

Multiparameter Respiratory Rate Estimation from the Photoplethysmogram

Walter Karlen, *Member, IEEE*, Srinivas Raman, J. Mark Ansermino, and Guy A. Dumont, *Fellow, IEEE*

Abstract—We present a novel method for estimating respiratory rate in real-time from the photoplethysmogram (PPG) obtained from pulse oximetry. Three respiratory induced variations (frequency, intensity, and amplitude) are extracted from the PPG using the Incremental-Merge Segmentation algorithm. Frequency content of each respiratory induced variation is analyzed using Fast Fourier Transforms. The proposed Smart Fusion method then combines the results of the three respiratory induced variations using a transparent mean calculation. It automatically eliminates estimations considered to be unreliable because of detected presence of artifacts in the PPG or disagreement between the different individual respiratory rate estimations.

The algorithm has been tested on data obtained from 29 children and 13 adults. Results show that it is important to combine the three respiratory induced variations for robust estimation of respiratory rate. The Smart Fusion showed trends of improved estimation (mean root mean square error 3.0 breaths/min) compared to the individual estimation methods (5.8, 6.2 and 3.9 breaths/min).

The Smart Fusion algorithm is being implemented in a mobile phone pulse oximeter device to facilitate the diagnosis of severe childhood pneumonia in remote areas.

Keywords—respiratory rate, pulse oximeter, photoplethysmogram, data fusion.

I. INTRODUCTION

PNEUMONIA kills more than 2 million children under five years old every year [1]. Almost all of these deaths occur in the developing world. Many of these deaths could be prevented by early detection and timely administration of simple treatments. While the integration of technology into health care has greatly improved the speed and accuracy of diagnosis and treatment of childhood pneumonia in the developed world, lack of access to clinical expertise and costly tests often delay diagnosis and treatment, and reduce survival rates across the developing world. An essential criteria integrated in many guidelines for diagnosis of pneumonia in ill children is the assessment of an elevated respiratory rate of ≥ 40 breaths/min (age 1-5 years) [2]. However, clinical measurement of respiratory rate has been shown to have poor reliability and repeatability [3].

Copyright ©2013 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending an email to pubs-permissions@ieee.org.

This work was supported by the Swiss National Science Foundation (SNSF) and the Canadian Institutes for Health Research (CIHR).

W. Karlen and G.A. Dumont are with the Department of Electrical and Computer Engineering, The University of British Columbia (UBC), 2332 Main Mall, Vancouver BC, V6T 1Z4, Canada (tel: +1 604 875 2000 x3646; e-mail: walter.karlen@ieee.org).

S. Raman and J.M. Ansermino are with the Department of Anesthesiology, Pharmacology & Therapeutics, University of British Columbia (UBC), Vancouver BC, V6T 1Z4, Canada.

An easy-to-use, inexpensive diagnostic device that can swiftly and accurately identify children with severe pneumonia would enable the timely administration of appropriate treatments. Prompt and effective application of life-saving antibiotics or oxygen therapy would, in turn, minimize wastage of scarce and costly resources. Pulse oximeters offer the possibility of achieving an accurate diagnosis of pneumonia.

Pulse oximeters apply the Beer-Lambert law to estimate hemoglobin oxygenation (SpO_2). The law describes that light intensity diminishes exponentially when traveling in an absorbing medium and the absorption is dependent on the wavelength. Oxygenated hemoglobin preferentially absorbs infrared light and transmits red light and deoxygenated hemoglobin behaves in the opposite manner. In a pulse oximeter, two light emitting diodes (LEDs) emitting red (660 nm) and infra-red (940 nm) light are used to actively illuminate the patient's tissue (usually at the finger tip) alternately. The intensity of the non-absorbed light from each LED is measured with a receiver photodiode. The ratio of the transmitted infra-red and red intensity is empirically related to the SpO_2 . In addition to blood composition, the light absorption and transmission depend on the traveled light path, optical density of the tissue and volume of blood present in the tissue [4]. This permits the display of the variation of blood volume in the finger over time with a photoplethysmogram (PPG), which has a pulsatile component and a constant component. From the pulsatile component of the PPG, heart rate (HR) can be easily deduced. In addition to SpO_2 and HR, respiratory rate (RR) is another important parameter that can potentially be estimated from the PPG [5]. All three parameters are strong predictors of critical illness in pneumonia that can differentiate it from mild respiratory tract infections [6]. However, commercial pulse oximeters are currently limited to measuring HR and SpO_2 and require bulky, expensive equipment to process and monitor the data, which restricts their widespread adoption in the developing world.

The aim of our research is to develop a fully featured, low cost pulse oximeter that operates on mobile phones and can be used as a multi-functional screening tool to improve diagnosis and treatment of children with severe pneumonia. We have previously demonstrated the *Phone Oximeter* [7], a commercial and Federal Drug Administration (FDA) approved pulse oximeter connecting to a smart phone. The *Phone Oximeter* was specifically designed for continuous monitoring and not spot-check applications, and only provided experimental RR measurements.

In this work we present an efficient processing algorithm for robust RR estimation with the aim of developing a com-

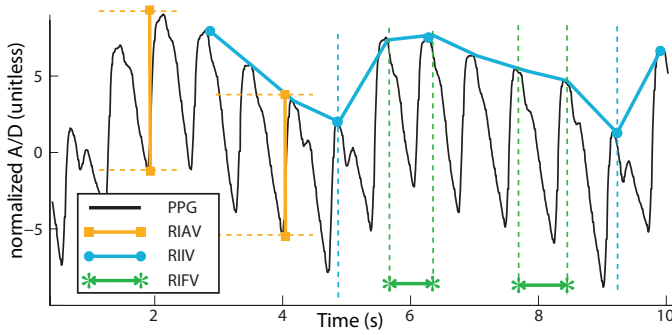


Fig. 1. Respiratory induced variations in the PPG. Respiratory induced amplitude variation (RIAV) is the change in pulse strength. Respiratory induced intensity variation (RIIV) is the variation of perfusion baseline. Respiratory induced frequency variation (RIFV) is a synchronization of the heart beat with the RR.

prehensive and portable pulse oximeter monitor to aid in the diagnosis of childhood pneumonia. This algorithm can be implemented into a novel spot-check version of the *Phone Oximeter*.

A. Background

RR modulates the PPG waveform in three ways (Figure 1): 1) An autonomic response to respiration causes the variation of HR to synchronize with the respiratory cycle. Also referred to as respiratory sinus arrhythmia (RSA), the HR increases during inspiration and decreases during expiration, causing respiratory induced frequency variation (RIFV) of the PPG waveform. 2) The intrathoracic pressure variation causes exchange of blood between the pulmonary circulation and the systemic circulation. This results in a variation of perfusion baseline, termed the respiratory induced intensity variation (RIIV). 3) A corresponding decrease in cardiac output due to reduced ventricular filling [8], causes the respiratory induced amplitude variation (RIAV), which is a change in peripheral pulse strength.

Different algorithms have been proposed for RR estimation from PPG since it was first suggested in the early 1990's [9]. Wavelet decomposition [10], digital filters [11], Fourier transforms [12], complex demodulation [13], and auto-regression [14] methods have been successfully used.

A recent study evaluated the correlation of each respiration-induced waveform variation with the respiratory cycle [15]. The results showed that there is no optimal variation, as the correlation performance is dependent on many factors including gender, respiratory rate and body position. In addition to the controlled variables, extraneous variables that mask the respiratory effect of the PPG signal can be caused by measurement artifacts and physiological variability between patients. For example, in dehydrated patients the amplitude variation may be stronger. The RSA may be negligible in chronic disease conditions (e.g. diabetes and cardiovascular diseases [16]). Other RR algorithms have also exploited the potential of extracting more than one waveform variation before the RR is determined. Johansson [17] combined RIIV, RIFV, and RIAV measurements using neural networks and Leonard et al. [10] compared the RIFV and RIAV before selecting the best

waveform for a final algorithm design. In [18], continuous wavelet decompositions were used for estimating RR for each of the three variations and a proprietary algorithm was used for merging and displaying this information.

Despite these advances, RR measurements have only recently been introduced in commercial pulse oximeters. This is partly due to the fact that more reliable RR estimation methods, such as spirometry or capnometry, are available in clinical environments where pulse oximeters are commonly used. However, in remote or resource-poor locations where no other RR measurement technique is readily available, RR estimation from PPG may offer the best practical solution. Nevertheless, current methods are computationally expensive (up to 4 s on a PC [13]) or do not allow on-line implementation. In addition, the methods do not have signal quality indices to indicate when they can be reliably used. Pulse oximetry is prone to artifacts. Further, in low perfused subjects, a reliable PPG signal is difficult to obtain. The absence of a quality indicator could lead to significant clinical error. To overcome this limitation, we have designed an algorithm that could achieve clinically significant accuracy and reliable RR readings using all the available respiration-induced waveform variations while still being optimized for power efficiency for on-line use on a mobile device or a microcontroller.

II. ALGORITHM DESCRIPTION

To perform automated multi-parameter RR estimation, two periodic components have to be extracted from the PPG signal: a) regular heart beat pulsations characterized by a maximal volume peak for the extraction of RSA and RIAV; and b) a lower frequency pattern belonging to the RIIV characterized by a sinusoidal variation of the signal baseline (Figure 1). To extract these two features, we process the waveform with a segmentation algorithm that splits the PPG into pulses. Since the PPG pulses are composed of morphological shapes that can be characterized by consecutive lines, this algorithm permits the desired trend calculation in an efficient manner. The respiratory frequency information is then extracted using a maximal spectral power approach for each parameter. Artifact identification and cancellation is essential in PPG processing. Our algorithm performs artifact detection within the segmentation processing steps. RR estimation is obtained by fusing the respiratory frequency from RIFV, RIIV, and RIAV (Figure 2). The quality of this estimation is evaluated by comparing the three components and taking the presence of artifacts into account. If the assessed quality is low, the RR estimation is not provided. In the following subsections we describe the RR estimation in more detail.

A. Preprocessing

A high-pass filter is applied to remove the DC component of the PPG signal. This step is usually performed before the PPG becomes available from the pulse oximeter and the cut-off frequency is controlled by the pulse oximeter manufacturer. Despite the DC filtering, most PPG signals from commercial pulse oximeters have the RIIV variation preserved.

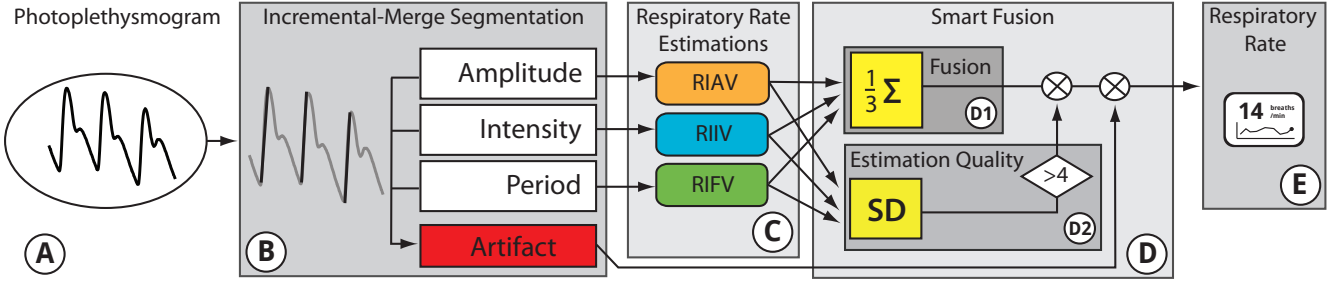


Fig. 2. Smart Fusion respiratory rate estimation algorithm. The photoplethysmogram (A) is processed with the Incremental-Merge Segmentation (B) algorithm. Amplitude, intensity and period features from pulse segments are used for artifact detection and calculating respiratory rate estimates (C) from the respiratory induced frequency (RIFV), amplitude (RIAV), and intensity (RIIV) variation. During the Smart Fusion process (D), RIFV, RIAV, and RIIV are combined by calculating the mean (D1) and only reliable estimations are selected. Estimations containing artifacts or low estimation quality (D2; standard deviations (SD) > 4) are eliminated and only robust respiratory rates (E) are displayed.

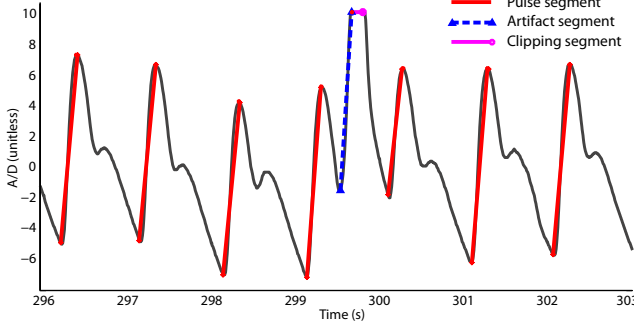


Fig. 3. Example output of the Incremental-Merge segmentation (IMS) algorithm. Up-slopes adjacent to clipping are automatically labeled as artifacts.

B. Pulse Segmentation

The PPG is segmented into pulses using an Incremental-Merge Segmentation (IMS) method [19] which is a mixture of Iterative-End-Point-Fit and Incremental algorithms. The Incremental-Merge algorithm has a sliding-window structure which is simple, fast and can be computed in real-time. The tuning of the algorithm requires the setting of only one parameter m (length of line segments in number of sample points minus one), which is dependent on the sampling rate. After segmentation, each PPG pulse is represented as a straight line originating at pulse start and ending at its primary peak (Figure 3). The length of this line corresponds to the pulse amplitude and is proportional to the peripheral perfusion of the patient. We have previously shown that this method is an efficient and accurate way to detect pulse peaks with high sensitivity (98.93%) and positive predictive value (96.68%) [19].

The IMS processing produces the pulse amplitude, maximum and minimum intensity of each pulse, and pulse period. All these PPG features are used for artifact detection and RR estimation.

C. Artifact Detection

Artifacts are assessed on-line within the IMS processing. Artifacts are identified by scanning for abnormal pulse periods. Data segments containing pulse periods outside the normal range (230 to 2400 ms) are automatically labeled as artifacts. In addition, the IMS detects artifacts by adaptively scanning

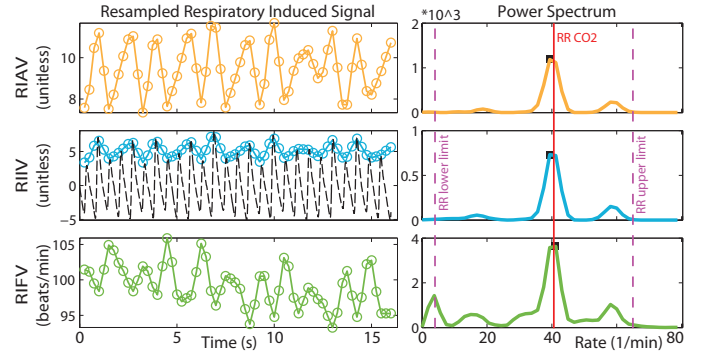


Fig. 4. Example of a 16 s PPG sliding window used for RR estimation. Respiratory induced amplitude variation (RIAV, top), respiratory induced intensity variation (RIIV, middle with PPG) and respiratory induced frequency variation (RIFV, bottom) are extracted and resampled to 4 Hz (left column). Using FFT, the power spectrum is calculated for each variation (right column) and the maximum power (black squares) is selected within a physiologically expected RR range (dashed vertical lines). The reference RR obtained from the CO₂ signal is represented by a vertical line.

for abnormal large and clipped pulses (Figure 3) [19]. This method is used to primarily eliminate motion artifacts given that individual abnormal pulses more easily identify motion artifacts than noise artifacts. The Smart Fusion method is designed to address noise artifacts by combining different features (that have different frequency modulations) from the PPG waveform as subsequently described.

D. RR Estimation using Smart Fusion

1) *RIFV*: RIFV is calculated using a Fast Fourier Transform (FFT) on evenly sampled beat intervals (tachograms). This is an improved procedure compared to that developed in [20]. Pulse periods are converted into tachograms. Fourier analysis requires evenly sampled data and, therefore, the tachogram is resampled onto an even 4 Hz grid using Berger's algorithm [21]. Berger's algorithm is computationally efficient, highly localized and ideal for real-time processing. The tachograms are then transformed to the frequency domain in pseudo real-time. Data are divided into sliding windows with power of two sample points (duration determined by experiments), which is optimal for computing the FFT. Each window is shifted by 1 s to simulate real-time analysis. A Hamming window is applied to minimize the first side lobe of the frequency response, and

the tachograms are then converted to the frequency domain using FFT. The resulting power spectrum of each window is then analyzed for the frequency with maximum power within the expected respiratory frequency range (0.067 to 1.08 Hz or 4 to 65 breaths/min, Figure 4).

2) *RIIV*: The maximum intensity of the PPG pulses is used for extracting the RIIV. Intensity trend data is resampled onto an even 4 Hz grid using linear interpolation. Equivalent to RIFV, data are divided into sliding windows with one-second shifts and multiplied with a Hamming window. The FFT power spectrum of each window is then analyzed for the maximum frequency content within the RR frequency range.

3) *RIAV*: Amplitude trend data is resampled onto an even 4 Hz grid using linear interpolation. Equivalent to RIFV and RIIV, data are divided into sliding windows with one-second shifts and multiplied with a Hamming window. The FFT power spectrum of each window is then analyzed for the maximum frequency content within the RR frequency range.

4) *RR estimation quality*: The RR estimation is analyzed for consistency. Standard deviations (SD) from the three RR estimations (RIAV, RIIV, and RIFV) are calculated continuously. SD greater than 4 breaths/min are considered to be aberrant and the corresponding window is labeled with a low RR estimation quality.

5) *RR Fusion*: RIAV, RIIV, and RIFV RR estimations are fused by calculating their mean.

6) *Smart Fusion*: The fusion process is enhanced by eliminating untrustworthy data windows that are labeled as artifact or as low RR estimation quality (Figure 2).

III. METHODS

A. Data Collection

Following institutional review board approval¹, we collected physiological data to calibrate the algorithms and evaluate the RR estimation performance. For this study, cases were randomly selected from a larger collection of physiological signals collected during elective surgery and routine anesthesia for the purpose of development of improved monitoring algorithms in adults and children. Vital signs from 59 children (median age: 8.7, range 0.8–16.5 years) and 35 adults (median age: 52.4, range 26.2–75.6 years) were selected. Pulse oximetry is a standard monitoring tool during general anesthesia and for the recording of vital signs no change in the monitoring equipment or procedures were required. The recordings obtained included capnometry (25 Hz) and PPG (100 Hz) signals. All signals were recorded with S/5 Collect software (Datex-Ohmeda, Finland) using a sampling frequency of 300 Hz (PPG and capnometry with lower sampling rates were automatically up-sampled by the software during export).

B. Data Preparation

Forty-two 8-min segments from 29 pediatric and 13 adults cases containing reliable recordings of spontaneous or controlled breathing were randomly selected from each case for

the Test Data Set. One hundred twenty-four 2-min segments were randomly selected from the remaining 52 cases for the Calibration Data Set. Both data sets are available for download from the on-line database CapnoBase.org [22]. The capnometric waveform was used as the reference gold standard recording for RR validation. A research assistant independently validated the reference recordings using the CapnoBase Signal Evaluation Tool [22]. The research assistant also labeled the the beginning and end of all potential artifacts in the PPG waveforms.

All the data processing and statistical analysis was performed using the Matlab software framework (Mathworks, Natick, USA).

C. Algorithm Evaluation

The RR algorithm's performance was assessed using the unnormalized root mean square (RMS) error (breaths/min) defined as

$$RMS\ error = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i^{ref} - x_i^{alg})^2} \quad (1)$$

where n is the number of observations and x^{ref} and x^{alg} are the reference and the algorithm observations respectively. A measurement error was calculated for each algorithm observation by comparing it with the reference observation that was nearest to it in time. The first 64 s were not used for the performance measurement since they were used to initialize the high-pass filters and the sliding window. Measurement errors for windows containing artifacts detected automatically by the algorithm were ignored for all RR estimation methods, including the single modulation methods. This allowed a more objective comparison between the different estimation methods. The changes in RIIV and RIAV are different for different modes of ventilation. Subjects receiving general anesthesia are frequently mechanically ventilated. The target population subject to pneumonia screening would be breathing spontaneously. Therefore, our analysis looks at ventilatory modes separately.

The distribution of RMS error was displayed using boxplots. To further illustrate the relative performance of each respiratory modulation, the most accurate method and least accurate method, as defined by the RMS error, were identified in the Test and Calibration Data Sets. The best performing method was assigned *High Score* and worst performing method was assigned *Low Score* for each data segment (2-min or 8-min recording respectively). Therefore, the *High Score* RR estimations illustrate the hypothetical maximum achievable performance with an ideal fusion process and the *Low Score* RR estimations illustrate the worst case performance of the individual methods. In addition, the overall contribution of each modulation to the *High Score* or *Low Score* categories in each data set was calculated. The *High Score* and *Low Score* selection was done using a-priori RR information from the reference CO₂ signal and was solely used for illustration purposes.

¹ID H06-70205 and H03-70451, University of British Columbia Research Ethics Board

D. Algorithm Calibration

The length of the line segment m was determined empirically using the Calibration Data Set. Choosing a smaller window size for the FFT computation provides higher time resolution and lower computational costs, and therefore better real-time performance; larger window sizes provide better accuracy. To determine the optimal window size, we repeated a RR estimation experiment by varying the FFT window size between 16, 32, and 64 s for each RR estimation method. These window sizes were selected as they did not need zero padding and were within reasonable physiological and clinical limits. The window size with the best performance and time delay was then chosen for algorithm testing.

E. Algorithm Testing

The algorithm configuration obtained from the calibration experiment was applied to the Test Data Set. RR accuracy and performance distribution over different RR estimation methods were calculated and displayed as boxplots. In addition, a scatter plot for all data points combined was drawn to illustrate the relationship between reference RR, Fusion and Smart Fusion algorithm output respectively.

All RMS errors were tested for Gaussian distribution using the Lilliefors test at a significance level of $p < 0.05$. A balanced analysis of variance (ANOVA) with fixed-effect was performed on the errors from RIFV, RIIV, RIAV, and Smart Fusion. Results from F-test and Tukey post-hoc comparison (significance levels of $p = 0.05$) were reported.

IV. RESULTS

A. Calibration

The calibration experiments showed that a large range of segment sizes m can be used to setup the IMS algorithm. For the sampling rate of 300 Hz used in this study, $m = 10$ was optimal. It offered a good compromise between computational load and time resolution for the pulse peak detection. Results on the accuracy of the beat and artifact detection using the IMS algorithm were presented in [19].

The different FFT window sizes used for estimating individual RR components did not show a significant difference in error (Figure 5). A positive trend for a lower error rate in larger windows was observed. However, larger windows delayed the real-time response of the algorithm. Further, larger windows increased the processing load at each time step. From a computational point of view, choosing the smaller window size was optimal. However, as the 16 s window limits the lowest detectable RR to 7.5 breaths/min, the 32 s window was chosen for further processing so that rates as low as 3.75 breaths/min can be detected.

B. Algorithm Testing

The ANOVA of the Test Data Set showed that there were differences between the RR estimation algorithms ($F=3.54$) which were significant at $p=0.02$. The RR estimation algorithms based on a single respiratory variation showed comparable RMS error performance (Figure 6). There was no

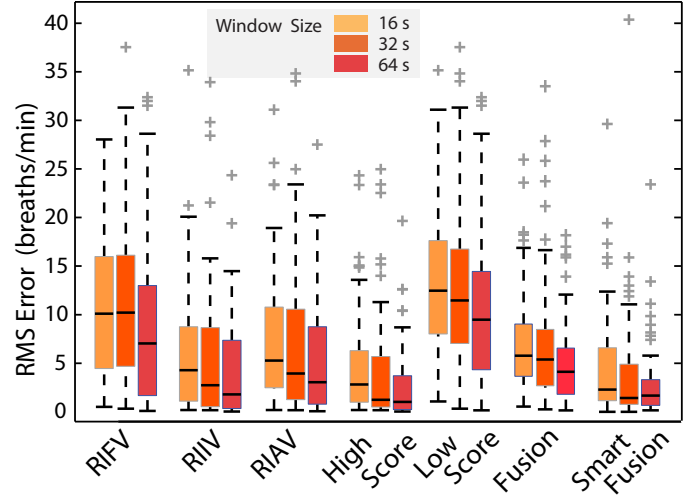


Fig. 5. Results from the sensitivity analysis for the selection of the best window size for the FFT processing using the Calibration Data Set. The boxplot show the RMS Error. Lower quartile, median, and upper quartile values were displayed as bottom, middle and top horizontal line of the boxes. Whiskers were used to represent the most extreme values within 1.5 times the interquartile range from the quartile. Outliers (data with values beyond the ends of the whiskers) were displayed as crosses. No significant difference between the three analyzed window sizes can be observed.

single variation method that was statistically significantly different from the others using the Tukey multiple comparisons (RIFV vs. RIIV: $p=0.2$; RIFV vs. RIAV: $p=0.98$; RIAV vs. RIIV: $p=0.4$). The Smart Fusion method incorporating the additional quality estimation outperformed the RIFV algorithm ($p=0.026$) and did not reach statistical significance compared to RIAV ($p=0.07$) and RIIV ($p=0.82$). There was a trend towards lower RMS errors for the Smart Fusion compared to the three variability measures in the sample tested (Figure 6). This may have reached statistical significance if a larger sample was used.

The Smart Fusion identified and eliminated a majority of high error estimations (Figure 7). The largest estimation errors that could not be resolved with the Smart Fusion were close to the minimal detectable RR (Figure 7 box A). These were caused by low frequency, non-respiratory processes whose spectrum was reaching into the expected respiratory frequency range that was defined for the RR estimation algorithm.

All three single respiratory variation methods contributed to the *High Score* and the *Low Score* group (Table I). The RIIV method produced the least *Low Score* estimations (14%) and also the most *High Score* estimations (52.4%) on the Test Data Set. The mean RMS error (± 1 SD) of the Smart Fusion (3 ± 4.7 breaths/min) was close to the *High Score* (2.8 ± 3.4 breaths/min). No statistical difference in RR estimation accuracy between ventilation modes (spontaneous breathing and mechanical ventilation) can be observed (Figure 6). There is indication that in case of mechanical ventilation, the use of RIIV is more favorable than the use of RIFV (Figure 6 and Table I).

V. DISCUSSION

We have presented a novel method for estimating RR in real-time from the PPG and validated the accuracy of the

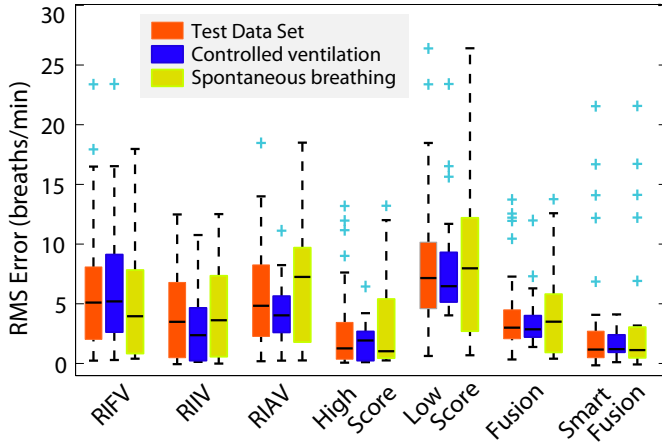


Fig. 6. Results for the Test Data Set using 32 s FFT windows. The boxplots give the RMS Error for the different respiratory rate estimation methods. Results are also displayed for the Test Data Set split up into ventilatory modes.

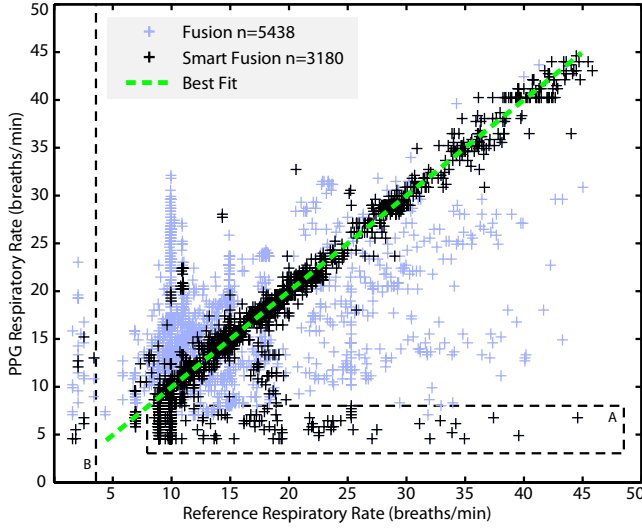


Fig. 7. Scatter plot comparing the reference respiratory rate obtained from capnometry with the PPG respiratory rate obtained from the Fusion and Smart Fusion algorithm. RR estimations during PPG artifacts are not shown. The Smart Fusion eliminates the estimations with large error (distance from Best Fit). Box A: Non-respiratory processes with low frequencies (<7 cycles/min) can overlap with the respiratory spectrum. The Smart Fusion is not able to completely eliminate the misidentification in this range. Line B: Lowest detectable RR (3.75 breaths/min) given by the selected FFT window (32 s).

method with capnometry as a comparative gold standard RR measurement. The Smart Fusion method is based on a segmentation algorithm (IMS) for pulse feature extraction and artifact detection and a RR estimation using the fusion of three different components, RIAV, RIFV, and RIIV. The IMS algorithm operates solely in the time-domain and is computationally efficient and robust to PPG artifacts [19]. The fusion of RR estimates is provided by the mean of the three RR estimations for 32 s windows that are not automatically labeled as artifact or as low RR estimation quality.

Our results demonstrate that it is important to combine the three respiratory induced variations for robust estimation of RR. The combination (with and without smart processing) showed improved estimation compared to the individual estimation methods. As shown in Table I, each variation has a

TABLE I
DISTRIBUTION OF RR ESTIMATION METHODS CONTRIBUTING TO THE High Score OR Low Score GROUPS IN NUMBER OF CASES (32 s WINDOW)

Data Set		Method		
		RIFV	RIIV	RIAV
Calibration	High Score	16	49	59
	Low Score	57	18	49
Test	High Score	12	23	7
	Low Score	20	7	15
Controlled ventilation	High Score	6	11	2
	Low Score	11	4	4
Spontaneous breathing	High Score	6	12	5
	Low Score	9	3	11

significant contribution to the best and worst performances. This is consistent with findings of [15] that illustrated the relative performance heterogeneity of each respiratory induced variation. It can be expected that some respiratory induced variations change behavior under different ventilatory condition. For example, RIIV is impacted by RR and respiratory volume [23]. It was not within the scope of the presented work to analyze the relationship between RR and the selection of estimation method and this will require further investigation.

We attempted to include a broad range of subjects into this study to generalize our findings. The data set included children and adults, under controlled ventilation or spontaneously breathing over a wide RR range. This study was limited to subjects receiving general anesthesia and that were continuously monitored. We did not observe a difference in RR estimation accuracy between ventilatory modes (Figure 6), which is consistent with previous findings [24], [20]. Although the performance of the Smart Fusion algorithm is promising in the studied population, further validation on a larger data set and the target population suffering from pneumonia will be required. For instance, our data sets contained only a limited number of cases with subjects breathing at abnormally high RR (> 30 breaths/min) that would be typical in childhood pneumonia. We are currently collecting pulse oximetry data from children admitted to the Acute Care unit at Mulago Hospital, Kampala, Uganda to extend our findings.

Low frequency processes other than respiration (e.g. Mayer waves [25] or artifacts) can modulate the PPG and become a confounding factor for accurately assessing RR. In our case, low frequency variations below 7 cycles/min contributed largely to the error in the RR estimation (Figure 7 box A). These variations were unanimously detected by the three respiratory induced variations and consequently, the Smart Fusion was not able to eliminate these errors. Major contributors were low frequency variations due to artifacts that remained undetected by the IMS algorithm. These variations appeared in the spectrogram with large power peaks that reached into the respiratory frequency range for short periods of time. In future versions of the Smart Fusion algorithm we will eliminate this behavior using an additional high-pass filter before the FFT. In applications where low RRs are unlikely, the lower limit of 4 breaths/min for the RR estimation algorithm could also be increased which would reduce the low frequency error.

The selected FFT window length of 32 s limited the lowest detectable RR to 3.75 breaths/min. Consequently, detection of lower rates was poor (Figure 7 line B). Such low rates can be considered apneic events and our algorithm is currently not designed to detect such events. For the detection of these events, we recommend the introduction of a signal processing method operating in the time domain that could overcome this FFT limitation.

For generating a single RR estimation from three sources, we have taken a simple, but efficient approach in using a mean of the three available values. This method is transparent and clearly understandable for users. Alternatively, more sophisticated approaches using time dependency like the hybrid median filter [26] could be used. Probabilities for the selection of the most accurate methods, as used in Kalman filters, also have the potential to combine information from different respiratory sources [27]. Machine learning approaches like neural network models have also been used to merge RR information [10]. However, these black box approaches are not transparent to the user who has to make decisions and diagnoses based on the presented information. In addition, due to the heterogeneous nature of the data sets used for generating the models, it is very unlikely to obtain a generalizable model for clinical practice. Heuristic selection of best RR sources, as has been suggested in [12] for selection between RIAV and RIFV, currently lacks objective, consistent and reliable selection methods for real-time use as the underlying principles are still not fully understood.

The Smart Fusion method efficiently eliminated RR estimations with large errors using estimation quality information. While this method largely improved the robustness of the estimation, it also introduced a major limitation. A RR estimation is only available during periods that do not contain artifacts or have an agreement between the three estimation methods. In particular, certain conditions (e.g. the influence of drugs) could alter a single respiratory induced variation and cause a permanent disagreement between the three estimation methods. In the Test Data Set, 7% of the windows were eliminated due to artifacts and an additional 38.5% due to low estimation quality. In practice, this 45.5% dropout can cause large time periods where no RR is available and the user has to extend the recording session to obtain a reliable value. Alternatively, a low confidence RR could be displayed, but supported with a measure of confidence.

This was not the first time all three respiratory induced variations have been used to estimate RR from PPG. For example, in [18], a continuous wavelet decomposition was used for estimation of each individual variation. However, it was not reported how the information from the three sources was combined into a single RR estimation. The individual contribution of each source to the final estimation and the benefit of combining these sources was not shown. In particular, the reported RR in [18] were averaged over a variable time frame (up to 2 min) using an undisclosed algorithm. This post-processing methodology diluted the effective performance of the estimation and introduces large real-time delay. This delay is unknown to the user and made comparison with our work impossible. We did not apply post-processing methods such as

averaging as we believe it is optimal to identify the source of error. We have also published raw data sources and algorithm outcomes on the Internet (CapnoBase.org), in the hope that other research groups can use this data set as a benchmark for comparison of new RR estimation algorithms.

The Smart Fusion algorithm can produce a RR estimation within 32 s. This real-time delay is required for the FFT calculation that uses a 32 s window size. It compares favorably to the delay of other suggested methods that require the analysis of data windows of 1 min [13] and up to 2 min with filtering [18]. For spot-check applications, such as the planned mobile phone pneumonia diagnosis, where a sensor reading has to be obtained quickly, a short recording window is beneficial.

The IMS and Smart Fusion are computationally efficient. The average processing time for 60 RR estimations (1 min and 32 s recording time) was 0.36 s on a single processor thread on a 1.8 GHz PC. The computational load of the presented algorithm could be further reduced by optimizing the FFT processes. The number of FFTs to be calculated is proportional to the update frequency of the trend estimations. Therefore, increasing the RR update rate from 1 s to an acceptable 2 s interval will almost half the processing effort.

We have presented the development of the Smart Fusion algorithm for on-line RR estimation from PPG. The algorithm combines the components of three respiratory induced variations (RIFV, RIIV, and RIAV) using a transparent mean calculation and eliminates estimations considered to be unreliable because of disagreement between the different individual RR estimations or detected presence of artifacts in the PPG.

We are working on a version of the IMS algorithm and the Smart Fusion RR estimation to be implemented on the Phone Oximeter [7]. The Phone Oximeter will provide the user with HR, RR and SpO₂ readings and facilitate the diagnosis of pneumonia in rural areas where high-end clinical devices such as X-ray machines are sparse or not affordable. It will provide a low-cost and multi-functional screening tool to improve treatment and increase survival of children with severe pneumonia.

ACKNOWLEDGMENTS

The authors would like to thank Joanne Lim for kindly helping to revise this manuscript and Erin Cooke for preparing the data sets.

REFERENCES

- [1] T. M. Wardlaw, E. W. Johansson, and M. Hodge, *Pneumonia: the forgotten killer of children*. New York / Geneva: The United Nations Children's Fund (UNICEF)/World Health Organization (WHO), 2006.
- [2] WHO, *Pocket Book of Hospital Care for Children - Guidelines for the Management of Common Illnesses with Limited Resources*, Geneva, CH, 2005.
- [3] P. B. Lovett, J. M. Buchwald, K. Stürmann, and P. Bijur, "The vexatious vital: neither clinical measurements by nurses nor an electronic monitor provides accurate measurements of respiratory rate in triage." *Ann Emerg Med*, vol. 45, no. 1, pp. 68–76, 2005.
- [4] P. D. Mannheimer, "The light-tissue interaction of pulse oximetry." *Anesth Analg*, vol. 105, no. 6 Suppl, pp. S10–7, 2007.
- [5] K. H. Shelley, "Photoplethysmography: beyond the calculation of arterial oxygen saturation and heart rate." *Anesth Analg*, vol. 105, no. 6 Suppl, pp. S31–6, 2007.

- [6] T. Duke, R. Subhi, D. Peel, and B. Frey, "Pulse oximetry: technology to reduce child mortality in developing countries." *Ann Trop Paediatr*, vol. 29, no. 3, pp. 165–75, 2009.
- [7] W. Karlen, G. A. Dumont, C. Petersen, J. Gow, J. Lim, J. Sleiman, and J. M. Ansermino, "Human-centered Phone Oximeter Interface Design for the Operating Room," in *HEALTHINF 2011 - Proceedings of the International Conference on Health Informatics*, V. Traver, A. Fred, J. Filipe, and H. Gamboa, Eds. Rome: SciTePress, 2011, pp. 433–8.
- [8] A. J. Buda, M. R. Pinsky, N. B. Ingels, G. T. Daughters, E. B. Stinson, and E. L. Alderman, "Effect of intrathoracic pressure on left ventricular performance." *New Engl J Med*, vol. 301, no. 9, pp. 453–9, 1979.
- [9] K. Nakajima, T. Tamura, T. Ohta, H. Miike, and P. Oberg, "Photoplethysmographic measurement of heart and respiratory rates using digital filters," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 1993, pp. 1006–7.
- [10] P. Leonard, 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 Monitor Comp*, vol. 18, no. 5-6, pp. 309–12, 2004.
- [11] K. Nakajima, T. Tamura, and H. Miike, "Monitoring of heart and respiratory rates by photoplethysmography using a digital filtering technique," *Med Eng Phys*, vol. 18, no. 5, pp. 365–72, 1996.
- [12] K. H. Shelley, A. A. Awad, R. G. Stout, and D. G. Silverman, "The use of joint time frequency analysis to quantify the effect of ventilation on the pulse oximeter waveform." *J Clin Monitor Comp*, vol. 20, no. 2, pp. 81–7, 2006.
- [13] 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, 2009.
- [14] S. Fleming and L. Tarassenko, "A comparison of signal processing techniques for the extraction of breathing rate from the photoplethysmogram," *Int J Biol Med Sci*, vol. 2, no. 4, pp. 232–6, 2007.
- [15] J. Li, J. Jin, X. Chen, W. Sun, and P. Guo, "Comparison of respiratory-induced variations in photoplethysmographic signals." *Physiol Meas*, vol. 31, no. 3, pp. 415–25, 2010.
- [16] C. M. Masi, L. C. Hawkey, E. M. Rickett, and J. T. Cacioppo, "Respiratory sinus arrhythmia and diseases of aging: obesity, diabetes mellitus, and hypertension." *Biol Psychol*, vol. 74, no. 2, pp. 212–23, 2007.
- [17] A. Johansson, "Neural network for photoplethysmographic respiratory rate monitoring," *Med Biol Eng Comp*, vol. 41, no. 3, pp. 242–8, 2003.
- [18] P. S. Addison, J. N. Watson, M. L. Mestek, and R. S. Mecca, "Developing an algorithm for pulse oximetry derived respiratory rate (RR(oxi)): a healthy volunteer study." *J Clin Monitor Comp*, vol. 26, no. 1, pp. 45–51, 2012.
- [19] W. Karlen, J. M. Ansermino, and G. A. Dumont, "Adaptive Pulse Segmentation and Artifact Detection in Photoplethysmography for Mobile Applications," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, 2012, pp. 3131–4.
- [20] W. Karlen, C. J. Brouse, E. Cooke, J. M. Ansermino, and G. A. Dumont, "Respiratory rate estimation using respiratory sinus arrhythmia from photoplethysmography." in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Boston, 2011, pp. 1201–4.
- [21] R. D. Berger, S. Akselrod, D. Gordon, and R. J. Cohen, "An Efficient Algorithm for Spectral Analysis of Heart Rate Variability," *IEEE Trans Biomed Eng*, vol. 33, no. 9, pp. 900–4, 1986.
- [22] W. Karlen, M. Turner, E. Cooke, G. A. Dumont, and J. M. Ansermino, "CapnoBase: Signal database and tools to collect, share and annotate respiratory signals," in *Annual Meeting of the Society for Technology in Anesthesia (STA)*, West Palm Beach, 2010, p. 25.
- [23] A. Johansson and P. A. Oberg, "Estimation of respiratory volumes from the photoplethysmographic signal. Part I: Experimental results." *Med Biol Eng Comp*, vol. 37, no. 1, pp. 42–7, 1999.
- [24] L. Nilsson, A. Johansson, and S. Kalman, "Respiration can be monitored by photoplethysmography with high sensitivity and specificity regardless of anaesthesia and ventilatory mode." *Acta Anaesth Scan*, vol. 49, no. 8, pp. 1157–62, 2005.
- [25] C. Julien, "The enigma of Mayer waves: Facts and models." *Cardiovasc Res*, vol. 70, no. 1, pp. 12–21, 2006.
- [26] P. Yang, G. A. Dumont, and J. M. Ansermino, "Sensor fusion using a hybrid median filter for artifact removal in intraoperative heart rate monitoring." *J Clin Monitor Comp*, vol. 23, no. 2, pp. 75–83, 2009.
- [27] S. Nemati, A. Malhotra, and G. D. Clifford, "Data Fusion for Improved Respiration Rate Estimation." *EURASIP J Adv Sig Pr*, vol. 2010, p. 926305, 2010.



Walter Karlen (M'08) holds a MSc degree in micro-engineering (2005) and a PhD in Computer, Communication and Information Sciences (2009) from Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland.

He is currently a post-doctoral researcher at the Biomedical Engineering Research Group at the Department of Mechanical and Mechatronic Engineering at the University of Stellenbosch, South Africa and at the Electrical and Computer Engineering in Medicine Research Group at the Child and Family Research Institute and the Department of Electrical and Computer Engineering of the University of British Columbia in Vancouver, Canada.

Dr. Karlen holds an Advanced Researcher Fellowship from the Swiss National Science Foundation and is an awardee of the Rising Stars in Global Health program of Grand Challenges Canada.



Srinivas Raman received his Bachelor's and Master's degree in mechanical engineering both from the University of British Columbia, Vancouver, Canada in 2007 and 2009 respectively. He is currently a medical student at the University of British Columbia.

His current research interests include biomedical device development and artificial intelligence applications in medicine.



J. Mark Ansermino received his Medical degree (1985) from the University of the Witwatersrand, Johannesburg, South Africa and a MSc degree in Health Informatics (2001) from City University, London, U.K.

He is an Associate Professor at the University of British Columbia's Department of Anesthesiology, Pharmacology and Therapeutics, and a Pediatric Anesthesiologist at the British Columbia's Children's Hospital, Vancouver, BC, Canada. He is co-director of the Electrical and Computer Engineering

in Medicine (ECEM) Research Group.

Dr. Ansermino is a Fellow of the Royal College of Physicians and Surgeons of Canada. He is a recipient of a Scholar Award from the Michael Smith Research Foundation, Canada and of the 2010 NSERC Brockhouse Canada Prize for Interdisciplinary Research in Science and Engineering together with Dr. Guy A. Dumont.



Guy A. Dumont (M'77-SM'84-F'99) holds a Dipl. Ing. degree (1973) from Ecole Nationale Supérieure d'Arts et Métiers, Paris, France, and a PhD in electrical engineering (1977) from McGill University, Montreal, Canada.

He worked with Tioxide, France (1973-1974; 1977-1979), then with Paprican (1979-1989). In 1989, he joined the Department of Electrical and Computer Engineering, University of British Columbia, where he is a Professor and co-director of the Electrical and Computer Engineering in

Medicine (ECEM) Research Group.

Dr. Dumont's awards and honors include the IEEE Control Systems Society 1998 Control Systems Technology Award, NSERC Synergy Awards in both 1999 and 2002, and the 2010 NSERC Brockhouse Canada Prize for Interdisciplinary Research in Science and Engineering.