35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

A New Movement Artifact Detector for Photoplethysmographic Signals*

Carlos A. Robles-Rubio, *Student Member, IEEE*, Karen A. Brown, and R. E. Kearney, *Fellow, IEEE*

Abstract— **Oximeters are commonly used in abbreviated cardiorespiratory studies (ACS) to monitor blood oxygen saturation and heart rate using the photoplethysmography (PPG) signal. These data are prone to movement artifacts, especially in infants who move or need to be handled often. Therefore segments of PPG data contaminated by movement artifact must be detected as a first stage of analysis. In ACS this identification is generally done manually, by having an expert visually assess the quality of the signal. This is subjective and very time consuming, especially for long data records. For this reason we present a novel detector of PPG movement artifacts that uses moving average filters to remove trends, reduce the effect of white noise, and notch filter pulse-related information. The normalized root mean square of the filtered signal is then used as a detection statistic. We demonstrate its detection properties using a data set from infants recovering from anesthesia, and show that it performs better than other automated methods based on entropy or higher-order statistics. Furthermore, the new method is more robust than the other methods in the presence of large noise.**

I. Introduction

Abbreviated cardiorespiratory studies (ACS) are commonly performed to assess/study respiratory conditions such as apnea, when full polysomnography (PSG) is not indicated or available (e.g., in the home [1], the recovery room [2], or the intensive care unit [3]). These studies acquire a subset of the PSG signals, typically including blood oxygen saturation (SaO_2) and photoplethysmography (PPG), measured with an oximeter, as well as the ribcage and abdomen respiratory movements from respiratory inductive plethysmography (RIP). SaO_2 is used to monitor oxygen levels with the intention to detect episodes of desaturation [4]. PPG is used to monitor heart rate, and its variability (PPGV) provides information about autonomic nervous system function similar to that provided by heart rate variability estimated from the electrocardiogram [5]. However, both PPG and SaO_2 are prone to movement artifacts, especially in infants, who move and/or need to be handled often. Therefore data segments corrupted by movement artifacts must be identified and removed as part of any analysis procedure. This has limited the utility of monitoring tools based on oximeters alone [6], and as a consequence the use of ACS is not widespread.

*Research supported in part by the Natural Sciences and Engineering Research Council of Canada. The work of CARR was supported in part by the Mexican National Council for Science and Technology.

CARR and REK are with the department of Biomedical Engineering, McGill University, Montreal, QC H3A 2B4, Canada (e-mail: carlos.roblesrubio@mail.mcgill.ca; robert.kearney@ mcgill.ca).

KAB is with the department of Anesthesia, McGill University Health Center, Montreal, QC H3A 2B4, Canada (e-mail: karen.brown@mcgill.ca).

Visual inspection of the PPG is the preferred method to identify movement artifacts [4]. However, it is subjective, suffers from low intra- and inter-operator repeatability, and is very time-consuming. Consequently, it is difficult to use the PPG for clinical monitoring and analysis of long recordings such as those acquired overnight. Consequently, there have been efforts to develop artifact reduction techniques to recover the underlying PPG data or estimate the heart rate and SaO_2 from corrupted segments (e.g., "motion-resistant" oximeters [7]). However, before any artifact reduction technique is applied, it is necessary first to detect corrupted segments so that artifact-free data do not undergo unnecessary processing that might alter relevant information. A variety of methods have been proposed for this, including methods based on higher-order statistics in the time and frequency domains [8, 9] (e.g., kurtosis, spectral kurtosis), as well as Shannon entropy [9]. However, performance of these methods varies widely and they have not been evaluated under the high noise conditions that prevail in environments such as intensive care units or the home.

This paper presents the design of a novel, robust detector of PPG movement artifacts suitable for real-time implementation, and compares its performance to that of other methods when applied to a representative infant data set. The paper is organized as follows: Section II summarizes existing PPG artifact detectors and describes the novel method; Section III describes the data acquisition and analysis procedures used to evaluate the detectors; Section IV reports the performance results; and Section V discusses the findings and provides concluding remarks.

II. Detectors of Movement Artifacts

This section first reviews three detection statistics commonly used for detection of PPG movement artifacts: entropy, kurtosis and skewness. Then, the new detection statistic is developed analytically.

A. Shannon Entropy

Movement artifacts generally appear as chaotic signals in contrast to the artifact-free quasi-periodic pattern of the PPG (see Fig. 1). The level of uncertainty in a stochastic signal, such as the PPG, can be quantified using Shannon entropy [10]. This concept was used to design an entropy-based detection statistic based on the hypothesis that segments contaminated by movement artifacts would have higher entropy than artifact-free segments [9].

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

The entropy of a discrete random variable is defined as the expected value of the logarithm of its probability mass function (PMF) [10]. Estimating the entropy of a continuous signal is achieved by quantizing it into k discrete values, then estimating a PMF, and finally estimating the expected value of the logarithm of this PMF. Estimates at each sample time are obtained by repeating this procedure for each sample in the record using a sliding window of length N_{ent}. Thus, for the PPG signal the entropy at each time point can be estimated as:

$$E[n] = \frac{-\sum_{i=1}^{k} P_{PPG[\theta]}(i) \log_b \left(P_{PPG[\theta]}(i) \right)}{\log_b (k)} \qquad (1)$$

where b is the base of the logarithm used for the calculation of entropy, and $P_{PPG[\theta]}$ is the estimate of the PMF of the PPG signal in the interval $\theta \in [n - (N_{ent} - 1) / 2, n + (N_{ent} - 1) / 2]$ quantized using k bins. Note that (1) is normalized with respect to $\log_b(k)$, the maximum possible value [10], so that $E[n]$ ranges from 0 to 1.

To follow the implementation of the entropy-based detector developed in [9], the PPG signal was band-pass filtered prior to (1) with a finite impulse response (FIR) filter of order 64 with cut-off frequencies of 0.1 and 10 Hz, to remove low frequency trends and higher frequencies where the PPG was not expected to contain any significant power.

B. Higher-order Statistics

Two higher-order statistics, kurtosis and skewness, have been applied for detection of PPG movement artifacts [8, 9]. Kurtosis is defined as the standardized fourth central moment of a stochastic signal. A continuous estimator for the kurtosis of the PPG is:

$$K[n] = \frac{\mu_4^{ppg}[n]}{\left(\mu_2^{ppg}[n] \right)^2} - 3 \qquad (2)$$

where μ_p^{ppg} corresponds to the p^{th} central moment of the PPG signal over a window of length $N_{HOS} = 2L + 1$ samples estimated as:

$$\mu_p^{ppg}[n] = \begin{cases} \dfrac{\sum_{i=n-L}^{n+L} PPG[i]}{2L+1}, & if\ p = 1 \\ \dfrac{\sum_{i=n-L}^{n+L} \left(PPG[i] - \mu_1^{ppg}[i] \right)^p}{2L+1}, & if\ p \neq 1 \end{cases} . \qquad (3)$$

Signals whose probability density functions (PDF) have a sharp peak and slow-decaying tails will have higher kurtosis compared to those with a broad peak and fast-decaying tails.

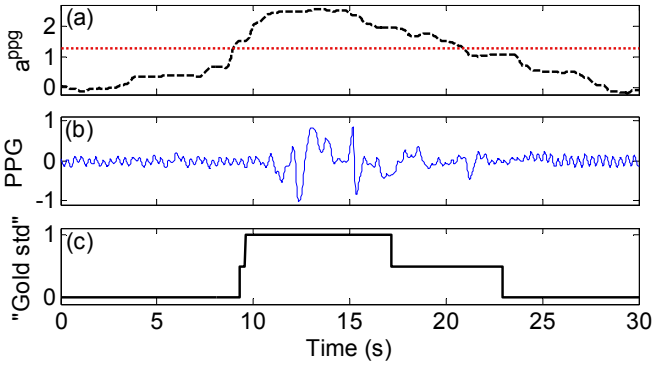

Fig. 1. (b) Representative epoch of infant photoplethysmography (PPG) signal; (a) corresponding output of our novel movement artifact detection statistic a^{ppg} (black dashed line) and the optimum detection threshold (red dotted horizontal line); and (c) "gold standard" classification from the manual analysis of an expert (artifact-free = 0, movement artifact = 1, uncertain = 0.5).

Skewness, the standardized third central moment, is a metric of the lack of symmetry in the PDF. PDFs with a longer left tail will have negative values of skewness, while those with a longer right tail will have positive values. PPG skewness can be estimated over a window of length N_{HOS} as:

$$S[n] = \frac{\mu_3^{ppg}[n]}{\left(\mu_2^{ppg}[n] \right)^{3/2}} . \qquad (4)$$

The hypothesis underlying the use of either kurtosis or skewness is that the shape of the PPG PDF will change in the presence of movement artifact, and that this change will be large enough to distinguish between the two states.

The PPG signal was band-pass filtered before estimating (2) and (4), using a digital zero-phase forward-backward Chebyshev Type I filter of order 6 with cut-off frequencies of 0.3 and 12 Hz. Note that this filter was implemented following the design from [8], where kurtosis and skewness were used to detect PPG movement artifacts.

C. Proposed Detection Statistic

In reviewing PPG data we noted that: (a) artifact-free PPG has the quasi-periodic waveform illustrated in Fig. 1; and (b) PPG movement artifacts comprise stochastic low frequency noise whose amplitude is typically larger than that of the PPG. We used these observations to design a detection statistic for the identification of artifacts in PPG.

The method attempts to estimate the residual component of the PPG signal after reducing the effect of white noise and notch filtering the pulse-related information. Then, the detection statistic a^{ppg} estimates the root mean square (RMS) of this residual. Since movement artifacts are expected to have larger amplitude than artifact-free signals, a^{ppg} will take lower values during no artifact and higher values when the signal is contaminated. Fig. 1 shows an example of a^{ppg} during artifact-free and corrupted PPG.

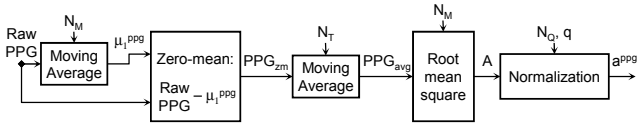

Fig. 2. Block diagram for the computation of a^{ppg}, the proposed statistic.

The complete procedure is illustrated in the block diagram in Fig. 2. First, a zero-mean version of the raw PPG signal is obtained as $PPG_{zm}[n] = PPG[n] - \mu_1^{ppg}[n]$, using a window of length $N_M = 2L + 1$ samples. PPG_{zm} is then filtered by a moving average filter of length N_T, whose frequency response is illustrated in Fig. 3. It is a low-pass filter with cut-off frequency of $f_c \approx f_s / (2N_T)$ and deep nulls at $f = mf_s / N_T$, where f_s is the sampling frequency and m is an integer ≥ 1. The value of N_T is set equal to f_s / f_{hr}, the inverse of the modal heart rate of the population (estimated from the complete data set) so that the first null occurs at f_{hr}. This filter reduces the amplitude of the most frequent pulse component in the signal as well as that of additive white noise. The RMS of the filtered signal PPG_{avg} is computed over a sliding window of length N_M as:

$$A[n] = \sqrt{\frac{1}{N_M} \sum_{i=n-(N_M-1)/2}^{n+(N_M-1)/2} PPG_{avg}^2[i]}\,. \quad (5)$$

The detection statistic is then defined as:

$$a^{ppg}[n] = \ln\left(\frac{A[n]}{A_{(q)}[n]}\right) \quad (6)$$

where $A[n]$ is normalized to account for nonstationary conditions in the amplitude of the PPG. To accomplish this, $A[n]$ is divided by $A_{(q)}[n]$, the q^{th} quantile of A in the sliding window $\theta \in [n - N_Q + 1\ ,\ n]$, $N_Q >> N_M$. To ensure normalization is done with respect to artifact-free data, at least $q \cdot N_Q$ samples in θ must be artifact-free at all times. In this study at least 10% of the samples ($q = 0.1$) in a 10 min long sliding window ($N_Q = 600\ f_s$) were expected to be artifact-free.

For windows that overlap by less than $N_Q - 1$ samples, $A_{(q)}[n]$ may be estimated efficiently as follows:

$$\hat{A}_{(q)}[\theta] = A_{(q)}[n],\ \ \theta \in [n, n + N_Q - N_o - 1] \quad (7)$$

where $N_o < N_Q$ is the number of overlapping samples. The first $N_Q - 1$ samples are set to $A_{(q)}[N_Q]$, the estimate of the first segment.

III. Evaluation Procedures

A. Subjects and Data Acquisition

The detectors were evaluated using data from 23 infants (18 male), postmenstrual age [11] 42.6±2.2 weeks and weight 3.6±0.9 kg, recorded in the postoperative period after elective inguinal herniorrhaphy with general anesthesia. Written informed parental consent was obtained and the study was approved by the Institutional Ethics Review Board of the Montreal Children's Hospital (MCH). Subjects were recruited for a prospective study of Postoperative Apnea in infants. Data from a subset of these patients was reported previously in a different study [12]. PPG and SaO_2, as well as ribcage and abdominal RIP, were recorded at a sampling rate of $f_s = 50$ Hz, for 6 to 12 hours in accordance with the MCH practice guidelines for apnea monitoring (see [12] for details).

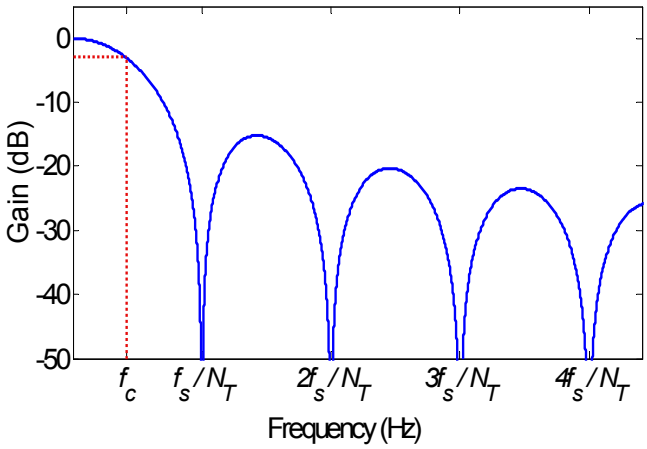

Fig. 3. Frequency response of a two-sided symmetric moving average filter of length N_T samples. Note that the filter is zero-phase so only the gain is shown. The cut-off frequency (f_c) corresponds to a gain of -3 dB (red dotted lines), and f_s is the sampling frequency.

B. "Gold Standard"

One of the investigators (KAB) used an interactive, graphical, manual classification tool, to visually identify movement artifact (MVT) segments in the PPG signal. MVT was defined as any motion not related to pulse, including chaotic and bad signal segments.

To account for the variability of manual classification, the scorer was asked to analyze each data set in two separate sessions. The "gold standard" scores were then determined as follows: (a) samples with two MVT scores were classified as MVT; (b) samples with no assignment to MVT were classified as artifact-free; and (c) samples with differing scores were considered to be uncertain and were excluded from the analysis.

Fig. 1 shows a representative epoch of a PPG signal with a segment corrupted by movement artifact, and illustrates the "gold standard" classification for that epoch.

C. Detection Performance

These "gold standard" scores were used to estimate two nonparametric PDFs for each detection statistic: one for samples classified as MVT (i.e., the alternative hypothesis H_1), and one for samples classified as artifact-free (i.e., the null hypothesis H_0). These PDFs were used to generate the receiver operating characteristics (ROC) curves, defining

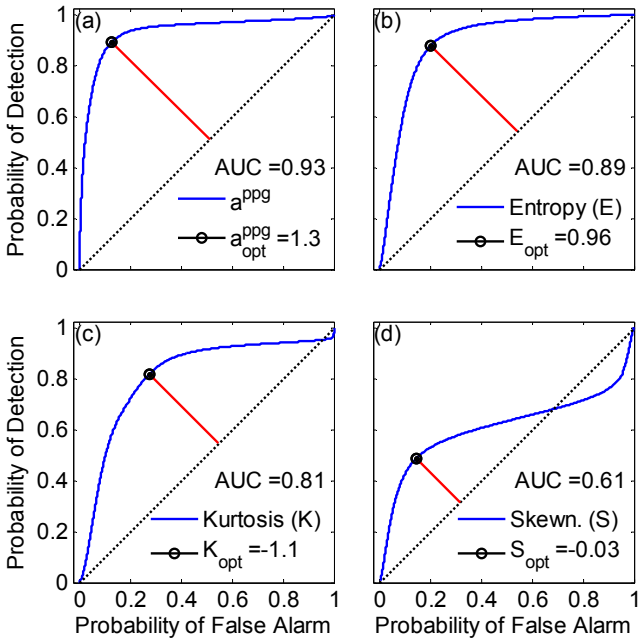

Fig. 4. Detection of photoplethysmography movement artifacts. Receiver operating characteristics (ROC) of: (a) proposed detector; (b) entropy; (c) kurtosis; and (d) skewness. The subscript *opt* indicates optimum threshold. AUC = Area under ROC curve.

how the probability of detection (P_D) and probability of false alarm (P_{FA}) varied as a function of the threshold value. For each statistic the optimal threshold γ_{opt} was selected as the point furthest from the chance line ($P_D = P_{FA}$); this defined the best tradeoff between P_D and P_{FA}. The area under the ROC curve (AUC) was used as a measure of performance. An $AUC = 1$ indicates perfect detection and $AUC = 0.5$ corresponds to the performance expected by chance.

D. Robustness with Noisy Data

To assess the performance of the detection statistics under noisy conditions, we generated independent sequences of white Gaussian noise to simulate electronic noise (using MATLAB™, The MathWorks Inc.), and pink (1/f) Gaussian noise [13] to simulate low frequency trends observed in PPG recordings. These sequences were scaled and added to the infant PPG data to generate a Signal-to-Noise Ratio (SNR) that ranged from *-20* to *20 dB*. The AUC was computed for each combination of SNR and noise type.

IV. Results

A. Manual Analysis

Of the 23 data sets (206 hrs), 22 were analyzed twice by the expert, and one was analyzed three times. This analysis yielded a total of 8,318 MVT (32.7 hrs) and 8,413 artifact-free (165.6 hrs) segments; 3.7% of the total data was classified as uncertain (7.7 hrs).

B. Detection Performance

The 23 data sets were analyzed with the four detectors described above. The new detector, a^{ppg}, used a sliding window of length $N_M = 251$, and the normalization

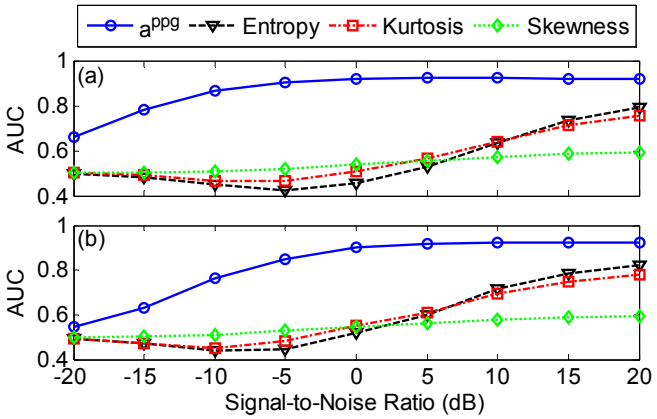

Fig. 5. Performance of photoplethysmography movement artifact detectors in the presence of added (a) white and (b) pink noise. AUC = Area Under receiver operating characteristics (ROC) Curve.

parameters were $q = 0.1$, $N_Q = 30{,}001$ and $N_o = N_Q - 2f_s$. The value of N_T was set to 21, corresponding to the inverse of f_{hr} = 2.4 Hz (i.e., 144 beats per minute), the most frequent heart rate value estimated from all of the 23 infant data sets.

The lengths of the sliding windows for the other statistics were set to $N_{HOS} = N_{ent} = N_M = 251$. We used the natural logarithm for the computation of entropy (i.e., $b = e$), and estimated the PMFs using $k = 16$ bins as in [9].

Fig. 4 shows the ROCs for the different detectors. The new statistic, a^{ppg}, performed best with an $AUC = 0.93$ (optimal operating point: $P_D = 0.89$, $P_{FA} = 0.13$), followed by the detectors based on entropy ($AUC = 0.89$, $P_D = 0.88$, $P_{FA} = 0.2$), kurtosis ($AUC = 0.81$, $P_D = 0.82$, $P_{FA} = 0.27$), and skewness ($AUC = 0.61$, $P_D = 0.49$, $P_{FA} = 0.15$).

C. Performance with Noisy Data

Detector performance as a function of SNR is shown in Fig. 5(a) for white and Fig. 5(b) for pink additive noise. The new detector was much more robust that the other detectors, maintaining an $AUC \geq 0.9$ for $SNR \geq 0$ *dB* for both white and pink noise. In contrast, the AUC of the entropy, kurtosis and skewness detectors never reached *0.9*, and decreased with increasing noise falling below $AUC \leq 0.55$ at $SNR \leq 0$ *dB*, a performance no better than chance.

V. Discussion

This paper presents a new method for detection of movement artifacts in photoplethysmography signals, and compares its performance to other detectors based on entropy [9] and higher-order statistics [8, 9]. The new method had the best discriminative ability ($AUC = 0.93$) and maintained an excellent performance in the presence of severe noise, due in part to the noise reduction obtained from the moving average filter. Indeed, its performance was independent of the noise level for $SNR \geq 0$ *dB*, as evidenced by the constant value of AUC in this SNR range. In contrast, the detection ability of entropy and higher-order statistics quickly decreased at lower noise levels ($SNR \leq 10$ *dB*). This property is especially relevant for signals acquired in intensive care units or the home; environments where high noise is common due to varying ambient light conditions [14] as well as

electromagnetic interference from other medical instruments and wireless telecommunication devices [15]. Note that this study worked with data acquired only from infants, but similar results could be expected in data from adults where movement artifacts have similar stochastic low frequency and large amplitude characteristics.

Previous studies have evaluated the performance of detectors using short data sets (< 15 hrs), acquired mostly under controlled conditions in laboratory settings [8, 9]. Moreover, they used an epoch-by-epoch approach, in which the recording was segmented into separate frames, and then each frame was classified as corrupted or artifact-free. This may bias the evaluation since subtle events may be masked if only a short portion of the epoch is corrupted, or artifact-free segments may be discarded if they happen to be grouped in an epoch with a few corrupted segments. In contrast, this work analyzed a much larger data set (> 200 hrs) acquired in a clinical setting, and obtained continuous estimates of all detection statistics. As a result performance was assessed on a sample-by-sample basis. This provided a detailed evaluation of performance treating the data as a continuous signal, which took into account the exact start and end times of events.

The new detector uses operations that can be readily implemented for real-time use. Additionally, it uses zero-phase (symmetric) FIR filters, so that a real-time implementation would have a response lag of only ½ of the length of the filters (a total of 5.2 s). An additional advantage is that it removes the local mean of the PPG and normalizes the value of the statistic to a local reference. As a result it will continue to work under nonstationary conditions where the offset, PPG amplitude, or both change with time.

The method provides improved detection of PPG movement artifacts even in high noise conditions. Since corrupted segments can be promptly discarded from further analysis, the new detector makes it possible to: (a) better analyze ACS data acquired in challenging environments such as intensive care units or the home; (b) study autonomic nervous system function based on variability of the heart rate estimated from the PPG signal (PPGV) [5]; and (c) perform those analyses in data acquired during long sessions (e.g., overnight). Additionally, if required by the application, artifact reduction techniques [16-18] could be applied to recover the underlying PPG data from corrupted segments only, while preventing additional unnecessary processing in artifact-free segments.

The new detector attempts to cancel the pulse-related information on the PPG signal to then estimate the RMS of the result. This is accomplished using a moving average to notch filter the most frequently observed heart rate value in the population. It might be possible to improve the notch filter performance and hence the distinction between artifact-free and corrupted segments varying the length of the moving average filter, making the null point adaptively track the heart rate at each time. Future work will evaluate this possibility.

REFERENCES

[1] G. M. Nixon and R. T. Brouillette, "Diagnostic techniques for obstructive sleep apnoea: is polysomnography necessary?," *Paediatr. Respir. Rev.,* vol. 3, pp. 18-24, 2002.

[2] A. Aoude, R. E. Kearney, K. A. Brown, H. Galiana, and C. A. Robles-Rubio, "Automated Off-Line Respiratory Event Detection for the Study of Postoperative Apnea in Infants," *IEEE Trans. Biomed. Eng.,* vol. 58, pp. 1724-1733, 2011.

[3] D. Precup, C. A. Robles-Rubio, K. A. Brown, L. Kanbar, J. Kaczmarek, S. Chawla, G. M. Sant'anna, and R. E. Kearney, "Prediction of Extubation Readiness in Extreme Preterm Infants Based on Measures of Cardiorespiratory Variability," in *Conf. Proc. 34th IEEE Eng. Med. Biol. Soc.*, San Diego, USA, 2012, pp. 5630-5633.

[4] C. F. Poets and V. A. Stebbens, "Detection of movement artifact in recorded pulse oximeter saturation," *Eur. J. Pediatr.,* vol. 156, pp. 808-811, 1997.

[5] G. Lu and F. Yang, "Limitations of Oximetry to Measure Heart Rate Variability Measures," *Cardiovasc. Eng.,* vol. 9, pp. 119-125, 2009.

[6] A. Blackman, C. Mcgregor, R. Dales, H. S. Driver, L. Dumov, J. Fleming, K. Fraser, C. George, A. Khullar, J. Mink, M. Moffat, G. E. Sullivan, J. A. Fleetham, N. Ayas, T. D. Bradley, M. Fitzpatrick, J. Kimoff, D. Morrison, F. Ryan, R. Skomro, F. Series, and W. Tsai, "Canadian Sleep Society/Canadian Thoracic Society position paper on the use of portable monitoring for the diagnosis of obstructive sleep apnea/hypopnea in adults," *Can. Respir. J.,* vol. 17, pp. 229-232, 2010.

[7] S. J. Barker, ""Motion-Resistant" Pulse Oximetry: A Comparison of New and Old Models," *Anesth. Analg.,* vol. 95, pp. 967-972, 2002.

[8] R. Krishnan, B. Natarajan, and S. Warren, "Two-Stage Approach for Detection and Reduction of Motion Artifacts in Photoplethysmographic Data," *IEEE Trans. Biomed. Eng.,* vol. 57, pp. 1867-1876, 2010.

[9] N. Selvaraj, Y. Mendelson, K. H. Shelley, D. G. Silverman, and K. H. Chon, "Statistical approach for the detection of motion/noise artifacts in Photoplethysmogram," in *Conf. Proc. 33rd IEEE Eng. Med. Biol. Soc.*, 2011, pp. 4972-4975.

[10] C. E. Shannon, "A mathematical theory of communication," *SIGMOBILE Mob. Comput. Commun. Rev.,* vol. 5, pp. 3-55, 2001.

[11] C. O. Fetus and Newborn, "Age Terminology During the Perinatal Period," *Pediatrics,* vol. 114, pp. 1362-1364, 2004.

[12] C. A. Robles-Rubio, K. A. Brown, and R. E. Kearney, "Automated Unsupervised Respiratory Event Analysis," in *Conf. Proc. 33rd IEEE Eng. Med. Biol. Soc.*, Boston, USA, 2011, pp. 3201-3204.

[13] M. Stoyanov, M. Gunzburger, and J. Burkardt, "Pink Noise, 1/f^alpha Noise, and Their Effect on Solutions of Differential Equations," *Int. J. Uncertainty Quantification,* vol. 1, pp. 257-278, 2011.

[14] 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, pp. 179-183, 1997.

[15] T. Kok-Swang, I. Hinberg, and J. Wadhwani, "Electromagnetic interference in medical devices: Health Canada's past and current perspectives and activities," in *IEEE Int. Symp. Electromagn. C.*, 2001, pp. 1283-1288 vol.2.

[16] J. M. Graybeal and M. T. Petterson, "Adaptive filtering and alternative calculations revolutionizes pulse oximetry sensitivity and specificity during motion and low perfusion," in *Conf. Proc. 26th IEEE Eng. Med. Biol. Soc.*, 2004, pp. 5363-5366.

[17] K. A. Reddy, B. George, and V. J. Kumar, "Use of Fourier Series Analysis for Motion Artifact Reduction and Data Compression of Photoplethysmographic Signals," *IEEE Trans. Instrum. Meas.,* vol. 58, pp. 1706-1711, 2009.

[18] B. S. Kim and S. K. Yoo, "Motion artifact reduction in photoplethysmography using independent component analysis," *IEEE Trans. Biomed. Eng.,* vol. 53, pp. 566-568, 2006.