization into pole terms. The real and complex conjugate poles define the power spectral peaks with the larger magnitude poles corresponding to higher magnitude peaks. The resonant frequency of each spectral peak is given by the phase angle of the corresponding pole. Among the poles, we set the region of interest for respiratory rates between 0.15 Hz and 1 Hz. The details of the respiratory algorithm using the AR model are described in [7].

C. Data Analysis

Using PPG signals with sampling rates of at least 250 Hz to derive heart rates has previously been shown to be a good alternative to ECG monitoring[14]. However, sampling rates for most smart phone and tablet video cameras range from 25-30 Hz. Given these low sampling rates, it is necessary to

determine the accuracy of the smart phone and tablet devices in estimating heart rates and respiratory rates. Comparisons of derived heart rates were made between the standard ECG, smartphones and tablets. We used our own peak detection algorithm to determine R- wave peaks from the ECG signals and cardiac pulse peaks from the phone camera PPG signal. Due to the frame rate variability, we interpolated the pulsatile signal to 30 Hz using a cubic spline algorithm followed by the peak detection. The peak detection algorithm incorporated a filter bank with variable cutoff frequencies, spectral estimates of the heart rate, rank-order nonlinear filters and decision logic.

Three minutes of data sampled at 30 Hz were low-pass-filtered to 1.78 Hz, and then downsampled to 15 Hz. We performed the extraction of the respiratory rate on every 1-minute segment of pulsatile signal, and then the data were shifted by every 10 seconds for the entire 3 minutes of recordings, i.e., each 1-minute dataset had a 50 second overlap. Thus, for each 3-minute segment, we had thirteen 1-minute segments to analyze for all methods to be compared. Thus, 3 minutes of data were sufficiently long to test the efficacy of each method but not too long in duration to fatigue the subjects as their task was to breathe on cue with a metronome-timed beep sound. For the VFCDM and CWT methods, for every 1-minute segment, the initial and final 5 seconds of the TFS were not considered because the TFS has an inherent end effect which leads to inaccurate time-frequency estimates. For the CWT method, the lower and upper frequency bounds of the analyzed signal were set to 0.01 and 0.5, respectively. The filter parameters of the VFCDM were set to the first cutoff frequency $F_w = 0.03$ Hz, second cutoff frequency $F_v = 0.015$ Hz, and filter length $N_w = 64$. We have previously shown that the parameter $F_v = F_w / 2$, and that N_w is chosen to be approximately half the data length. For each of these categories, detection errors were found for each frequency for all subjects using the four different methods. The error ε is calculated as follows:

$$\varepsilon = \frac{\sum_{i=1}^{n} |R_D^i - R_T^i|}{n}, \tag{1}$$

where R_D and R_T denote the detected breathing rate and the true breathing rate, respectively.

III. RESULTS

A. Selection of the Best Color Band and the Optimal Video Pixel Size for Estimation of Heart Rates

Fig. 2a shows the orientation of the Field of View (FOV) of each camera relative to the location of the camera flash. All references to "left" and "right" in this paper are relative to the camera FOV, regardless of whether the camera itself was on the front or rear of the device. Note that when a device's front video camera is on, what is displayed in the LCD display of the device is a mirror image of the FOV of the front camera. The stored video will revert to the FOV view, but until the videotaping is complete, the display in the LCD of the device will be the mirror image of the actual front camera FOV. This is to match people's expectations as they look in the display while photographing themselves. However, reversal in the display was not taken into account to avoid confusion, and because we used the video feed directly before it was processed for display on the device's LCD.

Figs. 2b and 2c provide details of the video pixel regions examined on all four devices and they consist of the following 12 video regions: left top (LT), left middle (LM), left bottom (LB), right top (RT), right middle (RM), right bottom (RB), middle top (MT), center (C), middle bottom (MB), vertical left half-VGA (vertical left HVGA), vertical right half-VGA (vertical right HVGA) and VGA.

All results shown are based on average values from 10 subjects. Table I shows pulse amplitudes for each of the three color band signals extracted by averaging 50×50 pixels of an iPhone 4S rear camera video signal. The red color band signal extracted from the LM and LB regions of 50×50 pixel size on the camera had a saturated *R* value of 256 at all times when the flashlight was turned on. Thus, pulse variations in red were not detected in the video signal. When the flashlight was on (back camera displays for iPhone 4S, iPod 5 and Galaxy S3), the green color consistently provided significantly higher mean amplitude values than either the blue or red color for all five regions, as shown in Table I. Higher mean amplitude values for the green color were also observed when the flashlight was off,

but for the RB region, the red color had significantly higher mean amplitude values than either the green or blue. In addition, for the RM region, similar to the green color, the blue color also had larger amplitude values when compared to the red color.

Table II shows experimental results of R-R intervals (RRIs) extracted from ECG and three-color band pulsatile signals (PS) from an iPhone 4S. As shown in Table II, the PS values from the smart phone are an excellent surrogate to RRIs derived from ECG for all colors. There was no statistical difference between RRIs derived from ECG and each of the three color PS; the median errors calculated using Eq. (1) are also very small for all three color band signals. Fig. 3 shows the Bland-Altman plot for the mean HR data from the iPhone 4S (green color) and the ECG. The Bland-Altman plot shows a mean difference of 0.074 and that most of the data are within the 95% confidence intervals.

Having established that the green color signal provides the best signal amplitude values for an iPhone 4S, we now systematically investigate which pixel regions of the green color give the most optimal signal quality as determined by the largest amplitude values for all four devices. Specifically, 9 different regions of 50×50 pixels, the left and right pixel regions of HVGA, and the entire VGA pixel region were investigated for the best signal quality. Since the iPhone 4S results shown in Table I were based on only the 5 regions of 50×50 pixels, we also investigated the left and right regions of HVGA and the VGA regions for determining the optimal signal quality. Table III shows the mean amplitude values of the green color pulse signal for different pixel regions of the 4 devices. For iPhone 4S, the left region of HVGA had the largest amplitude value among the twelve regions, as expected, since the LED flash is placed on the left side of the camera's FOV (see Fig. 2a). For the iPad 2, the device was held vertically on a desk, in landscape mode, so we chose also to consider the FOV in landscape mode. In this case, the right side of the portrait mode FOV was turned to be on top, and the left side was on the bottom. The RT and RM regions of 50×50 pixels and the right region of HVGA have among the largest amplitude values since the light source was from the ceiling of the room, i.e. closest to the top

in landscape mode. For the iPod 5, the LT and LM regions of 50×50 pixels and the VGA have the largest amplitude values. All left values exceed right values because the flash is on the left side of the camera's FOV (see Fig. 2). For the Galaxy S3, the RT, RM and RB regions have the largest amplitude values among the twelve regions as expected since the LED flash is placed to the right of the camera's FOV (see Fig. 2). Hence, depending on the location of the LED flash, the left or right HVGA or 50×50 regions of the green color pulsatile signal have the highest intensity value among all regions tested.

Table IV shows experimental results of the heart rate extracted from the ECG and three different color PS derived from an iPad 2 with only ambient light. The green and red colors have similar PS values and are not found to be different from RRI derived from the ECG signal. The blue color's PS significantly deviates from the RRI derived from the ECG signal. Hence, both red and green colors' PS are good surrogates for ECG; also the median errors are negligible.

Table V shows experimental results of heart rate extracted from ECG and three-color band signals from the left bottom region (50×50 pixels) of an iPod 5 when the flashlight was turned on or off. As noted earlier, the red color signal saturates when the flashlight is on, hence, there are no values to report. When the flashlight is on for the LB region, both green and blue colors' PS signals are not found to be statistically different from RRI derived from the ECG signals. When the flashlight is off, only the red signal is not statistically different from the RRI derived from the ECG signals. For the LM region, all colors' RRI are not statistically different from the ECG except for the blue color when the light is turned off.

B. Heart Rate, Frequency Spectrum and Power Spectrum

Figs. 4a-c show an example of a representative 1-minute segment of iPhone 4S PS data, its time-frequency spectrum of the green band signal via the VFCDM, and the power spectral density (PSD) of the AM and FM signals derived from the HR frequency band (e.g., ~1-1.5 Hz), respectively, while a subject was breathing at a metronome rate of 18 breaths/min. Note the similarity of the PS in Fig. 4a

to those of commercially-available PPG signals. As shown in Fig. 4c, the PSD of the extracted AM and FM time series show the largest peaks at 0.3 Hz; these peaks correspond accurately to the true respiratory rate of 18 breaths/minute.

C. Respiratory Rate

The true respiratory rates were derived by taking the PSD of the respiratory impedance trace signals during metronome breathing experiments. True respiratory rates from the respiration trace and the estimated breathing rates from the green signal using both the FM and AM sequences from the VFCDM and CWT were compared using metronome rates ranging from 0.2-0.9 Hz. In order to evaluate the four computational methods, we provide figures and tables that show the accuracy and repeatability of each method as a function of the true breathing rate. For tabulating results, we grouped the results for 0.2-0.3 Hz together and designated them as the low frequency (LF) breathing rates. Likewise, the results for 0.4-0.6 Hz breathing rates were lumped together and designated as the high frequency (HF) breathing rates. Since the percentage errors were found to be not-normally distributed, we report the median and inter-quarter range (IQR) values.

Fig. 5 shows the subjects' variations of percentage detection error in the form of box plots for the left region of the HVGA pixel resolution with flash on since this region was found to have the best signal quality as shown in Table III. The top and bottom panels of Fig. 5 represent results for the LF and HF breathing rates, respectively. The lower boundary of the box closest to zero indicates the 25th percentile, a line within the box marks the median, and the upper boundary of the box farthest from zero indicates the 75th percentile. Whiskers (error bars) above and below the box indicate the 90th and 10th percentiles. Therefore, the area of the blue box is an indication of the spread, i.e., the variation in median error (or IQR), across the population. These figures indicate how well the algorithms perform across the entire population. Red crosses represent the 5th and 95th percentiles.

As shown in Fig. 5, the AR model approach is the least accurate followed by CWT-AM, CWT-FM, and VFCDM (both AM & FM approaches) when we consider all breathing frequencies. Note that the

variances of the median values as determined by ε (the average respiratory estimation error as defined in Eq. 7) are significantly lower for both VFCDM and CWT than for AR model approach. Although there was no significant difference in the median error between CWT and VFCDM methods at 0.3 Hz, ε is found to be the lowest for VFCDM- FM at 0.2 Hz. In general, ε is larger for HF than LF breathing rates for all computational methods. For HF breathing rates, ε is lowest for CWT-FM, followed by VFCDM, CWT-AM, and AR model. While there is no significant difference in the variance between VFCDM-FM and CWT-FM, both methods have significantly less variance than either CWT-AM or VFCDM-AM or AR model. Thus, gauging the accuracy as defined by the median errors and their variances, as shown in Fig. 5, we observed that for HF breathing rates, CWT-FM consistently provides significantly lowest median errors and variance values.

Figs. 6 and 7 show the subjects' variation of percentage detection errors in the form of box plots, which were extracted from front cameras of an iPhone 4S and an iPad 2 (no flash), respectively, for the left HVGA region. While not shown, the left HVGA region also had the best signal quality with the flashlight off for an iPhone 4S. The top and bottom panels of Figs. 6 and 7 represent results from a front camera for the LF and HF breathing rates, respectively. The AR model approach is the least accurate followed by CWT and VFCDM methods when we consider all breathing frequencies. For LF breathing rates, there was no significant difference in the median error between VFCDM methods. However, the variances of the median values as determined by ε are significantly lower for both VFCDM and CWT than for AR model approaches. In general, ε is larger in HF than LF breathing rates. For HF breathing rates, ε is lowest for CWT-FM, followed by VFCDM, CWT-AM, and AR model. While there is no significant difference in the variance between VFCDM-FM and VFCDM-AM in LF breathing rate, median errors of VFCDM-FM are significantly lower than that of VFCDM-AM. Thus, gauging the accuracy as defined by the median errors and their variances, as shown in Figs. 6 and 7, we observed that for both LF and HF breathing rates, CWT-FM consistently provides the lowest median errors and variance values.

Figs. 8 and 9 show the subjects' variation of percentage detection error in the form of box plots, which were extracted from front cameras of a Galaxy S3 and an iPod 5, respectively, both from the 50×50 pixel resolutions in the LT for the former and LM regions for the latter. The top and bottom panels of Figs. 8 and 9 represent results from a rear camera for the LF and HF breathing rates, respectively. The AR model approach is the least accurate followed by CWT and VFCDM methods when we consider all breathing frequencies. For LF breathing rates, there was no significant difference in the median error between VFCDM methods. However, the variances of the median values as determined by ε are significantly lower for both VFCDM and CWT than for AR model approaches. ε is larger in HF than LF breathing rates. For HF breathing rates, ε is lowest for CWT-FM. While there is no significant difference in the variance between VFCDM-FM and VFCDM-AM in LF breathing rate, median errors of VFCDM-FM are significantly lower than that of VFCDM-AM. Thus, gauging the accuracy as defined by the median errors and their variances, as shown in Figs. 8 and 9, we observed that for both LF and HF breathing rates, VFCDM-FM most often provides the lowest median errors and variance values.

Table VI shows the numerical statistics (IQR) for the "repeatability" across the population of test subjects. The results for 0.2-0.4 Hz (LF breathing range) breathing rates are much better than for 0.5-0.6 Hz (HF breathing range), and in addition, the tracking ability of the breathing rate detection method is much better when CWT and VFCDM methods are used for the LF. Even though the AR method shows significantly lower values of IQR errors than all the other methods studied here, the AR method provided relatively high median errors. For each of the four different devices, the VFCDM-FM method has significantly lower IQR errors ($\varepsilon < 7$) and median errors ($\varepsilon < 6$) than those of any other devices in the 0.2-0.4 Hz breathing rate range.

ANOVA and the Bonferroni *t* test were used for analysis of differences between the medians for the seven different methods. Statistical significance was identified as $P < 0.05$. Tables VII and VIII provide a summary of the statistical analysis comparing the performance of the five methods (AR,

CWT-AM, CWT-FM, VFCDM-AM and VFCDM-FM) to each other. For Tables VII and VIII, we list only those comparison that show significant difference among the five computation methods for each device for both LF and HF breathing ranges.

Table IX summarizes these measures of median and IQR errors for 0.7, 0.8, and 0.9 Hz breathing rates – rates above what we termed HF rates. As presented numerically in the table, we observe that WT-FM provides the lowest median error at the 0.7 Hz breathing rate, and might be acceptable. However, no method provided reasonably good estimates of breathing rates above the 0.7 Hz breathing rate.

Fig. 10 shows the subjects' variation of percentage detection error in the form of box plots extracted from a rear camera (with flashlight on) of an iPhone 4S during spontaneous breathing. True respiration rate was found by computing the PSD of the impedance respiration trace signal and finding the frequency at the maximum amplitude using a respiration belt. The variances of the median values as determined by ε are significantly lower for both VFCDM and CWT than for the AR model approach. In the normal range (11-27 breaths/minute), VFCDM-FM consistently provides the lowest median errors and variance values. As shown in Table X, there was no significant difference in the median error among WT-AM, WT-FM, VFCDM-FM, and VFCDM-AM during spontaneous breathing, the accuracy of AR is lower than other approaches.

In general, the ability of the methods to provide consistent results is especially excellent (highest) for both the CWT-FM and VFCDM methods, for both LF and HF breathing rates. As with the accuracy results, the repeatability is also better for the LF than for the HF breathing rates for all four methods. Both CWT-FM and VFCDM provide significantly more repeatable results than either CWT-AM or AR model.

D. Computation Time

Table XI shows the computational time for heart rate extraction based on the choice of pixel resolution and the number of color bands used. As shown in the table, pixel resolutions of QVGA and

HVGA result in a frame rate of 25 frames/s when only one color is selected. The frame rates extracted from two and three colors are 23 and 20 frames/s, respectively, in the case of HVGA resolution.

The clock speed of the CPU used in the iPhone 4S and iPod 5 is 800MHz. The latest iPhone 5 is clocked at 1.02GHz. The recently released Samsung Galaxy S4 is equipped with a 1.9 GHz Quad-core processor. Thus, for most new smartphone and tablet cameras, higher than 30 frames/s can be achieved, suggesting that a choice of higher pixel resolution will not be a significant problem for accurate and real-time detection of heart rates and respiratory rates.

IV. DISCUSSION

In this work, we tested several smartphones and tablets for their feasibility in estimating respiratory rates using the pulsatile signals derived from a resident video camera and flashlight, when available. The motivation for this work is based on several recent works which showed that accurate respiratory rates, especially at normal breathing rates, can be obtained from pulse oximeters[11–13]. The characteristics of the pulsatile signal derived from cameras in smartphones and tablets are similar to PPG signals, hence, similarly-accurate respiratory rates can be obtained, theoretically. Our results do indicate that certainly for normal breathing ranges (0.2-0.3 Hz), this is feasible from pulsatile signals derived from smartphone and tablet video cameras.

We have optimized the accuracy of the respiratory rates by first systematically analyzing the optimal pixel resolution of the video signal for the attainment of the strongest pulsatile signal strength. It is logical to assume that the greater the amplitude of the pulsatile signal, the higher the signal's strength with the proviso that care is taken to minimize motion artifacts during measurements. Our results showed that a choice of larger pixel resolutions does not necessary result in higher pulsatile signal amplitude. For example, for the Galaxy S3, iPod5 and iPad2, 50×50 resolution provided either the highest pulsatile amplitude or was statistically equivalent to HVGA resolution. In fact, HVGA resolution was the best choice only for the iPhone 4S. The important implication of having a smaller

pixel region providing just as good or better signal quality than a larger pixel region is the significant reduction in the computational time so that real-time calculation of respiratory rates can be attained.

Commercial pulse oximeters in either transmittance or reflectance mode normally employ a single photodetector (PD) element, typically with an active area of about 6-10mm^2. The image sensor size of the iPhone 4S is 4.54×3.42=15.5268 mm^2. Consequently, when signals are extracted from HVGA (320x480 pixels) video mode, the active area is 2.27×3.42=7.7634 mm^2. Hence, we initially thought that motion artifact and noise can be significantly reduced by increasing the active area in the sensor. However, our investigation revealed that larger pixel resolutions do not necessary result in a higher signal-to-noise ratio.

We compared AR-based approaches, CWT, and VFCDM for respiratory rate estimation from smartphones and a tablet because these techniques have been shown to provide good results from PPG signals. Similar to PPG signal results, the VFCDM-FM provided the most accurate respiratory rate estimation with the fastest computational time than any of the methods compared in this study for the LF breathing rate. For HF breathing rates, both CWT and VFCDM methods provided comparable results. The CWT approach using either the FM or AM signals fared better than the AR method but at the expense of higher computational time.

Due to the inherent non-stationarity in the respiratory rate, a time-frequency method is needed and appears to be the most appropriate approach. Another advantage of the time-frequency spectral approach to estimating respiratory rates is that unlike most filtering approaches, tuning of a number of parameters specific to each subject is not required. Note that in our work, we have used the same parameters (as described in the Methods section) for both CWT and VFCDM for all subjects and for all breathing rates.

As was the case with respiratory rate estimation using the PPG signal, the computational speed of the VFCDM method is faster than that of the wavelet method for smartphone and tablet data. The average time to calculate the respiration frequency using the VFCDM method was found to be around

1.4 seconds, while using the wavelet method took 37.8 seconds on average (programs running on MATLAB R2012a). The AR spectral method was the fastest as it took only 0.2 seconds on average using MATLAB, and this computation time includes the time needed to calculate the model order based on an initial model order selection of 50. However, the AR method is the least accurate in respiratory rate estimation.

All three methods showed increased estimation errors as the breathing rates increased, for all devices tested. This observation was also noted for the PPG signal[2]. We have also examined breathing rates of 0.7 Hz, 0.8 Hz and 0.9 Hz, and the results showed significant deviation from the true breathing rates for all 3 methods. Both CWT and VFCDM methods provided comparable results with significantly worse estimates for the AR method which was also the case with both LF and HF breathing rates. Hence, our results show that it is feasible to obtain good results for the normal breathing rates but not higher breathing rates (i.e., > 0.5 Hz). We can speculate that there are two reasons for inaccurate results for high breathing rates. First, detection of both AM and FM phenomenon requires persistent oscillations for several cycles, but with faster respiratory rates, our decision to limit the data segment to 1 minute may not be sufficient. More importantly, with faster breathing rates, the AM or FM phenomenon becomes less apparent, and thus, it becomes more difficult to detect them no matter how sophisticated the detection may be.

In summary, our work was undertaken to determine the optimal pixel resolution and location as well as the color band for obtaining the best quality signal so that we maximize the accuracy of respiratory rate estimation from a video signal from either smartphones or tablets. It was found that a larger pixel resolution does not necessarily result in better signal quality. In fact in most scenarios, a 50×50 pixel resolution was just as good as or better than HVGA resolution. In addition, we found that the region closest to the flash in most cases resulted in a higher signal quality which is logical and expected. Finally, using the optimum pixel size, location and color band of the pulsatile signal, we found accurate respiratory estimates especially in the normal breathing ranges. We found increased breathing rate

estimation errors as the respiratory rates increased higher than 0.5 Hz with unreliable results at 0.6 Hz or higher. When both computational time and estimation accuracy are taken into account, the VFCDM-FM provided the best results among all approaches examined in this work. This work allows attainment of at least two vital sign measurements all directly from a finger pressed onto a video camera of either a smartphone or tablet: the heart rate and respiratory rate. It is expected that future work by either our laboratory or others will result in additional other vital sign capabilities directly from a video signal acquired from either a smartphone or tablet.

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

REFERENCES

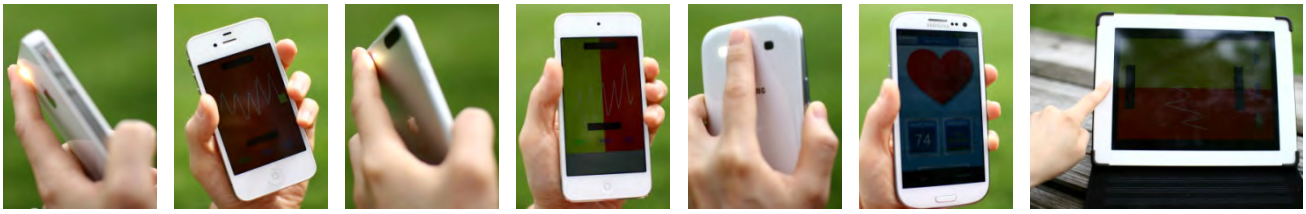
1. Allison, R., E. Holmes, and J. Nyboer. Volumetric dynamics of respiration as measured by electrical impedance plethysmography. *J. Appl. Physiol.* 19:166–173, 1964.
2. Chon, K. H., S. Dash, and K. Ju. Estimation of respiratory rate from photoplethysmogram data using time–frequency spectral estimation. *Biomed. Eng. IEEE Trans.* 56:2054–2063, 2009.
3. Fieselmann, J. F., M. S. Hendryx, C. M. Helms, and D. S. Wakefield. Respiratory rate predicts cardiopulmonary arrest for internal medicine inpatients. *J. Gen. Intern. Med.* 8:354–360, 1993.
4. Grimaldi, D., Y. Kurylyak, F. Lamonaca, and A. Nastro. Photoplethysmography detection by smartphone's videocamera. , 2011.
5. Hasselgren, M., M. Arne, A. Lindahl, S. Janson, and B. Lundbäck. Estimated prevalences of respiratory symptoms, asthma and chronic obstructive pulmonary disease related to detection rate in primary health care. *Scand. J. Prim. Health Care* 19:54–57, 2001.
6. Hirsch, J., and B. Bishop. Respiratory sinus arrhythmia in humans: how breathing pattern modulates heart rate. *Am. J. Physiol.-Heart Circ. Physiol.* 241:H620–H629, 1981.
7. Lee, J., and K. Chon. Respiratory rate extraction via an autoregressive model using the optimal parameter search criterion. *Ann. Biomed. Eng.* 38:3218–3225, 2010.
8. Lee, J., and K. H. Chon. An autoregressive model-based particle filtering algorithms for extraction of respiratory rates as high as 90 breaths per minute from pulse oximeter. *Biomed. Eng. IEEE Trans.* 57:2158–2167, 2010.
9. Lee, J., and K. H. Chon. Time-varying autoregressive model-based multiple modes particle filtering algorithm for respiratory rate extraction from pulse oximeter. *Biomed. Eng. IEEE Trans.* 58:790–794, 2011.
10. Lee, J., J. P. Florian, and K. H. Chon. Respiratory rate extraction from pulse oximeter and electrocardiographic recordings. *Physiol. Meas.* 32:1763, 2011.
11. Leonard, P. A., D. Clifton, P. S. Addison, J. N. Watson, and T. Beattie. An automated algorithm for

determining respiratory rate by photoplethysmogram in children. *Acta Paediatr.* 95:1124–1128, 2006.

12. Leonard, P., T. Beattie, P. Addison, and J. Watson. Standard pulse oximeters can be used to monitor respiratory rate. *Emerg. Med. J.* 20:524–525, 2003.
13. Leonard, P., N. R. Grubb, P. S. Addison, D. Clifton, and J. N. Watson. An algorithm for the detection of individual breaths from the pulse oximeter waveform. *J. Clin. Monit. Comput.* 18:309–312, 2004.
14. Liu, H., Y. Wang, and L. Wang. A review of non-contact, low-cost physiological information measurement based on photoplethysmographic imaging. , 2012.doi:10.1109/EMBC.2012.6346371
15. Lu, S., K. H. Ju, and K. H. Chon. A new algorithm for linear and nonlinear ARMA model parameter estimation using affine geometry. *Biomed. Eng. IEEE Trans.* 48:1116–1124, 2001.
16. Mason, K. P., P. E. Burrows, M. M. Dorsey, D. Zurakowski, and B. Krauss. Accuracy of capnography with a 30 foot nasal cannula for monitoring respiratory rate and end-tidal CO2 in children. *J. Clin. Monit. Comput.* 16:259–262, 2000.
17. McManus, D. D., J. Lee, O. Maitas, N. Esa, R. Pidikiti, A. Carlucci, J. Harrington, E. Mick, and K. H. Chon. A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation. *Heart Rhythm* 10:315 – 319, 2013.
18. Rantonen, T., J. Jalonen, J. Grönlund, K. Antila, D. Southall, and I. Välimäki. Increased amplitude modulation of continuous respiration precedes sudden infant death syndrome:–Detection by spectral estimation of respirogram. *Early Hum. Dev.* 53:53–63, 1998.
19. Scully, C. G., J. Lee, J. Meyer, A. M. Gorbach, D. Granquist-Fraser, Y. Mendelson, and K. H. Chon. Physiological parameter monitoring from optical recordings with a mobile phone. *Biomed. Eng. IEEE Trans.* 59:303–306, 2012.
20. South, M. Measurement of respiratory rate and timing using a nasal thermocouple. *J. Clin. Monit.*

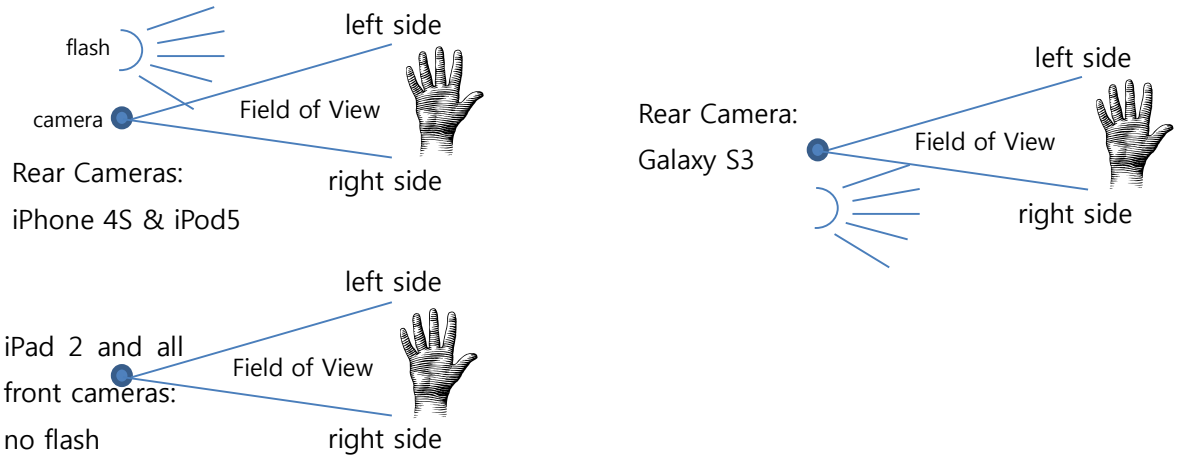
11:159–164, 1995.

21. Subbe, C., R. Davies, E. Williams, P. Rutherford, and L. Gemmell. Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions*. *Anaesthesia* 58:797–802, 2003.
22. Torrence, C., and G. P. Compo. A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* 79:61–78, 1998.
23. Wang, H., K. Siu, K. Ju, and K. H. Chon. A high resolution approach to estimating time-frequency spectra and their amplitudes. *Ann. Biomed. Eng.* 34:326–338, 2006.
24. Younes, M. Role of respiratory control mechanisms in the pathogenesis of obstructive sleep disorders. *J. Appl. Physiol.* 105:1389–1405, 2008.

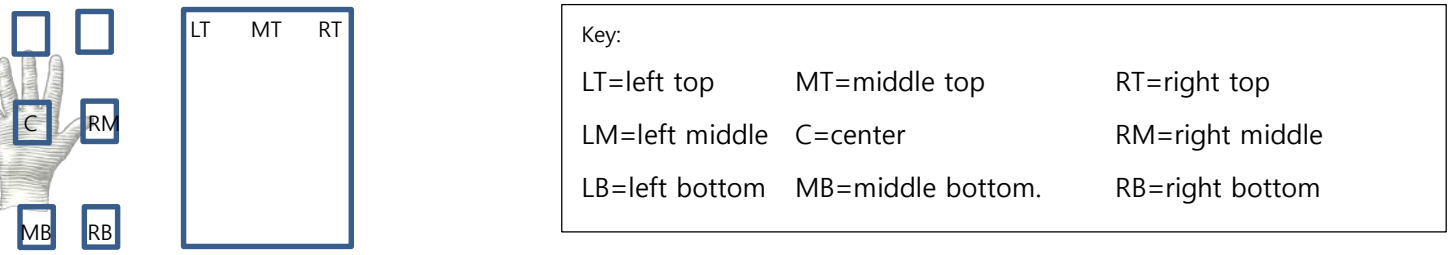


(a) iPhone 4S (Rear) (b) iPhone 4S (Front) (c) iPod 5 (Rear) (d) iPod 5 (Front) (e) Galaxy S3 (Rear) (f) Galaxy S3 (Front) (g) iPad 2 (Front)

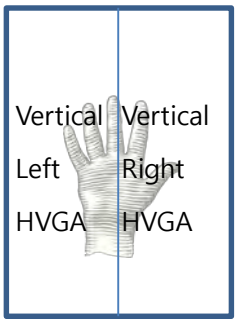
Fig. 1. General scheme to acquire video from the four devices



(a) Position of Flash Relative to Camera Field of View



(b) Selected 50x50 Pixel Regions (not to scale) within Camera's Field of View – All Devices Except iPad, which had Right on Top in Landscape



(c) Division of Camera's Field of View into Vertical Left Half VGA & Right Half VGA

Fig. 2. Example of different regions of iPhone 4S, iPad 2, iPod 5, and Galaxy S3. The top panel (Fig. 2a) represents the camera's Field of View (FOV) and relative position of flash LED's. The middle panel (Fig. 2b) shows the locations of the 9 50x50 pixel regions in the camera's FOV. The bottom panel (Fig. 2c) shows the division of the FOV into left and right vertical halves, each of HVGA resolution.

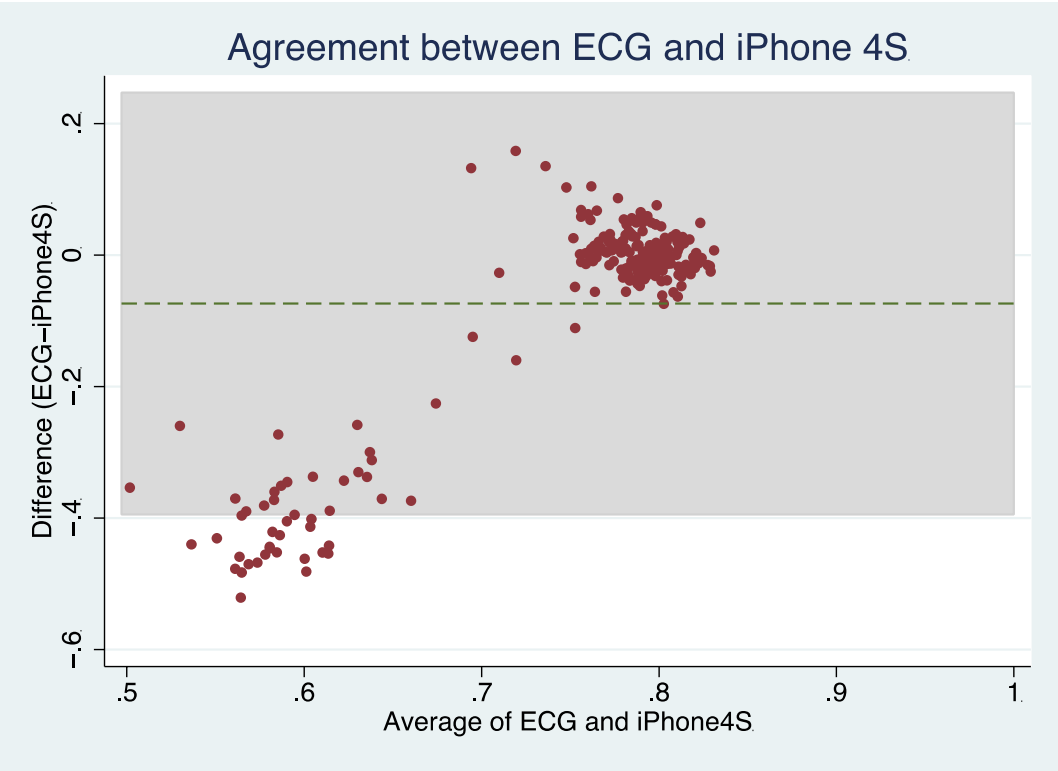
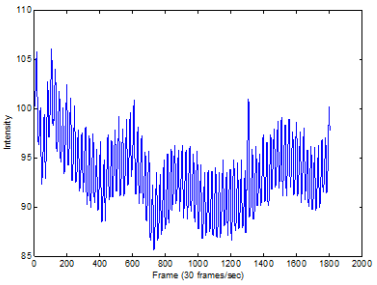
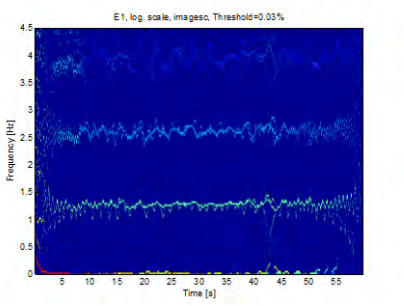


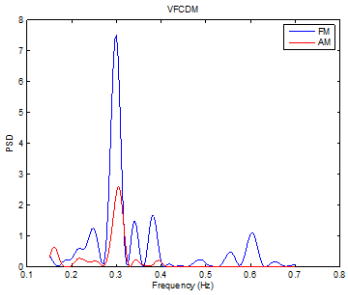
Fig. 3. Example Bland-Altman plot with a mean difference of 0.074 that shows the limit of agreement of 95% (dashed line is the mean difference ± the limit of agreement) between the continuous HR of a smart phone and the patient's corresponding ECG signal.



(a) Pulsatile Signal

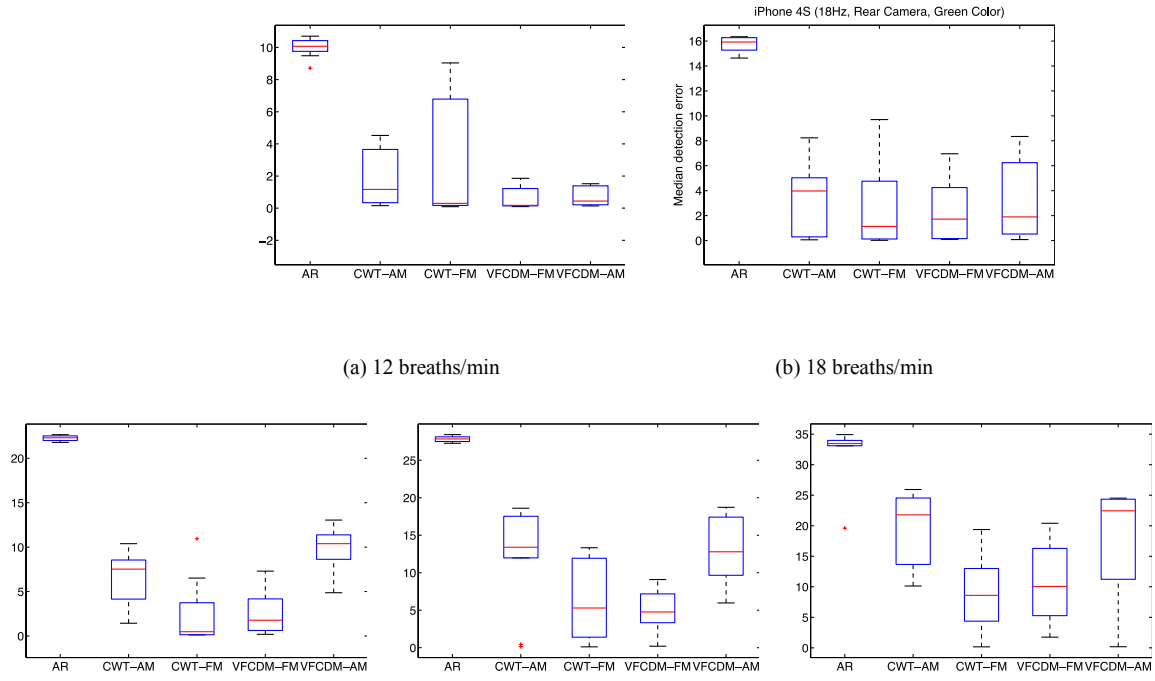


(b) Estimated instantaneous frequencies using VFCDM with prominent frequency oscillations seen near heart rate (1.3 Hz).



(c) PSD of PS signal

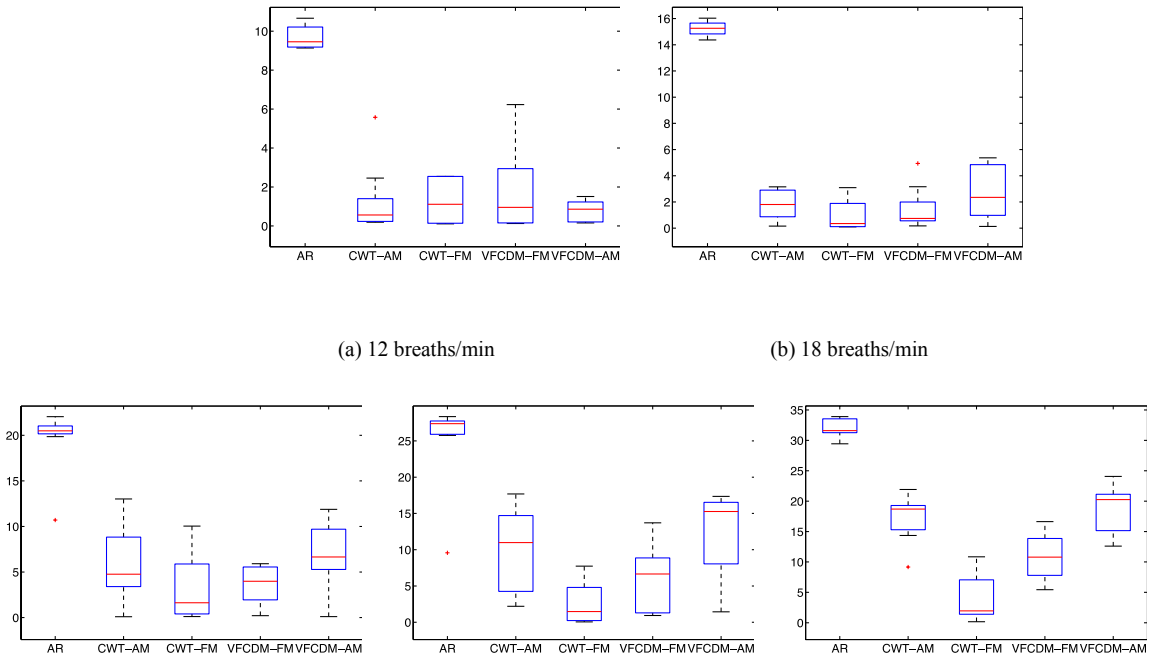
Fig. 4. PPG signal, Estimated instantaneous frequencies, PSD



(a) 12 breaths/min (b) 18 breaths/min

(c) 24 breaths/min (d) 30 breaths/min (e) 36 breaths/min

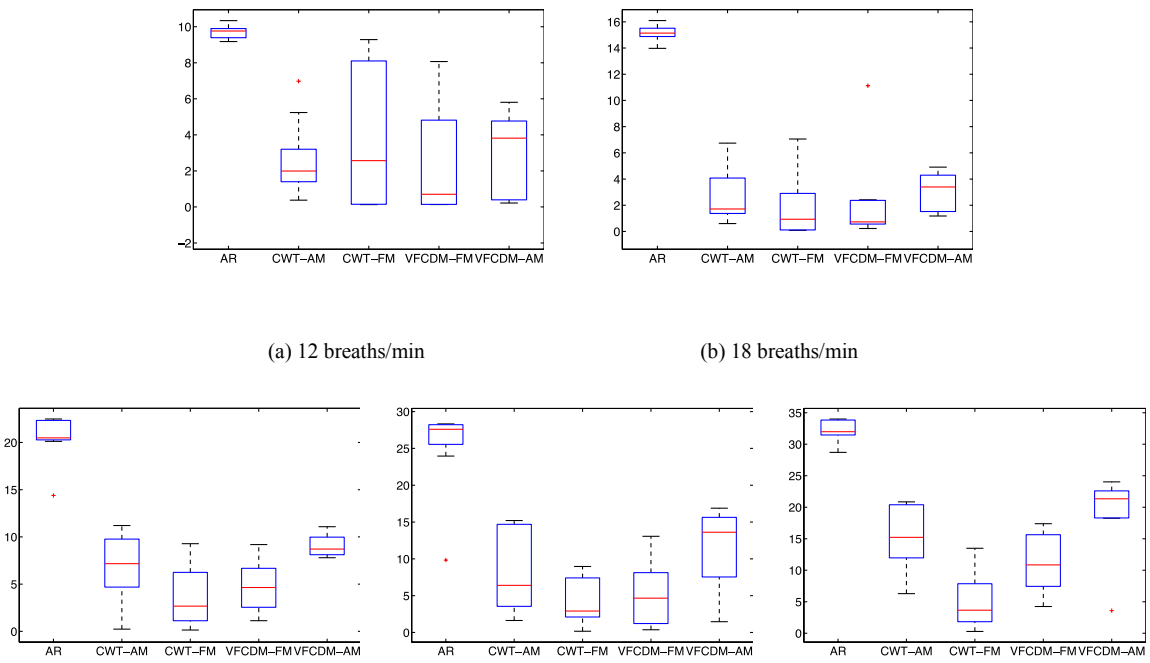
Fig. 5. Median and IQR errors measured from iPhone 4S when the flashlight was turned on (Resolution: HVGA). The top and bottom panels represent LH (12 and 18 breaths/min) and HF (24, 30 and 36 breaths/min) breathing rates, respectively.



(a) 12 breaths/min (b) 18 breaths/min

(c) 24 breaths/min (d) 30 breaths/min (e) 36 breaths/min

Fig. 6. Median and IQR errors measured from an iPhone 4S when the flashlight was turned off (Resolution: HVGA). The top and bottom panels represent LH and HF breathing rates, respectively.



(a) 12 breaths/min

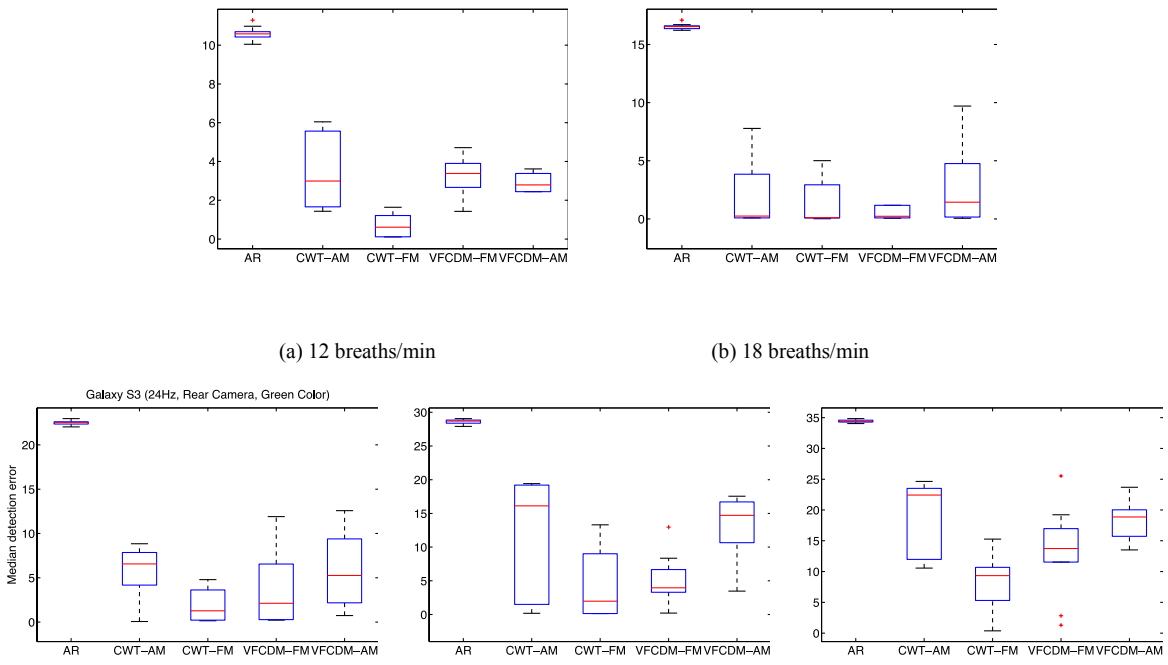
(b) 18 breaths/min

(c) 24 breaths/min

(d) 30 breaths/min

(e) 36 breaths/min

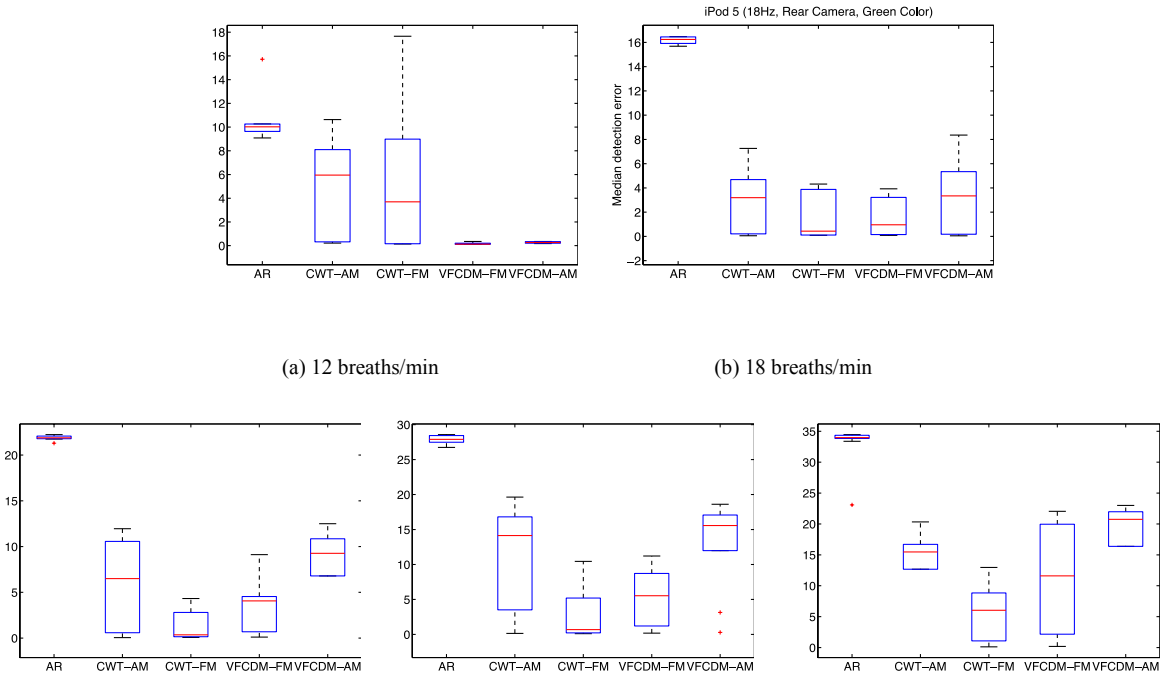
Fig. 7. Median and IQR errors measured from iPad 2 when the flashlight was turned off (Resolution: HVGA). The top and bottom panels represent LH and HF breathing rates, respectively.



(a) 12 breaths/min (b) 18 breaths/min

(c) 24 breaths/min (d) 30 breaths/min (e) 36 breaths/min

Fig. 8. Median and IQR errors measured from Galaxy S3 when the flashlight was turned on (Resolution: 50×50, region: LT). The top and bottom panels represent LH and HF breathing rates, respectively.



(a) 12 breaths/min

(b) 18 breaths/min

(c) 24 breaths/min

(d) 30 breaths/min

(e) 36 breaths/min

Fig. 9. Median and IQR errors measured from iPod 5 when the flashlight was turned on (Resolution: 50×50, region: LM). The top and bottom panels represent LH and HF breathing rates, respectively.

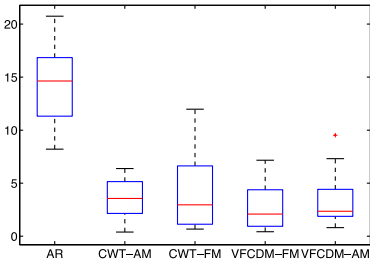


Fig. 10. Spontaneous respiratory rate

Table I Pulse amplitude values of each color band signal extracted from an iPhone 4S of 50×50 pixel regions consisting of: Center, Left Middle (LM), Left Bottom (LB), Right Middle (RM), Right Bottom (RB). *indicates $p<0.05$. # indicates $p<0.05$ for both blue and green vs. red.

Flash	Color	Mean amplitude value				
		Center	LM	LB	RM	RB
On	Blue	1.37±0.32	3.17±0.72	2.77±0.69	0.28±0.08	0.33±0.09
	Green	9.05±2.82*	10.49±3.26*	9.61±3.01*	7.02±2.19*	6.15±1.94*
	Red	1.71±0.56	0	0	4.71±1.5	4.69±1.49
Off	Blue	1.61±0.38	1.8±0.45	2.21±0.54	1.85±0.46#	2.16±0.53
	Green	3.01±0.89*	4.07±1.22*	3.43±1.02*	1.79±0.51#	0.76±0.21
	Red	0.21±0.06	0.19±0.05	1.04±0.32	0.24±0.06	4.08±1.26*

Table II Experimental results of heart rate extracted from ECG and three-color band signals obtained from iPhone 4S (Resolution: HVGA)

Color	PS	RRI	Median error
Blue	0.8124±0.23334	0.8103±0.0514	0.0021
Green	0.8149±0.19698		0.0047
Red	0.8121±0.22897		0.0018

Table III The mean amplitude values of the green color pulse signals with flash on except for iPad 2. *denotes p<0.05 to other pixel regions.

No	Resolution	Region	Mean amplitude value			
			iPhone 4S	iPad 2	iPod 5	Galaxy S3
1	50×50	RT	6.33±1.99	4.78±1.42*	2.67±0.82	9385.85±3140.96*
2		RM	7.02±2.19	4.77±1.42*	2.41±0.75	9326.86±3123.12*
3		RB	6.15±1.94	2.44±0.72	2.31±0.72	8583.78±2839.43*
4		MT	8.45±2.64	4.10±1.22	4.11±1.27	7066.07±2365.34
5		Center	9.05±2.82	3.88±1.16	2.79±0.88	6550.41±2173.4
6		MB	8.28±2.59	3.07±0.91	3.59±1.12	3459.99±1148.69
7		LT	9.42±2.94	3.53±1.06	5.79±1.79*	5682.13±1910.77
8		LM	10.49±3.26	2.89±0.85	6.23±1.92*	3969.18±1315.59
9		LB	9.61±3.01	4.05±1.21	5.04±1.57	1605.74±525.84
10	HVGA	Right	8.67±2.54	4.74±1.39*	3.53±1.02	7595.58±2521.62
11		Left	11.37±3.32*	3.78±1.11	5.17±1.49	2766.16±915.96
12	VGA	Full	9.05±2.65	3.11±0.91	5.75±1.66*	5168.72±1715.26

Table IV Experimental results of heart rate extracted from ECG and three-color band signals obtained from iPad 2 (Resolution: HVGA). * denote P<0.05 between pulsatile signals vs. RRI.

Color	Pulsatile Signals	RRI	Median error
Blue	0.5676±0.26*	0.7132±0.09	0.1456
Green	0.7124±0.04		0.0008
Red	0.7133±0.04		0.0001

Table V Experimental results of heart rate extracted from ECG (denoted RRI) and three-color band pulsatile signals obtained from iPod 5 (Resolution: 50×50). * denotes $p < 0.05$ between pulsatile signals vs. RRI.

Region	Flash	Color	Pulsatile Signal	RRI	Median error
LB	On	Blue	0.7684±0.1343	0.7522±0.0425	0.0162
		Green	0.749±0.0534		0.0032
	Off	Red	0.771±0.2755	0.7612±0.0842	0.0098
		Green	0.803±0.2874*		0.0418
		Blue	0.809±0.3571*		0.0478
LM	On	Blue	0.7832±0.2663	0.761±0.1303	0.0222
		Green	0.7675±0.0376		0.0065
	Off	Red	0.7858±0.3185	0.7942±0.1226	0.0084
		Green	0.8205±0.3291		0.0263
		Blue	0.8246±0.3756*		0.0304

Table VI Population statistics for IQR detection errors for each method. The error values listed for each method represent breaths/min.

Device	Breaths/min	AR	WT		CDM	
			AM	FM	FM	AM
iPhone 4S	12	1.06±0.55	1.52±0.8	3.65±1.89	3.17±1.65	1.04±0.53
	18	0.94±0.47	2.25±1.12	2.08±1.14	1.84±1.01	3.95±2
	24	1.28±0.65	6.12±3.12	5.86±3.08	3.76±1.89	5.24±2.67
	30	1.95±1.02	11.54±5.8	4.82±2.5	8.87±4.47	9.03±4.86
	36	2.48±1.32	4.57±2.43	6.38±3.46	7.02±3.51	7.44±3.94
iPad 2	12	0.59±0.3	2.69±1.38	7.96±4.08	5.18±2.84	4.58±2.39
	18	0.83±0.42	3.03±1.63	3.66±1.92	1.89±1.03	2.84±1.45
	24	2.15±1.17	5.94±2.98	6.25±3.22	4.4±2.2	2.01±1.02
	30	3.21±1.7	11.24±5.83	5.98±3.2	8.01±4.01	9.2±4.8
	36	2.45±1.28	8.93±4.48	6.95±3.54	9.15±4.6	4.34±2.23
Galaxy S3	12	0.42±0.22	1.26±0.64	2.1±1.05	1.68±0.92	1.09±0.55
	18	0.41±0.21	4.31±2.31	5.99±3	6.04±3.12	4.25±2.2
	24	0.15±0.08	7.96±4.03	6.48±3.24	5.28±2.66	5.79±2.89
	30	0.42±0.22	7.02±3.51	7.55±3.8	5.01±2.58	2.93±1.48
	36	0.69±0.35	9.94±4.97	14.07±7.13	7.79±3.99	7.93±4.18
iPod 5	12	8.64±4.72	7.48±3.79	4.42±2.21	3.29±1.66	5.49±3.05
	18	0.4±0.2	3.88±1.97	4.96±2.59	4.09±2.06	2.1±1.07
	24	0.67±0.34	4.54±2.27	2.38±1.29	4.57±2.4	7.87±4.05
	30	0.38±0.19	5.38±2.77	6.57±3.29	6.77±3.46	12.21±6.51
	36	0.9±0.45	9.34±4.68	19.38±9.97	11.19±5.74	11.8±5.96

Table VII Statistical significance (accuracy) among the five methods for four devices

Device	LF	HF	Device	LF	HF
IPhone 4S	AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM	AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM CDM-AM vs. CDM-FM CDM-AM vs. WT-FM CDM-FM vs. WT-AM WT-FM vs. WT-AM	IPod 5	AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM WT-AM vs. CDM-FM	AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM CDM-AM vs. CDM-FM CDM-AM vs. WT-FM WT-FM vs. WT-AM
iPad 2		AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM CDM-AM vs. CDM-FM WT-FM vs. CDM-AM WT-FM vs. WT-AM	Galaxy S3	AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM	AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM CDM-AM vs. CDM-FM CDM-AM vs. WT-FM WT-AM vs. WT-FM

Table VIII Statistical significance (repeatability across time) among the five methods for four devices

Device	LF	HF	Device	LF	HF
IPhone 4S	AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM	AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM CDM-AM vs. WT-FM	IPod 5	AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM	AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM CDM-AM vs. WT-FM
iPad 2		AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM WT-FM vs. CDM-AM	Galaxy S3		AR vs. CDM-AM AR vs. CDM-FM AR vs. WT-AM AR vs. WT-FM

Table IX Accuracy as determined by median errors at 42, 48, 54 breaths/min (iPhone 4S, flashlight: On). The error values listed for each method represent breaths/min.

Breaths/min	Error	AR	WT		CDM	
			AM	FM	FM	AM
42 (0.7 Hz)	Median	40.05±0.41	21.58±9.14	5.58±5.16	16.05±4.58	24.21±6.33
	IQR	0.72±0.38	19.89±10.15	7.22±3.87	9.27±4.72	5.17±2.59
48 (0.8 Hz)	Median	45.69±1.21	32.61±4.65	24.06±9.67	24.74±4.08	28.53±6.82
	IQR	0.68±0.35	9.3±4.97	14.07±7.04	4.61±2.32	6.25±3.15
54 (0.9 Hz)	Median	51.49±1.46	38.14±4.9	36.38±3.55	32.8±4.87	33.24±8.93
	IQR	0.41±0.22	6.79±3.68	6.93±3.51	6.05±3.07	11.77±6.28

Table X Statistical significance (accuracy and repeatability across time) among the five methods for spontaneous respiratory rate

Accuracy	Repeatability across Time
AR vs. CDM-AM	AR vs. CDM-AM
AR vs. CDM-FM	AR vs. CDM-FM
AR vs. WT-AM	AR vs. WT-AM
AR vs. WT-FM	AR vs. WT-FM

Table XI Computation time of heart rate extracted from color band signal of iPhone 4S depending on different resolutions

Resolution	Color	Computation time
320×240 (QVGA)	Green	25 frames/s
480×320 (HVGA)	Green	25 frames/s
480×320 (HVGA)	Green and Red	23 frames/s
480×320 (HVGA)	3 Colors	20 frames/s
640×480 (VGA)	Green or Red	19 frames/s

Multi-channel pulse oximetry for wearable physiological monitoring

Y. Mendelson, D. K. Dao, K. H. Chon
Department of Biomedical Engineering
Worcester Polytechnic Institute
Worcester, MA 01609
ym@wpi.edu

***Abstract*—Pulse oximetry is a widely accepted clinical method for noninvasive monitoring of arterial oxygen saturation and pulse rate. Significant improvements aimed at curbing motion artifacts and improving reliability in detecting sufficiently strong photoplethysmographic signals are required to reduce errant measurements before the pulse oximeter can be considered for wider mobile applications. The present work describes the development of a wearable multi-channel reflectance pulse oximeter to investigate if a motion artifact-free signal can be obtained in at least one of the multi-channels at any given time. Pilot findings provided a proof of concept to support the hypothesis that photoplethysmograms acquired concurrently from independent channels in a multi-channel pulse oximeter sensor respond differently to motion artifacts, thus laying the foundation for future development of robust active noise cancellation and data fusion based algorithms to mitigate the effects of motion artifacts.**

Keywords—Wearable sensors, pulse oximeter, motion artifacts

I. INTRODUCTION

Steady advances in noninvasive physiological sensing, hardware miniaturization and wireless communication are leading to the development of new wearable technologies that have broad and important implications for civilian and military applications. For example, the emerging development of compact, low-power, small-size, light- weight, and unobtrusive wearable devices may facilitate remote noninvasive monitoring of vital signs from soldiers during training exercises and combat. Telemetry of physiological information via a short-range wirelessly-linked personal area network can also be useful for particular categories of users, such as emergency first-responders, workers in harsh environments, including firemen and rescue patrols, or outdoors sportsmen, including high altitude mountaineers. The primary goals of such a wireless mobile platform would be to keep track of an injured person's vital signs, thus readily allowing the telemetry of physiological information to medical providers, and support emergency responders in making critical and often lifesaving decisions in order to expedite rescue operations. Having wearable physiological monitoring could offer far-forward medics numerous advantages, including the ability to determine a casualty's condition remotely without exposing the first responders to increased risks, quickly identifying the severity of injuries especially when the injured are greatly dispersed over large geographical terrains and often out-of-site, and continuously tracking the injured condition until they arrive safely at a medical care facility.

Several technical challenges must be overcome to address the unmet demand for long-term continuous physiological monitoring in the field. In order to design more compact sensors and improved wearable instrumentation, perhaps the most critical challenges are to develop more power efficient and low-weight devices. To become effective, these technologies must also be robust, comfortable to wear, and cost-effective. Additionally, before wearable devices can be used effectively in the field, they must become unobtrusive and should not hinder a person's mobility. Employing commercial off-the-shelf (COTS) solutions, for example finger pulse oximeters to monitor blood oxygenation and heart rate, or standard adhesive-type disposable electrodes for ECG monitoring, is not practical for many field applications because they limit mobility and can interfere with normal tasks. A potentially attractive approach to aid emergency medical teams in remote triage operations is the use of a wearable pulse oximeter to wirelessly transmit heart rate (HR) and arterial oxygen saturation (SpO_2) to a remote location.

Pulse oximetry is a widely accepted method that is clinically used for noninvasive monitoring of SpO_2 and HR. The method is based on spectrophotometric measurements of changes in the optical absorption properties of deoxyhemoglobin (Hb) and oxyhemoglobin (HbO_2). Noninvasive spectrophotometric measurements of SpO_2 are typically performed in the visible (600-700nm) and near-infrared (NIR) spectral regions between 800-950nm. Pulse oximetry relies on the detection of photoplethysmographic (PPG) signals produced by variations in the quantity of arterial blood that is associated with periodic contractions and relaxations of the heart. Hence, the technique relies on the presence of a stable peripheral arterial pulse.

Pulse oximetry can be performed in either transmission or reflection modes. In transmission pulse oximetry, the sensor is typically attached across a fingertip, foot, or earlobe. In this configuration, the light emitting diodes (LEDs) and photodetector (PD) are mounted on opposite sides of a peripheral pulsating vascular bed. Alternatively, in reflection-mode pulse oximetry, the LEDs and PD are both mounted side-by-side on the same planar substrate to enable readings from multiple body locations where trans-illumination measurements are not feasible. Clinically, reflectance pulse oximetry has long been recognized as a potential alternative method to transmission pulse oximetry in certain medical applications where peripheral perfusion might be compromised. Additionally, reflection-mode is attractive for body sensor networks (BSN) due to the flexibility in choosing various sensor mounting locations over conventional transmission-mode pulse oximetry.

978-1-4799-0330-6/13/$31.00 ©2013 IEEE

Several studies have reported that forehead oximeters are at least as accurate as finger mounted oximeters under normal testing conditions, and due to their central placement, are affected less by thermoregulatory vasoconstriction and are able to respond more quickly to desaturation events [1, 2]. Also, during conditions which lead to poor peripheral perfusion, forehead sensors have demonstrated greater accuracy than finger sensors [3, 4]. In addition, pulse oximetry measurements from the forehead offer a potential advantage in tactical settings that require extensive use of the hands that can introduce excessive motion artifacts. While reflectance mode pulse oximetry remains promising, significant improvements aimed at curbing motion artifact and improving reliability in detecting sufficiently strong PPG signals are required to identify and reduce errant measurements before they can be considered for wider and more reliable mobile applications.

II. Motion Artifacts

Although well accepted for use in resting subjects, using pulse oximetry outside of a more controlled hospital setting has been problematic for several reasons. Depending on the measurement site, sensors may be subjected to varying degrees of motion artifacts, resulting in signal corruption and thus inaccurate estimations of HR and SpO_2 [5, 6]. Many clinicians have cited motion artifacts in pulse oximetry as the most common cause of false alarms, loss of signal, and inaccurate readings [7]. While the intelligent design of sensor attachment, form factor and packaging can help to reduce the impact of motion disturbances by making sure that the sensor is securely mounted, it is rarely sufficient for noise removal.

In relation to pulse oximetry obtained from the forehead, it is speculated that the main source of motion artifact is due to changes in the relative position of the sensor with respect to the curved skull rather than the relative movements of the sensor with respect to the skin. Due to the rounded and optically inhomogeneous surface properties of the forehead, alterations in sensor position and orientation will cause changes in the distribution of backscattered light reaching the PD. Therefore, sudden changes in incident light intensity reaching the PD due to cyclical movement of the sensor will result in the corruption of the PPG signals. Some research has also suggested that there may be two other sources of motion artifacts. The first source of motion artifacts can be attributed to the formation of air gaps created between the skin and sensor during physical activity [8], which may cause measurement error. Another source of motion artifact can be attributed to low venous pressure blood "slosh" with back and forth movement which is seen when an individual is physically active. This local perturbation of venous blood adds to the AC component of the PPG signal and can result in low SpO_2 measurements [9].

Combating motion artifacts can be performed via both hardware and computational implementations:

i. Computational Approaches to Combat Motion Artifacts: Various computational algorithms attempt to isolate the effects of undesired motion-induced artifacts by rejecting suspect estimates of signal values [10]. Making matters worse in this case is that the noise can frequently fall within the same in-band frequency as the physiological signal of interest, thus rendering conventional linear signal filtering with fixed cut-off frequencies ineffective. Recently developed pulse oximeters offer potential advantages because they utilize advanced signal-processing methodologies in an attempt to provide continuous and accurate measurements when signals are weak (e.g., low perfusion) or corrupted by motion artifacts. Among the numerous signal processing techniques explored to address the confounding issue of in-band noise is adaptive noise cancellation (ANC). One example of a motion-tolerant algorithm is the Signal Extraction Technology (SET®) developed by Masimo [11].

ii. Hardware Approaches to Combat Motion Artifacts:

Since the introduction of pulse oximetry in the 1980s, improvements have been made to decrease the interference of motion artifacts on continuous, reliable estimation of oxygen saturation. New adhesive materials and mechanical design of the sensor housing placed against the skin have dramatically reduced problems with adherence and almost eliminated skin complications from sensor heat or reaction to adhesive materials. Improvements in sensor technology, particularly those related to minimizing motion artifacts, have progressively improved the accuracy and reliability of the devices during the past 20 years.

As PPG signals are highly susceptible to motion, various strategies have been employed to improve estimates of physiological variables derived from noisy PPG signals. Generally, motion artifacts in the recorded PPG signals are more difficult to remove than instrumental artifact as they do not have a predetermined narrow frequency band and their spectrum often overlaps with the desired signal. Thus, classical linear filtering with fixed cut-off frequencies to minimize the effect of motion artifacts cannot be implemented very effectively. Accelerometers (ACC) combined with ANC have been suggested as a promising approach for active noise cancellation of motion-corrupted biosignals [12, 13]. The most common approach employs an accelerometer sensor based on MEMS technology which offers a low-cost solution [14-16]. For example, Relente et al. [17] used an accelerometer as a motion reference for removing artifacts from a Nellcor pulse oximeter. However, despite these promising results, the effectiveness of an accelerometer-based automatic noise cancellation depends on the type of motion artifacts. For example, the reduction in noise may be limited during less repetitive sporadic movements. Moreover, if the motion frequency shifts rapidly over a wide spectral band, the approach is generally less effective due to a slower adaptation rate.

III. Prototype Sensor Configurations to Study the Effects of Motion Artifacts

Our laboratory has developed several prototype wearable reflectance-type pulse oximeters to investigate the effects of motion artifacts on different sensor configurations.

A. Dual-Wavelength and Single PD Configuration

Fig. 1 depicts a more conventional custom optical sensor configuration comprised of a pair of red (R) and NIR LEDs and a single PD. The wearable sensor contains an optical reflectance module, electronic circuitry and a tri-axial accelerometer. The PPG waveforms are acquired using a small

Fig. 1. Dual-wavelength forehead wearable pulse oximeter.

Silicon photodetector. The built-in accelerometer provides estimates of the wearer's posture and mobility. The sensor is housed in a rigid enclosure that is contoured to an average size adult head. The 30 mm x 70 mm x 15 mm sensor assembly is held in place by a compressive headband. The sensor is powered by a non-rechargeable Lithium-ion battery, providing approximately 100 hours of continuous operation. Data acquired by the wearable sensor are transmitted wirelessly to a USB-based receiver via low-power, peer-to-peer wireless communication over a short-range RF link using the 902-928 MHz ISM band.

Excessive contact pressure between a reflectance sensor and the skin is known to interfere with local blood flow, consequently leading to diminished or loss of the PPG signals. This interference can subsequently affect measurement accuracy. Several studies were conducted to aid in understanding how contact pressure affects pulse oximetry measurements. These studies provide a qualitative description of the effect of contact pressures on the PPG signal and its components [18, 19]. When the contact pressure used to secure the sensor to the body is too low, low amplitude PPG waveforms result in inaccurate measurements. On the contrary, if the contact pressure is too high, blood circulation can be compromised or necrosis could occur, leading to a complete loss of the PPG signal and the ability to obtain SpO_2 and HR measurements if the sensor is worn for extended periods of physical activity.

B. Multi-Channel LED and PD Configurations

Experience has shown that considerable variations in sensor position and tissue heterogeneities could cause large measurement errors. In addition, most of the light emitted from the LEDs is diffused by the underlying subcutaneous tissues predominantly in the forward direction (i.e. perpendicular to the emitting surface of the LEDs). Therefore, only a relatively small fraction of the light is diffused in a lateral direction. This suggests that to capture a larger proportion of the diffused backscattered light, the PD must be able to detect light from an area concentric with respect to the location of the LEDs. To minimize the dependency of backscattered light on local tissue inhomogeneity, a custom sensor has been designed based on a radially-symmetric arrangement of three pairs of identical R and IR LEDs surrounding a spectrally matched PD as depicted in Fig. 2.

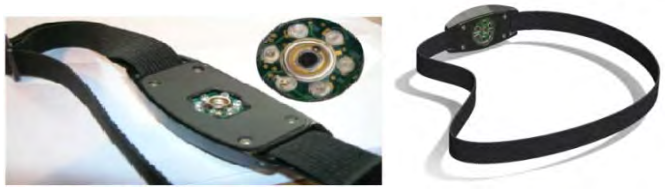


Fig. 2. Multi-channel wearable sensor configuration.

Our goal in designing these sensors was to create a multi-channel pulse oximeter (MCPO) that can be used to investigate how SpO_2 and HR readings may be affected by motion artifacts. The sensor design strategy is analogous to data fusion utilized for example in multi-channel electroencephalographic (EEG) analysis where noise may affect some of the channels but not all channels to the same degree at any given time. The rationale was based on the hypothesis that multiple channels will allow redundancy of data and will likely improve the confidence in making more robust decisions due to the use of complementary information, thus increasing the likelihood of maintaining the accuracy of physiological measurements even during the adverse induction of motion artifacts. Moreover, because of the relative differences in the spatial locations of the LEDs and PDs, we reasoned that local changes in sensor orientation can lead to perturbations in the coupling between the optical components and the skin during movement. These perturbations may have a different effect on the morphological similarities between correlated PPG signals acquired simultaneously by some channels compared to other channels at any given time.

We have developed two different MCPO configurations. The first sensor configuration is based on a single PD and 6 concentrically arranged LEDs. In this embodiment, the PD is used to acquire independently six PPG signals. Digital switching circuits are used to activate each LED in succession and synchronize signal detection using a time-multiplexing approach similar to the operation of a conventional dual-wavelength pulse oximeter, but expanded to include more channels. The second sensor configuration is comprised of six identical PDs arranged symmetrically in a radial configuration surrounding a pair of closely-spaced R and IR LEDs.

IV. Experimental Protocol

To analyze the ability of the MCPO sensor to reject motion artifact, different types of simple physical activities to introduce typical disturbances were tested. All tests were approved by our IRB. Data were first collected in a comfortable laboratory setting from 5 volunteers in an upright sitting position. In this setting, subjects were asked to rest for 5 minutes while PPG data were recorded from the forehead mounted sensor to establish motionless baseline readings. This was followed by a sequence of short activities that included slow left-right (L/R), up-down (U/D) and circular head movements to induce mild motion artifacts. In a second study we recorded PPG signals while the subjects wore the forehead sensor and walked straight, up/down a set of stairs and turned around in a circular pattern to simulate typical daily activities.

V. Data Analysis

Each digitized PPG signal was separated off-line into time-invariant (DC) and time-variant (AC) components using infinite impulse response (IIR) digital filters. Instantaneous heart rates were determined by computing the time interval between two successive peaks in the AC component of the IR PPG. A 5-point moving average was applied to account for variability in the instantaneous HR readings. SpO_2 were computed from the R/IR ratios using an empirical calibration relationship. Fig. 3 represents two typical examples of PPG signals recorded from different IR channels by the single PD

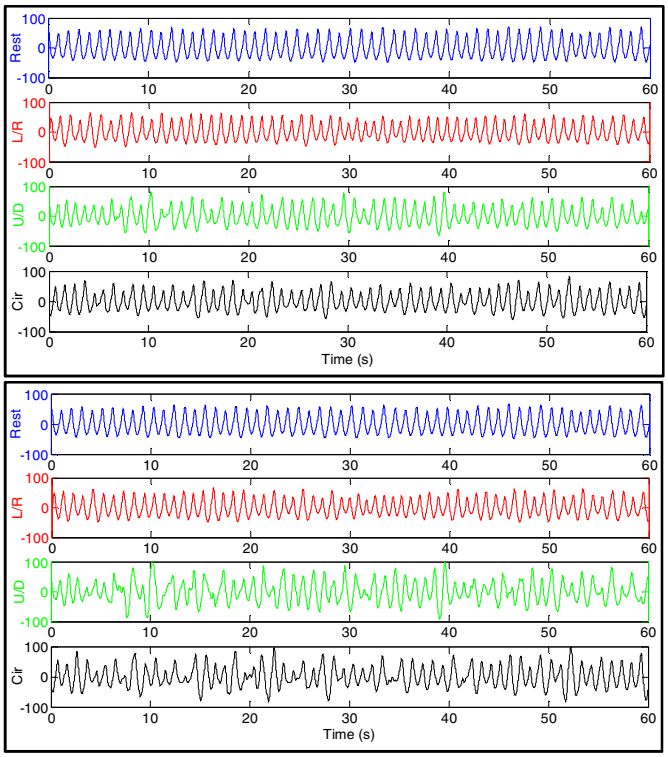


Fig. 3. Typical IR PPG signals recorded simultaneously from two different channels (top and bottom 4 traces) during rest, left-right (L/R), up-down (U/D) and circular head movements.

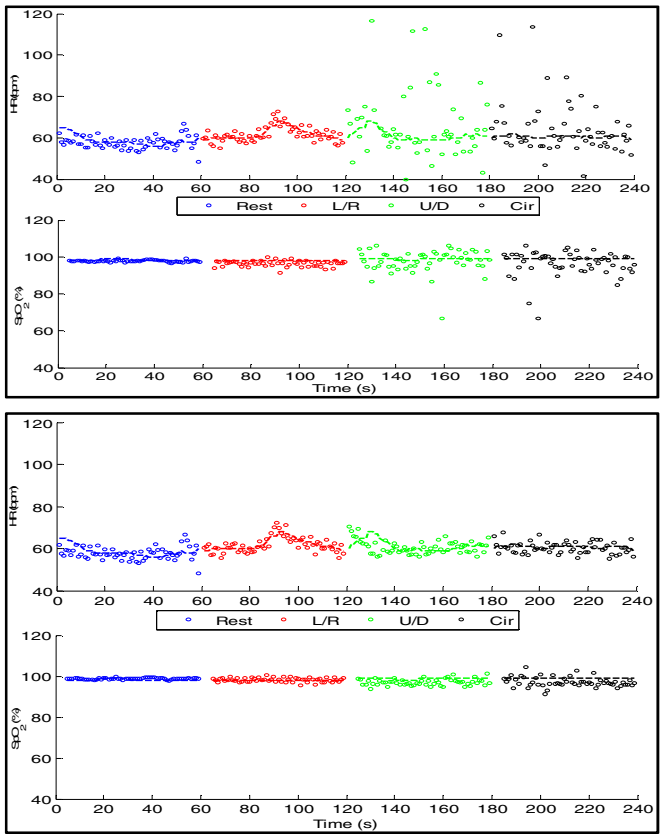


Fig. 4. SpO_2 and HR estimations derived from two different PPG channels recorded simultaneously during rest and motion induced activities. Horizontal traces denote average readings obtained by the reference pulse oximeter.

positioned in the center of the MCPO sensor. Similarly, Fig. 4 shows corresponding SpO_2 and HR estimations derived from two different PPG channels recorded simultaneously during rest and motion induced artifacts. Notice the overall changes in signal amplitude and morphology in the recorded PPG waveforms caused by typical left-right, up-down and circular head movements while the subject remained in a sitting position.

Fig. 5 summarizes the mean and SD corresponding to HR and SpO_2 derived readings obtained from every PPG channel in the MCPO prototype sensor. These data were recorded during rest and voluntary left-right, up-down and circular head movements while the subject remained in a sitting position. Horizontal lines represent mean HR and SpO_2 measurements obtained concurrently by a reference Masimo Radical SET™ pulse oximeter sensor mounted on the subject's finger while the hand was immobilized to limit motion artifacts.

The response of the MCPO to motion artifacts was also evaluated under more representative activities by recording PPG data from the forehead mounted sensor and Masimo finger pulse oximeter while the subject was walking casually, climbing a set of stairs and performing short turning manuvers. Fig. 6 summarizes the mean and SD corresponding to the HR and SpO_2 derived readings obtained during these activities.

Tables I and II compare average HR and SpO_2 measurements derived from different PPG channels in the MCPO sensor during voluntary head movements, while the subject was sitting and performing controlled head movements in the laboratory setting, with measurements obtained during free less restricted body movements outside the laboratory. These data clearly show that calculated HR values derived independently from certain PPG channels are within acceptable errors of ±1 bpm, while other channels produced clinically significant errors.

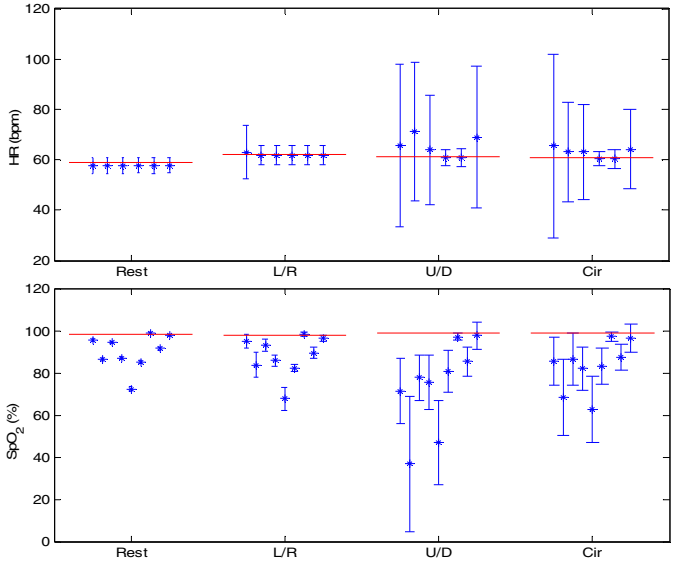


Fig. 5. HR and SpO_2 obtained from 6 independent PPG channels during rest, L/R, U/D and circular head movements. (*Top*) HR derived from each channel. (*Bottom*) corresponding SpO_2 readings derived from 9 R/IR channel pairs. Horizontal lines denote mean measurements obtained concurrently from a finger by the reference pulse oximeter.

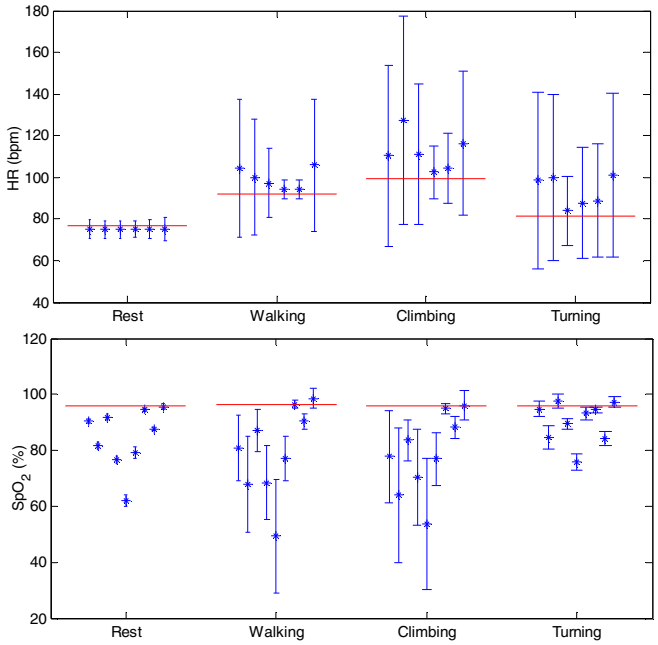


Fig. 6. HR and SpO_2 obtained from 6 independent PPG channels during sitting, walking straight, climbing stairs, and turning to simulate movement artifacts. (*Top*) HR derived from each channel. (*Bottom*) corresponding SpO_2 readings derived from 9 R/IR channel pairs. Horizontal lines represent mean measurements obtained concurrently from a finger by a reference Masimo pulse oximeter.

TABLE I. MEAN HR DIFFERENCES DERIVED FROM DIFFERENT PPG CHANNELS DURING VOLUNTARY HEAD MOVEMENTS (TOP) AND MEASUREMENTS OBTAINED DURING FREE MOVING EXERCISES (BOTTOM).

		LED 1	LED 2	LED 3	LED 4	LED 5	LED 6
Head	***Rest***	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
	L/R	1.0	-0.2	-0.1	-0.2	-0.2	-0.2
	U/D	4.2	9.7	2.5	-0.6	-0.5	7.5
	Circular	4.7	2.3	2.3	-0.4	-0.4	3.4
Body	***Rest***	-1.4	-1.4	-1.4	-1.4	-1.4	-1.3
	Walking	12.4	8.2	5.4	2.2	2.2	13.9
	Climbing	10.9	27.8	11.8	3.0	4.9	17.0
	Turning	16.9	18.2	2.4	6.2	7.1	19.5

TABLE II. MEAN SpO_2 DIFFERENCES DERIVED FROM DIFFERENT COMBINATIONS OF R/IR CHANNEL PAIRS DURING VOLUNTARY HEAD MOVEMENTS (TOP) AND FREE MOVING EXERCISES (BOTTOM).

R/IR pair		1	2	3	4	5	6	7	8	9
Head	***Rest***	-2.5	-11.7	-3.7	-11.1	-25.9	-13.0	0.6	-6.6	-0.4
	L/R	-3.1	-14.3	-5.1	-12.2	-30.3	-15.7	0.2	-8.6	-1.5
	U/D	-27.6	-62.1	-21.1	-23.4	-51.9	-18.2	-1.8	-13.6	-1.2
	Circular	-13.6	-30.7	-12.5	-17.0	-36.5	-15.8	-1.8	-11.6	-2.6
Body	***Rest***	-5.5	-14.2	-4.1	-19.2	-34.1	-16.8	-1.4	-8.3	-0.3
	Walking	-15.3	-28.4	-9.1	-27.9	-46.9	-19.3	0.2	-5.9	2.3
	Climbing	-18.3	-32.1	-12.5	-25.8	-42.4	-19.1	-1.2	-7.8	0.0
	Turning	-1.2	-11.2	1.6	-6.4	-20.0	-2.7	-1.4	-11.8	1.3

VI. CONCLUSIONS

The present work described the development of a MCPO that can be used to investigate how SpO_2 and HR readings may be affected by motion artifacts. These pilot findings showed evidence to support the hypothesis that PPG signals acquired concurrently from independent channels in a wearable reflectance-type MCPO sensor are affected differently by motion artifacts, allowing for automatic adjudication of which signal is likely to be a more accurate reflection of physiological changes, thus helping to reduce measurement errors. Future work will be focused on the development of advanced active noise cancellation algorithms to take advantage of the spatial diversity of different channels and fuse the data measured by the most reliable channels in the MCPO. If proven successful, this strategy will be used to improve real-time measurements of SpO_2 and HR by a wearable reflectance-mode pulse oximeter.

ACKNOWLEDGMENTS

The authors acknowledge the financial support provided by the MIT-Lincoln Laboratory. This work is supported by the US Army Medical Research and Materiel Command (USAMRMC) under Grants No. W81XWH-10-1-0529 and W81XWH-12-1-0541. The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REFERENCES

[1] S. Sugino, N. Kanaya, M. Mizuuchi, M. Nakayama, and A. Namiki, "Forehead is as sensitive as finger pulse oximetry during general anesthesia," *Canadian Journal of Anaesthesia-Journal Canadien D Anesthesie*, vol. 51, pp. 432-436, May 2004.

[2] S. J. Choi, H. J. Ahn, M. K. Yang, C. S. Kim, W. S. Sim, J. A. Kim, *et al.*, "Comparison of desaturation and resaturation response times between transmission and reflectance pulse oximeters," *Acta Anaesthesiologica Scandinavica*, vol. 54, pp. 212-217, Feb 2010.

[3] R. D. Branson and P. D. Mannheimer, "Forehead oximetry in critically ill patients: the case for a new monitoring site," *Respir Care Clin N Am*, vol. 10, pp. 359-67, vi-vii, Sep 2004.

[4] L. Schallom, C. Sona, M. McSweeney, and J. Mazuski, "Comparison of forehead and digit oximetry in surgical/trauma patients at risk for decreased peripheral perfusion," *Heart Lung*, vol. 36, pp. 188-94, May-Jun 2007.

[5] L. H. Norton, B. Squires, N. P. Craig, G. McLeay, P. McGrath, and K. I. Norton, "Accuracy of pulse oximetry during exercise stress testing," *Int J Sports Med*, vol. 13, pp. 523-7, Oct 1992.

[6] H. Benoit, F. Costes, L. Feasson, J. R. Lacour, F. Roche, C. Denis, *et al.*, "Accuracy of pulse oximetry during intense exercise under severe hypoxic conditions," *Eur J Appl Physiol Occup Physiol*, vol. 76, pp. 260-3, 1997.

[7] M. T. Petterson, V. L. Begnoche, and J. M. Graybeal, "The effect of motion on pulse oximetry and its clinical significance," *Anesth Analg*, vol. 105, pp. S78-84, Dec 2007.

[8] H. H. Asada, P. Shaltis, A. Reisner, S. Rhee, and R. C. Hutchinson, "Mobile monitoring with wearable photoplethysmographic biosensors," *Ieee Engineering in Medicine and Biology Magazine*, vol. 22, pp. 28-40, May-Jun 2003.

[9] A. Sola, L. Chow, and M. Rogido, "Pulse oximetry in neonatal care in 2005. A comprehensive state of the art review," *An Pediatr (Barc)*, vol. 62, pp. 266-81, Mar 2005.

[10] N. Selvaraj, Y. Mendelson, K. Shelley, D. Silverman, and K. Chon, "A computational approach for the detection and rejection of motion/noise artifacts in PPG," *IEEE Trans Biomed Eng*, 2011.

[11] J. M. Goldman, M. T. Petterson, R. J. Kopotic, and S. J. Barker, "Masimo signal extraction pulse oximetry," *J Clin Monit Comput*, vol. 16, pp. 475-83, 2000.

[12] L. B. Wood and H. H. Asada, "Noise cancellation model validation for reduced motion artifact wearable PPG sensors using MEMS

accelerometers," *Conf Proc IEEE Eng Med Biol Soc*, vol. 1, pp. 3525-8, 2006.

[13] J. Y. 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, pp. 140-5, Mar 2006.

[14] M. J. Mathie, A. C. Coster, N. H. Lovell, and B. G. Celler, "Detection of daily physical activities using a triaxial accelerometer," *Med Biol Eng Comput*, vol. 41, pp. 296-301, May 2003.

[15] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Trans Inf Technol Biomed*, vol. 10, pp. 156-67, Jan 2006.

[16] Y. Mendelson, R. J. Duckworth, and G. Comtois, "A wearable reflectance pulse oximeter for remote physiological monitoring," *Conf Proc IEEE Eng Med Biol Soc*, vol. 1, pp. 912-5, 2006.

[17] A. R. Relente and L. G. Sison, "Characterization and adaptive filtering of motion artifacts in pulse oximetry using accelerometers," *Second Joint Embs-Bmes Conference 2002, Vols 1-3, Conference Proceedings*, pp. 1769-1770, 2002.

[18] A. C. M. Dassel, R. Graaff, M. Sikkema, A. Meijer, W. G. Zijlstra, and J. G. Aarnoudse, "Reflectance Pulse Oximetry at the Forehead Improves by Pressure on the Probe," *Journal of Clinical Monitoring*, vol. 11, pp. 237-244, Jul 1995.

[19] X. F. Teng and Y. T. Zhang, "The effect of contacting force on photoplethysmographic signals," *Physiological Measurement*, vol. 25, pp. 1323-1335, Oct 2004.

Arrhythmia Discrimination using a Smart Phone

Jo Woon Chong[1], David D. McManus[2], and Ki H. Chon[1]
[1]Department of Biomedical Engineering, Worcester Polytechnic Institute, Worcester, MA, USA;
[2]Department of Medicine, University of Massachusetts, USA;

***Abstract*—We propose an arrhythmia discrimination algorithm for a smart phone that can reliably distinguish among normal sinus rhythm (NSR), atrial fibrillation (AF), premature ventricular contractions (PVCs) and premature atrial contraction (PACs). To evaluate the algorithm in clinical application, we recruited 27 subjects with 3 PVC and 4 PAC subjects as well as 20 AF pre- and post- electrical cardioversion. From each subjects, two-minute pulsatile time series from a fingertip is measured using a smart phone. Our arrhythmia discrimination approach combines Poincare plot and Kulback-Leibler (KL) divergence with Root Mean Square of Successive RR Differences (RMSSD) and Shannon Entropy (ShE). Clinical results show that our algorithm discriminates PVC and PAC with accuracy of 100% and 97.87%, respectively.**

Keywords—atrial fibrillation; Kullback-Leibler divergence; Poincare plot; premature ventricular contraction; premature atrial contraction; Shannon entropy; turning point ratio

I. INTRODUCTION

Atrial Fibrillation (AF) is the most prevalent arrhythmia worldwide and is increasing with the aging of population. Since AF is associated with heart failure, hospitalization and mortality, AF detection as well as AF treatment is in high demand for the longevity of lives. Since AF can be paroxysmal and asymptomatic in its nature [1], the major challenge is to detect paroxysmal and asymptomatic AF in an efficient way. Currently, undiagnosed AF population is reported to be considerable [2] and frequent monitoring is shown to improve AF detection [3]. Hence, an accurate AF detection method, which is readily available, is highly demanded to improve longevity of lives and reduce healthcare cost. However, current AF algorithms misclassifies non-AF into AF. For example, normal sinus rhythm (NSR) with frequent premature ventricular contraction (PVC) or premature atrial contraction (PAC) episodes can lead to misclassification as AF. This is because the PAC and PVC episodes with NSR mimic the AF dynamics and current AF algorithms are unable to discriminate the dynamics of NSR with PVCs/PACs from that of AF.

For PVC/PAC detection algorithms from ECG signals, template matching is widely used [4]. However, it needs data memory to store templates and requires high computational complexity for template matching between input signal and templates. Due to these limitations, this method is not applicable to real-time computing devices. Our previous AF discrimination algorithm misclassifies NSR with frequent PVC and PAC episodes to be AF. Hence, a new algorithm discriminating PVC and PAC from NSR and AF is needed.

In this paper, we developed an automated algorithm discriminating among NSR, AF, PVCs, and PACs. The developed algorithm is based on a smartphone without additional hardware. Digital camera and flash light embedded in a smart phone monitors skin optically and enables sensing the variability in heart rate signal as shown in Fig. 1 [5]. A 60-beat segment from smart phone is used as an input for our arrhythmia algorithm to discriminate among NSR, AF, PVC, or PAC rhythms. Our previous algorithm is based on statistical metrics of root mean square of successive RR differences (RMSSD) and Shannon entropy (ShE) [6]. We combine Poincare plot and Kullback-Leibler (KL) divergence with RMSSD and ShE to additionally detect bigeminy, trigeminy, quadrigeminy associated with PVCs and PACs as well as to improve accuracy of AF detection. Poincare plot is applied to detect specific patterns such as bigeminy, trigeminy, quadrigeminy associated with PVC or PAC while KL is for discriminating PVCs from PACs. We measured pulsatile time series of 20 NSR, 20 AF, 3 PVC, and 4 PAC subjects using digital camera and flash in an iPhone 4S.

II. METHODS

A. AF, PVC, and PAC Databases and Clinical Data Collection

The 20 NSR, 20 AF, 3 PVC, and 4 PAC subjects are recruited by the UMASS Medical Center (UMMC). Pulsatile time series data are collected using an iPhone 4S and the data collection protocol was approved by the Institutional Review Boards of Worcester Polytechnic Institute (WPI) and UMMC, respectively. The subjects are instructed to place their first (index) or third (middle) finger on the camera. Our data collection program automatically turns on flash when we start measurement. During two minutes of measurement, the subjects are instructed to breathe spontaneously in the supine position.

A current prototype of NSR, AF, PVC, and PAC discrimination application for iPhone 4S is shown in Fig 1. A patient can monitor iPhone's measurement procedure on a screen showing blood flow intensity amplitude, heart rate and remaining progress time in real-time. After two minutes of measurement, the application displays heart rhythm identification as well as average hear rate on the screen.

B. Preprocessing

We record videos of the human fingertip to measure blood flow intensity. From the video, we made use of the green band among the RGB band. This is due to our recent study that the green band among RGB band shows the best signal fidelity.

978-1-4799-0330-6/13/$31.00 ©2013 IEEE

Fig. 1. A smart phone application for data recording (the application uses the camera lens and illumination to acquire information about heart rate and rhythm).

The sampling rate is 30 frames per second, and the resolution of one frame is 640x480 pixels. Since our experiment shows that the upper half of the video signal (320x480 pixels) has the best signal fidelity, we average the intensity value of the upper half of the frame.

After obtaining the intensity value from each frame, pulse-to-pulse detection is performed by incorporating interpolation, sudden DC change elimination, two stages of band pass filter, derivative rank filter and matching of original peaks.

C. AF, PVC, and PAC Discrimination

The proposed arrhythmia discrimination algorithm takes a 64 pulse beat series as an input in detecting and determining arrhythmia of AF, PVC, and PAC. Combined with our previous AF detection algorithm [7], the proposed arrhythmia discrimination algorithm discriminates various patterns of PVCs and PACs.

The algorithm first derives RMSSD and Shannon entropy from the pulsatile time series and then compares those statistics with their corresponding threshold values $RMSSD_{\text{th}}$ and ShE_{th}. If the derived RMSSD and ShE are less than their thresholds, then the subject is determined to have NSR without PAC or PVC. On the other hand, if at least one of the statistics is larger than their threshold values, then the algorithm constructs a Poincare plot. Poincare plot discriminates specific patterns, e.g,. bigeminy, trigeminy, and quadrigeminy patterns of PVCs/PACs from the NSR and AF. After finding major patterns of PVCs and PACs, the algorithm updates the pulsatile time series subtract them from the original time series and derive RMSSD' and ShE' from the updated pulsatile time series. If at least one of the derived RMSSD and ShE are larger than their corresponding threshold values, then the subject is discriminated to have AF. Otherwise, the subjects are classified to have PVC or PAC.

The KL determines whether a subject has PVC or PAC. The algorithm makes use of two KL divergences values, obtained from the measured pulses of the unclassified subject as well as trained PVC and PAC data, in its determination. The detailed procedure of the Poincare plot and KL methods are explained in the following section.

1) Poincare Plot

A Poincare plot is used to visualize and quantify the self-similarity of a time series x_n for $n = 1, 2, ..., N$. For a two-dimensional Poincare plot, (x_{n-1}, x_n) is plotted on a two-dimensional Euclidean space. The Poincare plot approach is appropriate in determining the specific patterns of PVCs/PACs due to the observed relation between ECG and pulsatile time series obtained by a smart phone as shown in Fig. 2. The top panel is the ECG data with a PAC episode (noted by an arrow) while a bottom panel is its corresponding pulsatile time series. When a PAC occurs after a normal beat, the pulse signal is elongated and two peaks of the PAC and normal beats are merged into one peak in a pulsatile time series domain. Hence, the difference between consecutive beats (ΔPI) during a PAC/PVC episode is larger than the difference during normal episodes.

Our Poincare plot is divided into six sections with section IDs from 0 to 5. Section 0 is centered on the origin and surrounded by four boundaries, $x \le x_{\text{bound}}$, $x \ge -x_{\text{bound}}$, $y \le y_{\text{bound}}$ and $y \ge -y_{\text{bound}}$. The "short-short-short" and "long-long-long" PI sequences are confined to this section 0. Section 1 is bounded by $x > x_{\text{bound}}$, $y \le y_{\text{bound}}$ and $y \ge -y_{\text{bound}}$, and "short-short-long" PI sequences are matched to the section 1. Similarly, we define sections 2 to 5 as follows: section 2 for $x \le -x_{\text{bound}}$ and $y \ge y_{\text{bound}}$, section 3 for $x \le x_{\text{bound}}$, $x \ge -x_{\text{bound}}$

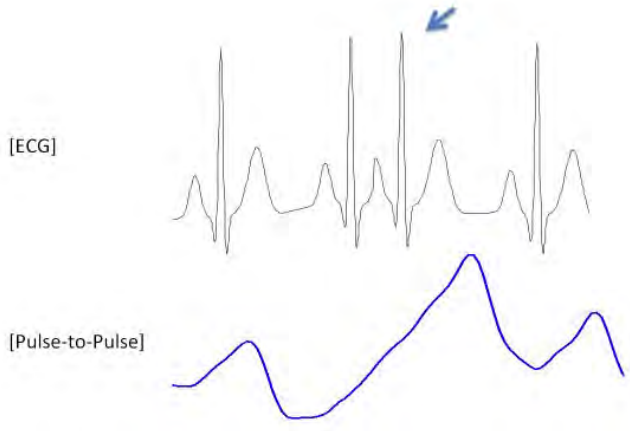


Fig. 2. ECG RR intervals versus pulse-to-pulse intervals from an iPhone 4S for a PAC subject. A PAC episode has longer pulse interval and larger amplitude compared to a NSR episode.

and $y < -y_{\text{bound}}$, section 4 for $x > x_{\text{bound}}$ and $y < -y_{\text{bound}}$, section 5 for $x > -x_{\text{bound}}$ and $y > y_{\text{bound}}$, and section 6 for $x < -x_{\text{bound}}$ and $y < y_{\text{bound}}$. The section boundaries x_{bound} and y_{bound} for x and y axis are set considering the pulse time series interval dynamics of NSR, Af, PVC, and PAC subjects. Then, these six sections covers two-dimensional Euclidean space and every combination of three consecutive PIs, i.e., "short-short-short" and "long-long-long" for section 0, "short-short-long" for section 1, "short-long-short" for section 2, "long-short-short" for section 3, "long-short-long" for section 4, "short-long-long" for section 5, and "long-long-short" for section 6.

Hence, NSR's Poincare plot is plotted within the sections 0 since NSR has regular R-R intervals (RRIs) and its corresponding PI sequence is "short-short-short-…" with small variance. On the other hand, AF's Poincare plot is irregularly spanned over 6 section. This is because AF has irregular RRIs and corresponding PI sequence are irregular with high variance.

For PVCs and PACs, we define that PAC that occurs every 2nd, 3rd, and 4th pulse, as the bigeminy, trigeminy, and quadrigeminy, respectively. For PVC/PAC quadrigeminy, the PI sequence is "short-short-long-short-short-long-…" where long PI exists every 3rd PI. Hence, its Poincare plots trajectory has a regular pattern of triangle spanning sections 1, 2, and 3. Similarly, the PVC/PAC trigeminy has a periodic PI sequence of "short-long-short-long-…" where short and long periodically oscillates. Hence, its Poincare plot also oscillates between sections 2 and 4. The patterns of quadrigeminy and trigeminy are specific compared to the patterns of NSR and AF. The PVC/PAC bigeminy shows similar pattern with NSR in that their paths are within section 0. However, the bigeminy can be discriminated from NSR using the observation that the bigeminy has longer PI and larger amplitude than NSR.

We applied these patterns of bigeminy, trigeminy, quadrigeminy of PVCs/PACs to discriminate them from NSR and AF since these patterns are regular and discernible from those of NSR and AF.

2) Kullback-Leibler Divergence

The KL divergence is used to measure the difference between probability distributions $p(x)$ and $q(x)$, and is defined by: $KL(q \| p) = -\int p(x) \log \left\{ \frac{q(x)}{p(x)} \right\} dx$. The KL divergence approach is appropriate for determining whether the unclassified PVC/PAC from Poincare plot is PVC or PAC due to the specific characteristics of PVC and PAC pulses. We first build $p_1(x)$ and $p_2(x)$ from the PVC and PAC training pulses, respectively. We then construct $q(x)$ from the unclassified PVC/PAC pulse measurement data. We determine the pulse of $q(x)$ is PVC if $KL(q \| p_1)$ is smaller than $KL(q \| p_2)$. Otherwise, the pulse is determined to be PAC.

The representative PVC pulse $p_1(x)$ and PAC pulse $p_2(x)$ obtained from PVC and PAC subjects have clear difference. Hence, the unclassified PVC/PAC data from Poincare Plot can be classified into PVC or PAC based on KL divergence method with $p_1(x)$, $p_2(x)$ and $q(x)$.

D. Performance Evaluation

The performance of the proposed arrhythmia discrimination for smart phone is evaluated with PAC and PVC as well as NSR, AF, PVC and PAC subjects. We set the thresholds $RMSSD_{\text{th}}$ and ShE_{th} based on ROC curve having the largest area, and set the boundaries of Poincare plot reflecting the NSR, AF, PVC, and PAC training data. We evaluate our discrimination algorithm in terms of sensitivity, specificity, and accuracy.

III. RESULTS

Using an iPhone, the pulsatile time series of 20 NSR, 20 AF, and 3 PVC, 4 PAC subjects are measured at UMass Medical Center [6]. We set the threshold values of RMSSD and Shannon entropy by $\text{RMSDD}_{\text{th}} = 0.1300$, $\text{ShE}_{\text{th}} = 0.7913$, respectively. Moreover, the boundary values of Poincare Plot are set to $x_{\text{bound}} = 0.1$ and $y_{\text{bound}} = 0.1$. Sensitivity, specificity and accuracy of the proposed arrhythmia discrimination algorithm is obtained. Our arrhythmia discrimination algorithm with the Poincare plot and KL divergence method combined with statistical metrics of RMSSD and ShE shows an sensitivity of 100% in detecting PVCs and PACs. For the discrimination of PVC, the algorithm shows specificity and accuracy of 1.0000 and 1.0000, respectively. Similarly, the proposed algorithm discriminates PAC with specificity and accuracy of 0.9767 and 0.9787, respectively.

IV. DISCUSSION AND CONCLUSION

In this paper, we have shown that NSR, AF, PVC, and PAC can be discriminated from pulsatile signal of the fingertip obtained by a smart phone. Considering that significant number of AF episodes can be paroxysmal and asymptomatic, arrhythmia discrimination algorithms, which are accurate, readily available and cheap, are in a high demand. With growing prevalence of smart phone, our approach using a smart phone in discriminating arrhythmia can address the needs to monitor arrhythmia in an efficient way. Specifically, since our smart phone-based approach does not required additional hardware such as ECG sensor, it is cost-effective and readily available. The proposed arrhythmia algorithms for smart phones performs AF detection with high accuracy and discriminates PVC and PAC with their specific types. Further study will evaluate usability of our algorithm in a diverse cohort of patients.

REFERENCES

[1] I. Savelieva, AJ. Camm, "Clinical relevance of silent atrial fibrillation: prevalence, prognosis, quality of life, and management," *J Interv Card Electrophysiol*, vol. 4 no. 2, pp. 369-82, Aug. 10, 2000.

[2] K. H. Humphries, C. R. Kerr, S. J. Connolly, G. Klein, J. A. Boone, M. Green, et al., "New-onset atrial fibrillation: sex differences in presentation, treatment, and outcome," *Circulation*, vol. 103, no.19, pp. 2365-70, May 23, 2001.

[3] P. Defaye, F. Dournaux, E. Mouton, "Prevalence of supraventricular arrhythmias from the automated analysis of data stored in the DDD pacemakers of 617 patients: the AIDA study. The AIDA Multicenter

Study Group. Automatic Interpretation for Diagnosis Assistance," *Pacing Clin Electrophysiol*, vol. 21, pp.250-5, 1998.

[4] V. Krasteva, I. Jekova, "QRS template matching for recognition of ventricular ectopic beats," *Ann Biomed Eng*. vol. 35, no. 12, pp. 2065-76, Sep. 7, 2007.

[5] C. Scully, J. Lee, J. Meyer, A. Gorbach, D. Granquist-Fraser, Y. Mendelson, et al, "Physiological Parameter Monitoring from Optical Recordings with a Mobile Phone," *IEEE Trans Biomed Eng*. Aug. 2, 2011.

[6] J. Lee, B. A. Reyes, D. D. McManus, O. Mathias, K. H. Chon, "Atrial fibrillation detection using an iPhone 4S," *IEEE Trans Biomed Eng*, vol 60, no. 1, Aug. 8. 2013.

[7] D. D. McManus, J. Lee, O. Maitas, N. Esa, R. Pidikiti, A. Carlucci, et al., "A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation. Heart rhythm," *the official journal of the Heart Rhythm Society*, Dec. 12, 2012.