33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Statistical Approach for the Detection of Motion/Noise Artifacts in Photoplethysmogram

Nandakumar Selvaraj*, Yitzhak Mendelson, Kirk H. Shelley, David G. Silverman, and Ki H. Chon, *Senior Member, IEEE*

Abstract—**Motion and noise artifacts (MNA) have been a serious obstacle in realizing the potential of Photoplethysmogram (PPG) signals for real-time monitoring of vital signs. We present a statistical approach based on the computation of kurtosis and Shannon Entropy (SE) for the accurate detection of MNA in PPG data. The MNA detection algorithm was verified on multi-site PPG data collected from both laboratory and clinical settings. The accuracy of the fusion of kurtosis and SE metrics for the artifact detection was 99.0%, 94.8% and 93.3% in simultaneously recorded ear, finger and forehead PPGs obtained in a clinical setting, respectively. For laboratory PPG data recorded from a finger with contrived artifacts, the accuracy was 88.8%. It was identified that the measurements from the forehead PPG sensor contained the most artifacts followed by finger and ear. The proposed MNA algorithm can be implemented in real-time as the computation time was 0.14 seconds using Matlab®.**

***Index Terms*—Motion and noise artifacts, Photoplethysmography, Shannon entropy, kurtosis.**

I. Introduction

There is growing interest in the real-time, wearable and ambulatory monitoring of vital signs using pulse oximeter sensors. However, motion and noise artifacts (MNA) are a serious obstacle in realizing this quest. Artifacts have been recognized as an intrinsic weakness of using the Photoplethysmogram (PPG) that limits its practical implementation and reliability for real-time monitoring applications. Artifacts are the most common cause of false alarms, loss of signal, and inaccurate measurements in clinical monitoring [1], where artifacts are more likely due to the voluntary and involuntary movements of the patient.

While the intelligent design of hardware elements such as PPG 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 to entirely avert artifacts. Various algorithms have also been attempted to isolate the effects of undesired artifacts with the outcomes being less than desired [2]. An algorithm based on the comparison between the heart rate (HR) measured from electrocardiogram (ECG) and pulse rate calculated from PPG for very short segments has been reported for reliable artifact detection in PPG signals [3]. This approach is not very efficient and practical, since it requires the additional recording of the ECG to achieve artifact detection in the PPG signal. Statistical measures such as skewness, kurtosis, Shannon entropy and Renyi's entropy have been shown useful for automatic detection of artifacts [2, 4]. However, no detailed quantitative results have been reported to verify their accuracy and suitability for successful detection of artifacts in PPG waveforms. Hence, a comprehensive and quantitative approach is needed to accurately and automatically detect the presence of artifacts in PPG data.

Our MNA detection algorithm is based on the computation of kurtosis and Shannon Entropy (SE) measures from the PPG segments. We hypothesize that our algorithm may provide accurate discrimination between artifact-free and artifact-contaminated PPG data. We tested the efficacy of our computational approach on multi-site PPG data containing involuntary artifacts recorded under clinical settings and on finger-PPG data containing controlled voluntary movements recorded in a laboratory setting.

II. Materials And Methods

A. Experimental protocol

We tested our algorithm on PPG signals obtained from two distinct scenarios as follows.

1. Involuntary movements: Multi-site PPG signals recorded from 10 healthy volunteers under supine resting condition for 5 to 20 minutes in clinical settings was used for our analysis. Three identical reflective infrared (940nm) PPG-probes (MLT1020; ADI Instruments, CO Springs, CO, USA) were placed at the finger, forehead and ear. While the finger and ear PPG probes were attached with a clip, the forehead probe was securely covered by a clear dressing. The PPG signals were recorded at 100 Hz with a Powerlab/16SPdata acquisition system equipped with a Quad Bridge Amp (ML795 & ML112; ADI Instruments) and a high-pass filter cut-off of 0.01 Hz. The subjects were

Manuscript received March 23, 2011. *Asterisk indicates corresponding author.* This work was funded in part by the Office of Naval Research work unit N00014-08-1-0244.

*N. Selvaraj, Y. Mendelson, and K. H. Chon are with the Department of Biomedical Engineering, Worcester Polytechnic Institute, Worcester, MA 01609 USA. (e-mail: nselvaraj@wpi.edu; ym@wpi.edu; kichon@wpi.edu).

K. H. Shelley, and D. G. Silverman are with the Department of Anesthesiology, Yale University, New Haven, CT 06520 USA. (e-mail: david.silverman@yale.edu; kirk.shelley@yale.edu).

not restricted from making any sort of movements during the recording procedure.

2. Voluntary movements: Finger-PPG signals were obtained from 14 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 0.05-10 Hz. The MP100 (BIOPAC Systems Inc., CA, USA) was used to acquire finger PPG signals at 100 Hz. After baseline recording for 5 min without any movements (i.e. clean data), motion artifacts were induced in the PPG data by left-right movements of the index finger 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. Such controlled movements were carried out 5 times for each level of noise. The recorded PPG signals from both protocols were analyzed offline using Matlab®.

B. Data Preprocessing

The PPG data were partitioned into 60s segments and shifted every 10s for the entire data. Each 60s PPG segment was subjected to a finite impulse response (FIR) band pass filter of order 64 with cut-off frequencies of 0.1 Hz and 10 Hz. To account for the time-dependent low-frequency trends associated with the PPG signal, either a low- or high-order polynomial detrending was used. We used in some cases as high as the 32nd-order polynomial fit to eliminate nonstationary dynamics in the PPG signal. The use of a high-order polynomial detrend is the key to an effective classification between clean and artifact-containing signals, which will be demonstrated in the Results Section. We visually examined the PPG waveforms in each data segment and classified them into clean vs. corrupted segments. Any sort of disruption in the pulse characteristics was labeled as corrupted segments.This was done in order to later determine the accuracy of the method.

C. Computational measures for artifact detection

Following the preprocessing of each PPG data segment, our approach for the detection of artifacts involves the computation of the following two parameters.

1. Kurtosis: Kurtosis is a statistical measure used to describe the distribution of observed data around the mean. It represents a heavy tail and peakedness or a light tail and flatness of a distribution relative to the normal distribution. The kurtosis is defined as:

$$k = \frac{E(x-\mu)^4}{\sigma^4} \tag{1}$$

where μ is the mean of x, σ is the standard deviation of x, and $E(t)$ represents the expected value of the quantity t.

2. Shannon entropy: SE quantifies how much the probability density function (PDF) of the signal is different from a uniform distribution and thus provides a quantitative measure of the uncertainty present in the signal [5]. SE can be calculated as

$$SE = -\sum_{i=1}^{k} \frac{\mathrm{p(i)}*\log\left(p(i)\right)}{\log\left(\frac{1}{k}\right)} \tag{2}$$

where i represents the bin number, and $p(i)$ is the probability distribution of the signal amplitude. Presently, we have used 16 bins $(k = 16)$ to obtain a reasonably accurate measure of SE [6].

D. Statistical analysis of computational measures

The data are reported as mean ± SE. The nonparametric Mann Whitney test was conducted to find the significance levels ($p<0.05$) for the SE and kurtosis measures obtained from clean vs. corrupted PPG segments of the involuntary motion protocol. Meanwhile, the nonparametric Kruskal-Wallis test with Dunn's multiple comparison post test was conducted to find the significance ($p<0.05$) for the two measures obtained from clean vs. 10-50% noise-corrupted PPG segments of the voluntary motion protocol.

E. Detection of motion/noise artifacts

By varying kurtosis values from 0 to 10 with an increment of 0.1, and SE values from 0.5 to 1.0 with an increment of 0.01, we conducted receiver-operator characteristic analysis for the population of SE and kurtosis values obtained from the respective pool of clean and corrupted PPG segments of both protocols. We evaluated the threshold values for kurtosis and SE that produced the optimal sensitivity and specificity for the detection of artifacts.

The decision rules for the detection of artifacts were formulated as follows:

$$DK_i = \begin{cases} 1 & if\ K_i \le K_{Th} \\ 0 & if\ K_i > K_{Th} \end{cases} \tag{3}$$

where DK_i refers to the decision for artifact detection based on K_i, kurtosis for the ith segment. '1' represents clean data, whereas '0' represents corrupted data. K_{Th} refers to the Kurtosis threshold.

$$DS_i = \begin{cases} 1 & if\ SE_i \ge SE_{Th} \\ 0 & if\ SE_i < SE_{Th} \end{cases} \tag{4}$$

where DS_i refers to the decision for artifact detection based on SE_i, SE for the ith segment. '1' represents clean data whereas '0' represents corrupted data. SE_{Th} refers to the SE threshold.

We further investigated the fusion of kurtosis and SE metrics with their optimal threshold values for the artifact detection and quantified the sensitivity and specificity for the fusion of these two metrics. The decision rule for the detection of artifacts using a fusion of kurtosis and SE is:

$$FD_i = \begin{cases} 1 & if\ DK_i + DS_i = 2 \\ 0 & if\ DK_i + DS_i \ne 2 \end{cases} \tag{5}$$

where FD_i refers to the fusion decision for artifact detection based on both DK_i and DS_i for the ith segment. '1' represents clean data whereas '0' represents corrupted data.

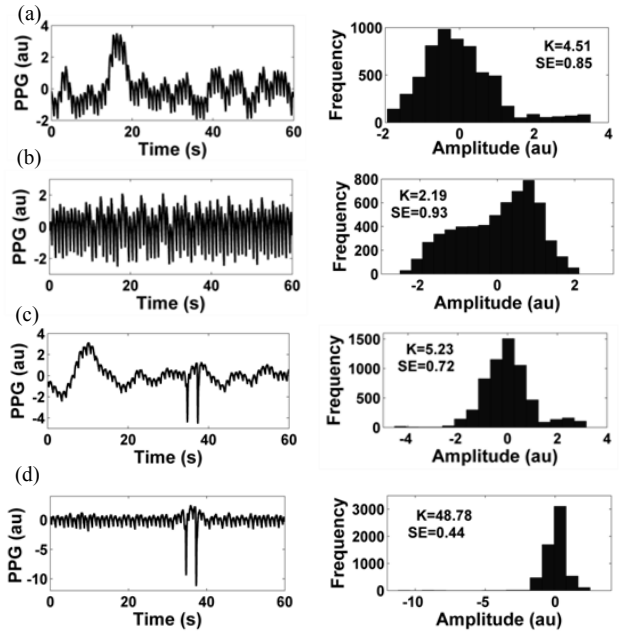

Fig. 1. Sample clean (a-b) and corrupted (c-d) ear-PPG segments applied with linear (a, c) and 32nd order polynomial detrends (b, d) are shown along with their respective histograms and calculated kurtosis (K) and Shannon entropy (SE) values. The higher-order polynomial detrending is critical to enhance the specificity in the presence of physiological baseline drift (a) and the sensitivity in the presence of artifacts (c).

III. Results

Our use of a high-order polynomial detrend for artifact detection is illustrated in Fig. 1. In a sample of a clean ear-PPG segment with strong baseline drift, the linear detrend (Fig.1a) resulted in a long tail in its PDF. Meanwhile, the high-order (e.g., 32nd-order) polynomial detrend (Fig. 1b) on the same data resulted in SE and kurtosis values which were similar to that of clean data. In a segment of corrupted PPG data subjected to a linear detrend (Fig. 1c), the low frequency trend masked the high-frequency artifacts. But, the high-order polynomial detrend (Fig. 1d) drastically changed the PDF, SE and kurtosis values, and clearly differentiated from those of clean data. Thus, the high-order polynomial detrend is a critical component in enhancing the specificity in the presence of strong baseline drift and the sensitivity in the presence of artifacts.

Detection of artifacts in multi-site PPG data with involuntary motion/noise artifacts

Followed by the higher-order polynomial detrending, the statistical measures were obtained for artifact detection. The SE values obtained for the clean vs. corrupted data segments were (0.90±0.0 vs. 0.59±0.03), (0.88±0.0 vs. 0.65±0.01) and (0.88±0.0 vs. 0.72±0.0) for ear, finger and forehead PPG signals, respectively. The kurtosis values obtained for the clean vs. corrupted data segments were (2.29±0.01 vs. 20.77±3.68), (2.60±0.02 vs. 12.77±1.42) and (2.30±0.02 vs. 5.48±0.16) for ear, finger and forehead PPG signals. The corrupted PPG segments showed significant decrease (P<0.001) in SE value and significant increase (P<0.001) in kurtosis value in all three probe sites as compared to their respective clean PPG segments. The optimal threshold values of SE (SE_{Th}) and kurtosis (K_{Th}) were found to be 0.8 and 3.5, respectively. Their specificity, sensitivity and accuracy values for the artifact detection in all three probe sites are given in Table 1. SE (SE_{Th}=0.8) offered an accuracy of 99.0%, 94.4% and 91.3%, and Kurtosis (K_{Th}=3.5) offered an accuracy of 99.6%, 97.0% and 94.0% to classify clean vs. corrupted segments in ear, finger and forehead PPG signals, respectively. Table 1 also provides the specificity and sensitivity values obtained for the three PPG sites using the fusion detection with S_{Th}=0.8 and K_{Th}=3.5. The fusion detection of SE and kurtosis metrics offered an accuracy of 99.0%, 94.8% and 93.3% for artifact detection for ear, finger and forehead PPG signals, respectively.

TABLE I

THE PERFORMANCE OF SE_{TH}=0.8, K_{TH}=3.5 AND FUSION OF THESE TWO METRICS FOR THE DETECTION OF MNA IN MULTI-SITE PPG RECORDED WITH INVOLUNTARY MOVEMENTS

	SE_{Th}=0.8			K_{Th}=3.5			Fusion detection		
	Ear PPG	Fgr PPG	Fhd PPG	Ear PPG	Fgr PPG	Fhd PPG	Ear PPG	Fgr PPG	Fhd PPG
Sp (%)	98.9	94.9	92.1	99.8	96.8	99.4	98.9	93.8	91.9
Se (%)	100	92.1	89.6	95	97.8	83.0	100	99.3	96.3
Acc (%)	99.0	94.4	91.3	99.6	97.0	94.0	99.0	94.8	93.3

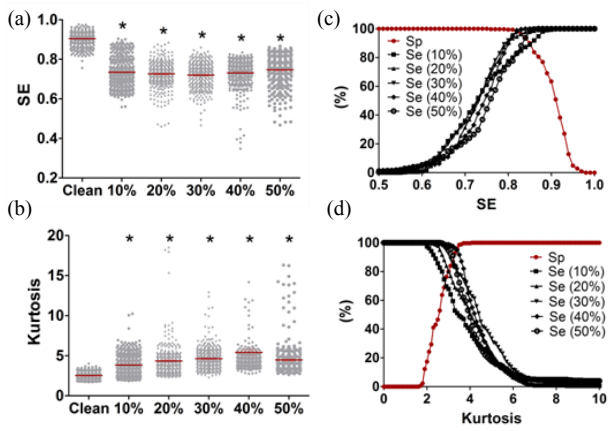

Fig. 2. The values of (a) SE and (b) kurtosis measures obtained for clean and corrupted finger-PPG segments with 10% to 50% noise (350 segments in each state). Mean±SEM values are denoted as bars;*P<0.001. The specificity (Sp) and sensitivity (Se) analysis are shown for (c) SE and (d) kurtosis measures for the classification of clean vs. corrupted finger-PPG segments recorded with voluntary MNA.

Detection of artifacts in finger-PPG data with voluntary motion/noise artifacts

The SE and kurtosis values obtained for clean (n=350 segments) and corrupted segments with varying levels (10-50%) of added artifacts (each with 350 segments) and their specificity and sensitivity analysis are shown in Fig. 2. Similar to the data with involuntary artifacts, a significant (P<0.001) decrease in SE, and a significant increase

TABLE II

THE PERFORMANCE OF SE_{TH}=0.8, K_{TH}=3.5 AND FUSION OF THESE TWO METRICS FOR THE DETECTION OF MNA IN FINGER-PPG RECORDED WITH VOLUNTARY MOVEMENTS

	Sp (%)	Se (%)					Overall Se (%)	Acc (%)
		10% noise	20% noise	30% noise	40% noise	50% noise		
SE_{Th}=0.8	99.4	77.5	90.3	91.9	89.7	75.4	85.0	87.4
K_{Th}=3.5	98.6	49.4	72.9	77.5	91.4	71.7	72.6	77.0
Fusion detection	98.3	77.2	90.0	91.6	95.7	79.7	86.9	88.8

($P<0.001$) in kurtosis were found for all levels of noise (10% to 50%) as compared to the clean PPG segments (Fig. 2a-b). Note that both SE and kurtosis values do not reflect the varying level of noise present in the PPG segments. The sensitivity and specificity values obtained for the optimal threshold values SE_{Th}=0.8 and K_{Th}=3.5 are given in Table 2. SE offered specificity of 99.4% and sensitivity of 85.0%, whereas kurtosis offered specificity of 98.6% and sensitivity of 72.6% for the finger-PPG signals induced with voluntary left-right movements. When kurtosis and SE measures were combined for the artifact detection with the threshold values of SE_{Th} =0.8 and K_{Th} =3.5, we obtained a specificity of 98.3% and sensitivity of 86.9%.

IV. Discussion

A quantitative and accurate approach for real-time artifact detection in raw PPG data has been elusive to date. Our algorithm based on the computation of SE and kurtosis offered very high accuracy for automatic detection of artifacts in multi-site PPG data with involuntary and voluntary artifacts. Our success is predicated the use of preprocessing stage with higher order polynomial detrend that accentuated the differences in SE and kurtosis values between clean and corrupted data segments as shown in Fig. 1. Note that voluntary and involuntary movements are significantly different from each other as the latter are recorded in a clinical setting while the former are contrived in a controlled laboratory environment. Certainly, the involuntary movements are a subset of the wide ranges of artifacts encountered in voluntary movements. Thus, they represent only a small fraction of the true artifact dynamics. Despite the different noise dynamics, we obtained the same optimal threshold values for SE and kurtosis for both voluntary and involuntary movements. This is important because it suggests that our derived threshold values are nearly optimal and, therefore, further tuning may not be required. Often, threshold values are derived from training data, so the accuracy becomes degraded with independent test data, but our results suggest that this is another advantage of our method.

While the specificity of SE and kurtosis measures was very high in finger-PPG data with voluntary movements, the sensitivity was relatively lower (Table 2). This is primarily because repetitive motions such as the left-right movements exhibit some degree of periodicity and repetitiveness which do not significantly alter the peakedness or tail characteristics of the PDF as compared to the clean data segments. Hence, we observed overlap in SE and kurtosis values between the clean and artifact-corrupted PPG data especially at the 10% noise level (Fig. 2b), which decreased the sensitivity to a great extent. While the finger-PPG sensor employed in the voluntary artifact protocol is from a different manufacturer than the sensor used during clinical experiments, we do not believe this has led to lower sensitivity values with the former data.

In our second experimental protocol, we simulated artifacts via voluntary left-right finger movements for 10% to 50% of the total 60 s duration of a PPG segment, since episodes of motion probably occur for a short duration of time, typically not exceeding 30s in real clinical situations [7]. The segment size of the PPG signal for our computational algorithm has been chosen to be 60s instead of shorter segments of a few seconds duration due to the fact that the statistical measures perform more accurately on longer segments than on very short segments. Our algorithm is able to detect the artifact-corrupted data segments with good accuracy when the data length is 1 minute in duration. While we have chosen the duration of the sliding window length to be 10 seconds, it can be reduced to 1 second but at the expense of a significant computational load. Moreover, the purpose of our algorithm is to identify artifacts in PPG data rather than to localize them so that replacement data can be inserted in place of the corrupted segment. Thus, a smaller sliding window segment does not provide any better accuracy or meet our needs.

In conclusion, our computational algorithm was presently tested for the detection of involuntary and voluntary motion artifacts. Our approach offered very high sensitivity and specificity for the detection of voluntary and involuntary artifacts, and is real-time realizable as it only takes 0.137 second a 1.66 GHz Intel Core2 processor using Matlab®.

References

[1] 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.

[2] R. Krishnan, B. B. Natarajan, and S. Warren, "Two-stage approach for detection and reduction of motion artifacts in photoplethysmographic data," *IEEE Trans Biomed Eng,* vol. 57, no. 8, pp. 1867-76, Aug, 2010.

[3] C. F. Poets, and V. A. Stebbens, "Detection of movement artifact in recorded pulse oximeter saturation," *Eur J Pediatr,* vol. 156, no. 10, pp. 808-11, Oct, 1997.

[4] N. Mammone, and F. C. Morabito, "Enhanced automatic artifact detection based on independent component analysis and Renyi's entropy," *Neural Netw,* vol. 21, no. 7, pp. 1029-40, Sep, 2008.

[5] S. Tong, Z. Li, Y. Zhu, and N. V. Thakor, "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.

[6] S. Dash, K. H. Chon, S. Lu, and E. A. Raeder, "Automatic real time detection of atrial fibrillation," *Ann Biomed Eng,* vol. 37, no. 9, pp. 1701-9, Sep, 2009.

[7] 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.