Extracting Instantaneous Respiratory Rate From Multiple Photoplethysmogram Respiratory-Induced Variations

Parastoo Dehkordi [1]*, Ainara Garde [2†], Behnam Molavi [3†], J. Mark Ansermino [4] and Guy A. Dumont [1]

[1] Electrical and Computer Engineering, Faculty of Applied Science, The University of British Columbia, Vancouver, BC, Canada, [2] Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, Netherlands, [3] Lionsgate Technologies, Vancouver, BC, Canada, [4] Anesthesiology, Pharmacology and Therapeutics, The University of British Columbia, Vancouver, BC, Canada

In this study, we proposed a novel method for extracting the instantaneous respiratory rate (IRR) from the pulse oximeter photoplethysmogram (PPG). The method was performed in three main steps: (1) a time-frequency transform called synchrosqueezing transform (SST) was used to extract the respiratory-induced intensity, amplitude and frequency variation signals from PPG, (2) the second SST was applied to each extracted respiratory-induced variation signal to estimate the corresponding IRR, and (3) the proposed peak-conditioned fusion method then combined the IRR estimates to calculate the final IRR. We validated the implemented method with capnography and nasal/oral airflow as the reference RR using the limits of agreement (LOA) approach. Compared to simple fusion and single respiratory-induced variation estimations, peak-conditioned fusion shows better performance. It provided a bias of 0.28 bpm with the 95% LOAs ranging from −3.62 to 4.17, validated against capnography and a bias of 0.04 bpm with the 95% LOAs ranging from −5.74 to 5.82, validated against nasal/oral airflow. This algorithm would expand the functionality of a conventional pulse oximetry beyond the measurement of heart rate and oxygen saturation to measure the respiratory rate continuously and instantly.

Keywords: instantaneous respiratory rate, pulse oximetry, photoplethysmogram, respiratory-induced variation, synchrosqueezing transform

1. INTRODUCTION

Respiratory rate (RR), along with other vital signs like heart rate (HR) and blood pressure, is monitored for primary or continuous assessment of patient wellness. There is significant evidence that an abnormal respiratory rate is an important predictor of serious illness. For example, in children aged 1–5 years old, an elevated RR (>40 breaths/min) is an important criterion for the diagnosis of pneumonia (WHO, 2013). Furthermore, Fieselmann et al. analyzed the measurements of vital signs during the 72 h prior to cardiac arrest and showed that a high respiratory rate (>27 breaths/min) was a significant predictor of cardiac arrest in hospitals (Fieselmann et al., 1993). In addition, Subbe et al. showed that relative changes in respiratory rate are much more significant than changes in HR or systolic blood pressure in unstable patients and therefore the respiratory rate is more likely to be a better predictor for identifying the patient at risk (Subbe et al., 2003).

RR can be measured by a nurse counting the number of times the chest rises in 1 min (Lovett et al., 2005). Continuous monitoring of RR, though, needs a monitoring device and can be performed using capnography, transthoracic impedance pneumography, nasal/oral pressure transducers and abdominal/thoracic respiratory inductance plethysmography belts, among others. However, recent studies have found that neither the nurses nor the monitoring devices provide accurate and reliable measurements of RR (Lovett et al., 2005). Therefore, there is a clear need for a robust, automatic, reliable and non-invasive measure of RR for performing a spot-check and for continuous monitoring.
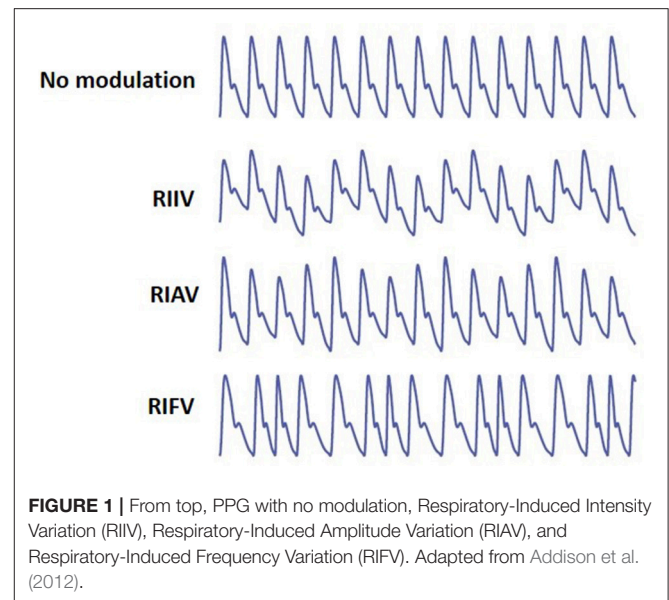
Analysis of the photoplethysmogram (PPG) recorded using a pulse oximeter could offer an alternative method for monitoring RR. The PPG waveform contains information about a wide range of physiological parameters such as HR, heart rate variability (HRV), oxygen saturation (SpO_2), vascular tone, blood pressure, cardiac output and respiration (Shelley, 2007). However, most conventional pulse oximeters only provide information about HR and SpO_2. In this study, we have presented an algorithm for robust estimation of instantaneous respiratory rate (IRR) from PPG with the aim of developing a portable solution based on pulse oximetry, suitable for both continuous monitoring and spot-check applications.

1.1. Background

A pulse oximeter measures the blood volume changes or PPG, based on the light absorption characteristics of the blood at the measuring site on the body (e.g., finger, forehead, and earlobe). A conventional transmission pulse oximeter sensor consists of two LEDs, red and infra-red, and a photo-detector. The LEDs emit light at both red and infrared wavelengths (650 and 950 nm, respectively) and the photo-detector captures the transmitted light. Based on the Beer-Lambert's law, the transmitted light density decreases during systole when the peripheral arterial blood volume is at its maximum value and increases during diastole. The PPG signal generated in the photo-detector, then, has a pulsatile waveform (AC) whose peaks and troughs reflect light transmitted through the tissue when blood volume is minimal and maximal, respectively. PPG also has a large baseline component (DC) mainly rises because of constant absorption and scattering of light traveling through skin, bones and tissues. The small variation observed in DC is mostly due to venous blood variation which changes the intensity of the light captured by the photo-detector (Webster et al., 1997).

Respiration may induce variation in PPG in three different ways (Meredith et al., 2012) (**Figure 1**):

(1) Respiratory-Induced Intensity Variation (RIIV): Changes in venous return due to changes in intra-thoracic pressure throughout the respiratory cycle cause a baseline (DC) modulation of the PPG signal. During inspiration, decreases in intra-thoracic pressure result in a small decrease in central venous pressure increasing venous return. The opposite occurs during expiration. As the venous bed at probing site cyclically fills and drains, the baseline is modulated accordingly.

FIGURE 1 | From top, PPG with no modulation, Respiratory-Induced Intensity Variation (RIIV), Respiratory-Induced Amplitude Variation (RIAV), and Respiratory-Induced Frequency Variation (RIFV). Adapted from Addison et al. (2012).

(2) Respiratory-Induced Amplitude Variation (RIAV): During inspiration, left ventricular stroke volume decreases due to changes in intra-thoracic pressure leading to the decreased pulse amplitude. The opposite happens during expiration.

(3) Respiratory-Induced Frequency Variation (RIFV): HR varies throughout the respiratory cycle; HR increases during inspiration and decreases during expiration. This phenomenon well-known as respiratory sinus arrhythmia (RSA) is mainly due to the autonomic regulation of HR during respiration.

Respiration may induce variation in PPG differently among different individuals in health and disease. For instance, RIFV, as an indicator of autonomic activity, maybe affected by diseases and disorders (e.g., myocardial infarction, diabetic neuropathy or sleep breathing disorders Dehkordi et al., 2016). RIAV and RIIV are also very sensitive to dehydration and hypovolemia. In addition, respiratory-induce variations are different for women and men (Li et al., 2010). Estimation of IRR by combining the information from three respiratory-induce waveform variations, then, improves the algorithm performance and increases the robustness of results (Karlen et al., 2013).

Many algorithms have been proposed to estimate RR from PPG. Auto-regression (Thayer et al., 2002), Fourier transform analysis (Karlen et al., 2013), correntropy spectral density (Garde et al., 2014), digital filters (Nakajima et al., 1996) and empirical mode decomposition (Garde et al., 2013) were successfully used, among others. These algorithms have mostly focused on estimating average RR from a window of PPG. For example, Karlen et al. (2013) and Garde et al. (2014) estimated RR every second using 16, 32, 64-s windows of PPG data.

Few algorithms, however, have proposed to estimate RR instantaneously (IRR), mostly performed by the time-frequency approaches based on a continuous wavelet (Clifton et al., 2007; Addison et al., 2012), variable frequency complex demodulation

methods (VFCDM) (Chon et al., 2009) and short-time Fourier analysis (STFT) (Shelley et al., 2006).

In this study, we have proposed a novel method for extracting the instantaneous respiratory rate (IRR) from PPG. The method was performed in three main steps: extraction of RIIV, RIAV, and RIFV signals from PPG, estimation of IRR from each extracted respiratory-induced variation signals and fusion of IRR estimates. A time-frequency transform called synchrosqueezing transform (SST) (Daubechies et al., 2011) was used to extract RIIV, RIAV, and RIFV from PPG. Later, a second SST was applied to estimate IIR from respiratory-induced variation signals (Addison and Watson, 2004). To combine IRR estimates corresponding to each respiratory-induced variation signal, a novel method, called peak-conditioned fusion algorithm was proposed.

2. ALGORITHM DESCRIPTION

2.1. Instantaneous Frequency (IF)

The instantaneous frequency is the frequency at a given time. Consider a multicomponent signal f that can be modeled as

$$f(t) = \sum_{k=1}^{K} f_k(t) = \sum_{k=1}^{K} A_k(t) \cos(2\pi \phi_k(t)) \quad (1)$$

where $A_k(t)$ and $\phi_k(t)$ are the time-varying amplitude and phase of kth frequency component, respectively.

In theory, the instantaneous frequency (IF) is defined as the derivative of the phase function with respect to time as

$$IF_f = \{\phi'_k(t)\}_{1 \leq k \leq K} \quad (2)$$

2.2. Synchrosqueezing Transform

Synchrosqueezing Transform (SST) was first introduced by Daubechies et al. (2011) in 1996 and then implemented by Thakur et al. (2013). SST is a combination of wavelet analysis and a reallocation method which sharpens a time-frequency representation by allocating its points to another locations in the time-frequency plane. SST can provide an accurate estimation of IF.

As defined in Daubechies et al. (2011), SST involves three steps:

Step 1: Estimation of the continuous wavelet transform (CWT)

The CWT of f is calculated as

$$W_f(a,b) = \int f(t) a^{-1/2} \overline{\psi\left(\frac{t-b}{a}\right)} \mathrm{d}t \quad (3)$$

where ψ is a wavelet with $\hat{\psi}(\xi) = 0$ for $\xi \leq 0$ and a and b are scale and location variables, respectively. $\overline{\psi(\xi)}$ is the complex conjugate of $\psi(\xi)$ and $\hat{\psi}(\xi)$ is the Fourier transform of $\psi(\xi)$ estimated as

$$\hat{\psi}(\xi) = \int \psi(\xi) e^{-i(2\pi\xi)t} \mathrm{d}t \quad (4)$$

Step 2: Estimation of the instantaneous frequency

If $\hat{\psi}(\xi)$ is concentrated around $\xi = \omega_0$, then $W_f(a,b)$ will be spread out around the horizontal line $a = {\omega_0}/{\omega}$ on the time-scale presentation for a given frequency of ω. However, Daubechies et al. (2011) showed that the oscillation of $W_f(a,b)$ around b tends to the original frequency ω, irrespective of the value of a. Therefore, for any (a,b) where $W_f(a,b) \neq 0$, the instantaneous frequency $\omega_f(a,b)$ for signal f can be defined as

$$\omega_f(a,b) = -\frac{i}{2\pi}((W_f(a,b))^{-1}\frac{\partial}{\partial b}W_f(a,b)) \quad (5)$$

Step 3: Transfer to the time-frequency plane

In this step, each point on the time-scale plane is allocated to a point on the time-frequency plane using the map $(a,b) \rightarrow (\omega_f(a,b), b)$. The frequency variable ω and the scale variable a are both binned: $W_f(a,b)$ is computed only at discrete values a_k, with $a_k - a_{k-1} = (\Delta a)_k$ and its SST, $T_f(\omega,b)$, is estimated only at the centers ω_l of the successive bins $[\omega_l - \frac{1}{2}, \omega_l + \frac{1}{2}]$, with $\omega_l - \omega_{l-1} = \Delta\omega$, by summing different points:

$$T_f(\omega_l,b) = (\Delta\omega)^{-1} \sum_{a_k : |\omega(a_k,b)-\omega_l| \leq \frac{\Delta\omega}{2}} W_f(a_k,b) a_k^{\frac{-3}{2}} (\Delta a)_k. \quad (6)$$

3. MATERIALS AND METHODS

3.1. Data Set

3.1.1. Capnobase Data Set

The Capnobase contains test and calibration data sets (Karlen et al., 2010). Test data set contains forty-two 8-min segments of recordings obtained from 29 pediatric and 13 adults receiving general anesthesia at the British Columbia Children's Hospital and St. Paul's Hospital, Vancouver, BC, respectively. Calibration data set contains one hundred twenty-four 2-min segments of recordings used for tuning the parameters of the proposed algorithm.

In both data sets, the recordings included ECG, capnometry, and PPG (sampled at 300, 300, and 100 Hz, respectively) obtained with S/5 collect software (Datex-Ohmeda, Finland). The capnography waveform was used as the reference gold standard recording for RR. A research assistant manually labeled each breath in the capnogram. The beginning and end of all artifacts in the PPG waveforms were also manually labeled. Both datasets can be downloaded from the on-line database, CapnoBase.org.

3.1.2. Sleep Data Set

Sleep database contains forty-three 20-min segments of recording from 43 children (age = 9.1 ± 4.1, Apnea/Hypopnea Index = 8.9 ± 17.9) referred to the British Columbia Children's Hospital for overnight standard polysomnography (PSG). The children had been recruited following approval by the University of British Columbia Clinic Research Ethics Board (H11-01769) and informed parental consent. Parental/guardian written

informed consent was obtained for all children. Children with a cardiac arrhythmia or abnormal hemoglobin were excluded from the study.

Standard PSG recordings included overnight measurements of ECG, electroencephalography (EEG), oxygen saturation (SpO_2), PPG, chest and abdominal movement, nasal and oral airflow, left and right electrooculography (EOG), electromyography (EMG) and video capture. The PSG recordings were performed with the Embla Sandman S4500 (Embla Systems, ON, Canada).

In addition to PSG, the PPG was recorded simultaneously using the Phone Oximeter™ (Karlen et al., 2011) sampled at 62.5 Hz with 32-bit resolution.

For each subject, the 20 min of recordings were selected after 2 h of sleep. Apnea/hypopnea events were annotated as the artifacts and no specific sleep stage was selected.

The nasal/oral airflow waveform was used as the reference gold standard recording for RR. Two experts manually labeled each breath in nasal/oral airflow waveform. The beginning and end of all artifacts in the oral/nasal waveforms were also manually labeled.

Sleep database is available for download at https://figshare.com/s/552ec33f37ae8d99c032 (doi: 10.6084/m9.figshare.6683807).

3.2. Estimation of IRR From PPG

To perform IRR estimation, after a preprocessing stage, a first SST was applied to PPG to extract RIIV, RIAV and RIFV. Later, a second SST was performed to estimate IIR from a respiratory-induced variation signals. The peak-conditioned fusion algorithm was then used to fuse simultaneous IRR estimates. This procedure, inspired by the method known as secondary wavelet feature decoupling (SWFD) (Addison and Watson, 2004), involves the following steps (**Figure 2**):

(1) The first SST is applied to the PPG signal.
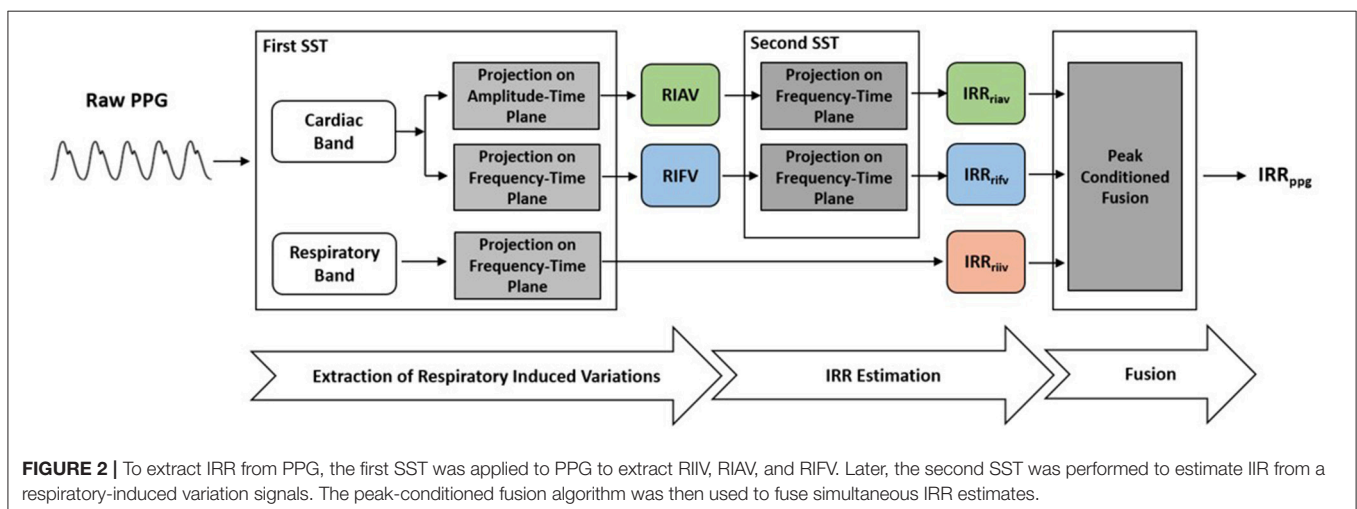(2) In the STT surface, two components are identified: a strong cardiac component in the cardiac band (0.5–3 Hz, 30–180 beats/min) and a respiratory component in the respiratory band (0.14–0.9 Hz, 8–54 breaths/min) (**Figure 3**).
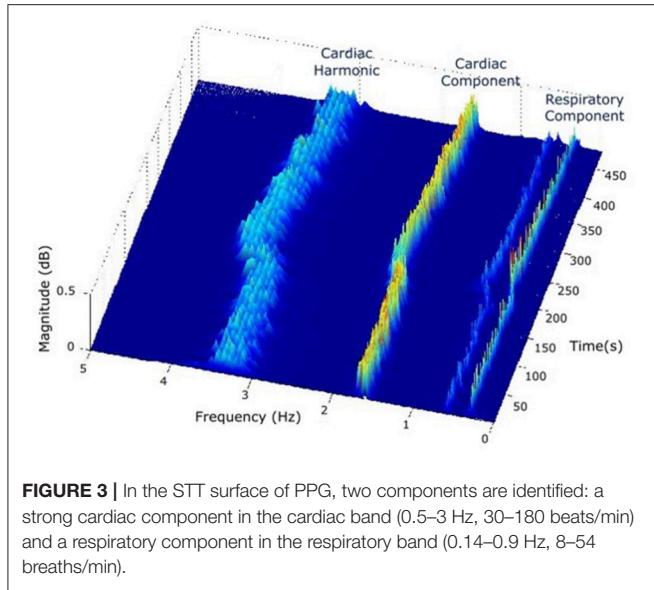
In this study, reference ranges of cardiac and respiratory bands were extracted from a review of observational studies that used HR from 143,346 children and RR data from 3,881 children (from 6 months to 18 years old) (Fleming et al., 2011). Based on 99th and 1st centiles for children and young adults, the HR could range from 30 to 180 beats/min (0.50 to 3 Hz, respectively) and RR from 8 to 54 breaths/min (0.14 to 0.9 Hz, respectively). The range in adults is much more restricted, thus it would be included in this range.

(3) The respiratory component in SST surface shows RIIV and its ridge in the frequency-time plane represents RIIV-derived IRR (IRR_{riiv}) (**Figure 3**).
(4) The ridge of cardiac component is followed either in the amplitude-time plane to get RIAV or in the frequency-time plane to get RIFV. This is done by projecting the cardiac ridge points onto the amplitude-time or frequency-time planes, respectively.
(5) The second SST applied to RIAV results in a dominant single component in the respiratory band (0.14–0.9 Hz, 8–54 breaths/min) whose ridge represents RIAV-derived IRR (IRR_{riav}).
(6) A second SST is applied to the RIFV signal as well to get a dominant single component in the respiratory band whose ridge represents RIFV-derived IRR (IRR_{rifv}).
(7) Estimation of final IRR (IRR_{ppg}) is performed using a proposed peak frequency tracking method (so-called peak-conditioning fusion) which combines the instantaneous frequency information from (IRR_{riiv}), (IRR_{riav}) and (IRR_{rifv}).

3.2.1. Preprocessing
The PPG signals were lowpass filtered by a lowpass Chebyshev Type I IIR filter of order 8 and down sampled to 10 Hz.

FIGURE 2 | To extract IRR from PPG, the first SST was applied to PPG to extract RIIV, RIAV, and RIFV. Later, the second SST was performed to estimate IIR from a respiratory-induced variation signals. The peak-conditioned fusion algorithm was then used to fuse simultaneous IRR estimates.

FIGURE 3 | In the STT surface of PPG, two components are identified: a strong cardiac component in the cardiac band (0.5–3 Hz, 30–180 beats/min) and a respiratory component in the respiratory band (0.14–0.9 Hz, 8–54 breaths/min).

3.2.2. Estimation of IRR_{riiv}

Consider a PPG signal as a vector $ppg \in R^n$, $n = 2^{L+1}$ where L is a nonnegative integer. The CWT of ppg, W_{ppg}, was calculated using the Morlet wavelet, ψ, where its Fourier transform was concentrated around 1.25 Hz. The W_{ppg} was sampled at the location (a_j, b), where $a_j = 2^{j/n_v}$, $j = 1, ..., Ln_v$, $n_v = 32$ and $b = 1, ..., n$. The result is a $Ln_v \times n$ matrix denoted $\tilde{W}_{ppg}$.

When $\tilde{W}_{ppg} > 0$, $\tilde{\omega}_{ppg}$ was implemented as follow

$$\tilde{\omega}_{ppg} = -\frac{i}{2\pi} D_b \tilde{W}_{ppg}(a_j, b) \tilde{W}_{ppg}(a_j, b)^{-1} \quad (7)$$

where $D_b \tilde{W}_{ppg}$ was the finite differences of $\tilde{W}_{ppg}$ with respect to b.

Then frequency variable, ω, was binned into frequency division $\omega_l = 2^{l\triangle\omega}\underline{\omega}$, $l = 0, ..., Ln_v - 1$, where $\triangle\omega = \frac{1}{Ln_v - 1} log_2(\frac{n}{2})$, $\underline{\omega} = \frac{1}{n\triangle t}$ and $\bar{\omega} = \frac{1}{2\triangle t}$. $\bar{\omega}$ and $\underline{\omega}$, were maximum and minimum frequencies respectively and were chosen based on Nyquist sampling theorem.

The SST of PPG was calculated as

$$T_{ppg}(\omega_l, b) = \sum_{a_j : |\omega(a_j,b)-\omega_l| \leq \frac{\triangle\omega}{2}} \frac{log2}{Ln_v} \tilde{W}_{ppg}(a_j, b) a_j^{\frac{-1}{2}}. \quad (8)$$

T_{ppg} over time shows both cardiac and respiratory bands (**Figure 3**).

A ridge fitting the dominant area of T_{ppg} in the respiratory band (0.14–1 Hz) represented IRR_{riiv} and was extracted by tracking the local maximum values in this region.

3.2.3. Estimation of IRR_{riav}

Consider $RIAV$ as a vector $riav \in R^n$, where n is the length of ppg. In the amplitude-time plane of T_{ppg}, $riav$ estimated as a ridge fitting the dominant area of T_{ppg} in the cardiac band (0.5–3 Hz, 30–180 beats/min). The ridge extracted by finding the local

maximum values which minimize the following cost function (Abid et al., 2007):

$$Cost = \sum_{b=1}^{n} [-|T_{ppg}(riav(b), b)|^2 + |riav(b) - riav(b-1)|^2] \quad (9)$$

The SST of $riav$, T_{riav} was calculated using the same implementation described in the previous section.

A ridge fitting the dominant area of T_{riav} in the respiratory band (0.14–0.9 Hz) represented the RIAV-derived IRR (IRR_{riav}) and can be extracted by tracking the local maximum values in this region.

3.2.4. Estimation of IRR_{rifv}

Consider $RIFV$ as a vector $rifv \in R^n$, where n is the length of ppg. In the frequency-time plane of T_{ppg}, $rifv$ estimated as a ridge fitting the dominant area of T_{ppg} in the cardiac band (0.5–3 Hz, 30–180 beats/min). The ridge extracted by finding the local maximum values which minimize the following cost function (Abid et al., 2007):

$$Cost = \sum_{b=1}^{n} [-|T_{ppg}(rifv(b), b)|^2 + |rifv(b) - rifv(b-1)|^2] \quad (10)$$

The SST of $riav$, T_{riav} was calculated using the same implementation described in the section 3.2.2.

A ridge fitting the dominant area of T_{riav} in the respiratory band (0.14–0.9 Hz) represented the RIFV-derived IRR (IRR_{rifv}) and can be extracted by tracking the local maximum values in this region.

3.2.5. Peak-Conditioned Fusion

The peak-conditioned fusion method, inspired by Lázaro Plaza (2015), was proposed to combine the IRR estimates from three respiratory-induced variations to provide the final IRR_{ppg}.

The calculated T_{ppg}, T_{riav} and T_{rifv} are two-dimensional matrices $\in R^{Ln_v n}$, $n = 2^{L+1}$ where L is a nonnegative integer and $n_v = 32$. Each column of T_{ppg}, T_{riav} and T_{rifv} matrices shows the frequency distribution of PPG, RIAV and RIFV signals at each time instance, respectively. To reduce the variance, each matrix is averaged in time dimension using a moving window of length $T_m = 16$ s every $t_s = 5$ s. The averaged matrix is denoted as $\hat{T}_k$, where k refers to ppg, $riav$, or $rifv$ (**Figure 4**).

At instant b, the location of the largest peak in respiratory band of each $\hat{T}_k(:, b)$ column (for k = ppg, riav or rifv) is detected and denoted as $ir_k^I(b)$. Then, a reference frequency interval, $\Omega_k(b)$, was defined as

$$\Omega_k(b) = [f(b-1) - \delta, f(b-1) + 2\delta] \quad (11)$$

where $(b-1)$ was a respiratory rate reference estimated from the $b-1$ previous step.

All peaks larger than 85% of $ir_k^I(b)$ inside $\Omega(b)$ were detected and $ir_k^{II}(b)$ was chosen as the nearest to $f_{(b-1)}$. By reaching to this point, $ir_{riiv}^{II}(b)$, $ir_{riav}^{II}(b)$ and $ir_{rifv}^{II}(b)$ were available simultaneously.

The final respiratory peak at instant b, $IIR_{ppg}((b))$, was then chosen among $ir_{riiv}^{II}(b)$, $ir_{riav}^{II}(b)$ and $ir_{rifv}^{II}(b)$ estimates with the
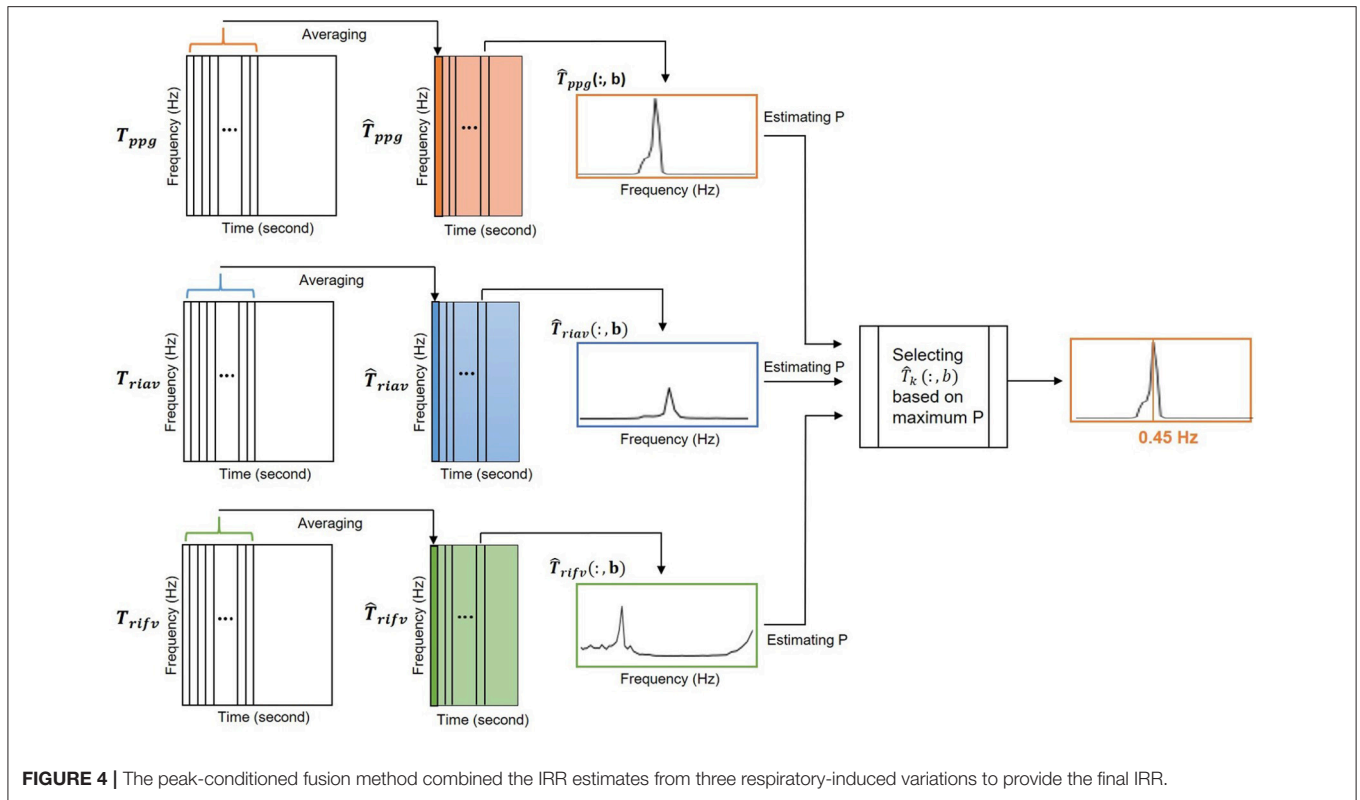
FIGURE 4 | The peak-conditioned fusion method combined the IRR estimates from three respiratory-induced variations to provide the final IRR.

largest P_k. P_k is a measure of the peakness and was defined as the ratio of power contained in an interval centered around the largest peak to the power of $\Omega_k(b)$. P mathematically calculated as

$$p_k(b) = \frac{\sum_{max\{if_k^{II}(b)-0.6\delta,f(b)-\delta\}}^{min\{if_k^{II}(b)+0.6\delta,f(b)+2\delta\}} T_k(\hat{:}, b)}{\sum_{f(b)-\delta}^{f(b)+2\delta} \hat{T}_k(:, b)} \qquad (12)$$

Estimation of respiratory rate as the largest peak in the respiratory band would increase the risk of choosing the location of false peaks. To decrease this risk, the search for the largest peak was limited to the reference frequency interval, $\Omega_k(b)$ (Lázaro Plaza, 2015). This is an asymmetric interval of 3δ centered around a reference frequency. At each step the respiratory rate reference was updated using

$$f(b + 1) = \beta * f(b) + (1 - \beta) * IRR_{ppg}(b) \qquad (13)$$

where $f(b) = arg\ max(\hat{T}_k(:, 1))$ in the frequency band of [0.2Hz, 0.7Hz].

Value of δ was set as 0.1 and the value of a was tuned as 0.6 over the calibration data set.

3.3. Algorithm Evaluation
To evaluate the performance of SST-based algorithms, agreement between reference IRR and estimated IRR (using peak-conditioned fusion, simple fusion, single respiratory-induce

variation) was assessed using the limits of agreement (LOA) technique. The bias and 95% LOA were estimated using the Bland-Altman plot. Since for each subject multiple measurement were observed, the Bland-Altman method for multiple observations per individual (Zou, 2013) was used instead of the standard Bland-Altman method. The bias was calculated as mean of IRR_{est} - IRR_{ref} and the 95% LOAs as mean bias $\pm$ 1.95 standard deviations. Two standard deviations (2SD) were also estimated in the purpose of ranking the proposed algorithm in this study based on the statistical analysis reported by Charlton et al. (2016).

The coverage probability (CP_2) was also reported as the probability of measurement error falling within pre-defined bounds, set as 2 breaths per minute (bpm) in this study (Barnhart et al., 2007).

In addition, the performance of algorithms was assessed using the root mean square error (RMSE) (breaths/min) defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(IRR_{ref} - IRR_{est})^2} \qquad (14)$$

where n is the number of observations and IRR_{ref} and IRR_{est} are the reference and the algorithm estimates, respectively.

4. RESULTS

4.1. Capnobase Data Base

IRR extracted from the capnography waveform (IRR_{CO2}) was used as the reference gold standard. The distribution of the respiratory rates contained 3,542 data points estimated every 5 s from IRR_{CO2} for the 16 s moving windows over the whole dataset. The respiratory rates ranged from the lowest value of 3.6521 bpm to the highest value of 44.22 bpm. The mean rate was 15.02 bpm with standard deviation of 7.66 bpm. About 7.7% of the data points were excluded from the further analysis due to to poor signal quality of the capnography signals.

For each algorithm, the measures of agreement between the estimated IRR from PPG (IRR_{est}) and IRR_{CO2} and also RMSE were estimated (**Table 1**). For peak selection algorithm, bias was estimated as 0.28 bpm with the 95% LOAs from −3.62 to 4.17 (**Figure 5**). The value of 2SD was estimated as 3.97 bpm. The value of CP_2 showed that for 89% of the IRR estimates, the measured error was less than 2 breaths/min. RMSE was estimated as 1.8 bpm.

The values of 2SD of the other algorithms ranged from 8.32 bpm to 16.00 bpm.

4.2. Sleep Database

IRR extracted from the nasal/oral airflow waveform (IRR_{nas}) was used as the reference gold standard in the sleep dataset. The distribution of the respiratory rates contained 10,553 data points estimated every 5 s from IRR_{nas} over the 16 s moving window for all subjects. The respiratory rates ranged from the lowest value of 9.561 bpm to the highest value of 50.85 bpm. The mean rate was 18.64 bpm with standard deviation of 5.66 bpm. About 0.66% of the data points were excluded from the further analysis due to to poor signal quality of the nasal/oral airflow signals.

The measures of agreement between the estimated IRR from PPG (IRR_{est}) and IRR_{nas} also RMSE were estimated for each algorithm (**Table 1**). For peak selection algorithm, bias was estimated as 0.04 bpm with the 95% LOAs from −5.74 to 5.82 (**Figure 6**). The value of 2SD was estimated as 5.90 bpm. The value of CP_2 showed that for 85% of the IRR estimates, the

measured error was less than 2 breaths/min. RMSE was estimated as 2.3 bpm.

The values of 2SD of the other algorithms ranged from 8.4 bpm to 21.3 bpm.
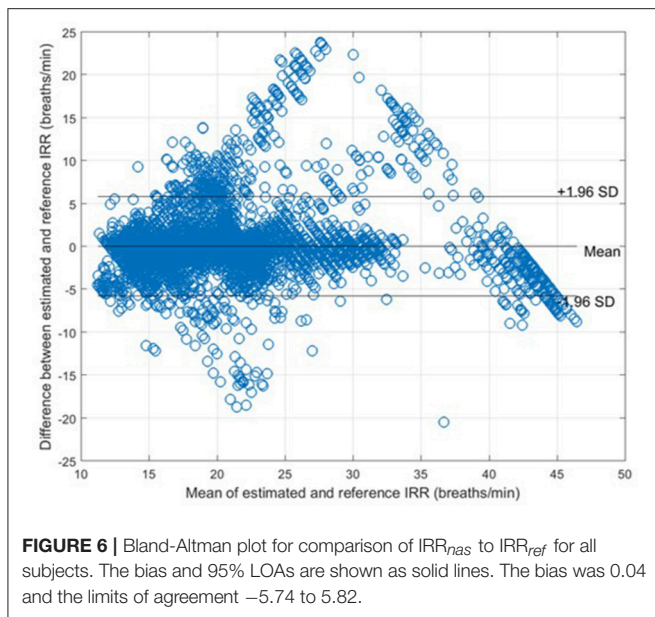
5. DISCUSSION AND CONCLUSION

In this study, we presented an algorithm to extract IRR from PPG. We extracted RIIV, RIAV, and RIFV from PPG using Synchrosqueezing Transform (SST), a sharpening time-frequency method which provides instantaneous frequency rate. The peak-conditioned fusion was proposed to combine the extracted information from three respiratory induced variations waveforms to estimate respiratory rate at each instance. We validated the implemented method with capnography and nasal/oral airflow as the reference RR. Compared to simple fusion and single respiratory-induced variation estimations, peak-conditioned fusion shows better performance (**Table 1**). It
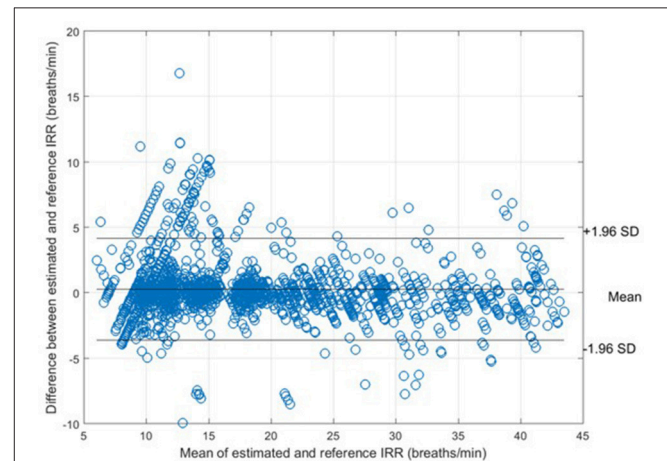
FIGURE 5 | Bland-Altman plot for comparison of IRR_{CO2} to IRR_{ref} for all subjects. The bias and 95% LOAs are shown as solid lines. The bias was 0.28 and the limits of agreement −3.62 to 4.17.

TABLE 1 | The performance of different method for estimation IRR from PPG.

	Different IRR estimation method	2SD	Bias	95% LOA	CP₂	Proportion of windows with IRR estimate (%)	RMSE
	RIIV	8.8	0.3	−8.3 to 9.0	88	100	4.1
	RIAV	16.0	1.3	−14.5 to 16.9	60	100	7.5
Capnobase	RIFV	9.2	0.04	−9.0 to 9.1	74	100	4.2
dataset	Simple Fusion	8.3	0.5	−7.6 to 8.7	63	100	4.0
	Peak-Conditioned Fusion	4.0	0.3	−3.6 to 4.2	89	100	1.8
	RIIV	11.0	0.7	−10.1 to 11.4	80	100	5.1
	RIAV	21.3	5.6	−15.4 to 26.5	31	100	11.3
Sleep	RIFV	8.4	−0.1	−8.4 to 8.2	79	100	3.7
dataset	Simple Fusion	9.5	2.0	−7.3 to 11.3	41	100	4.6
	Peak-Conditioned Fusion	5.9	0.04	−5.7 to 5.8	85	100	2.3

FIGURE 6 | Bland-Altman plot for comparison of IRR$_{nas}$ to IRR$_{ref}$ for all subjects. The bias and 95% LOAs are shown as solid lines. The bias was 0.04 and the limits of agreement −5.74 to 5.82.

provided a bias of 0.28 bpm with the 95% LOAs ranging from −3.62 to 4.17, validated against capnography (in the Capnobase dataset) (**Figure 5**) and a bias of 0.04 bpm with the 95% LOAs ranging from −5.74 to 5.82, validated against nasal/oral airflow (in the Sleep dataset) (**Figure 6**).

In this study, the proposed method estimated IRR from three sources of respiratory-induced variation and fused the estimated rates to measure the final IRR. Our findings showed that fusion of estimation rates would increase the accuracy and robustness of RR estimation. Even the simple fusion compared to single respiratory-induced variation estimations showed higher rank (narrower 2SD and greater CP2). It is consistent with the findings of Li et al. (2010) that respiratory activity may induce variation in PPG differently in different individuals. As discussed by Karlen et al. (2013), ventilatory conditions (spontaneous or mechanical ventilation) can change the behavior of respiratory induced variations.

In this study, we applied the proposed algorithm to two different data sets to include a broad range of subjects into the study. The Capnobase data set includes children adults, under controlled ventilation or spontaneously breathing over a wide RR range. The subjects were under general anesthesia and were continuously monitored. The sleep dataset includes children from 1-month to 17 years old spontaneously breathing during the overnight sleep in a sleep lab. During the overnight recordings, respiratory rates might change significantly during different sleep stages (light sleep, deep sleep or rapid eye movement (REM) sleep). In addition, some of the children may have experienced periods of breathing cessation, or obstructive sleep apnea. Since the apnea/hypopnea events induce changes on the morphology of nasal/oral airflow, i.e., flat line during apnea or very fast and low-amplitude oscillations during hypopnea, to extract the accurate reference RR from the nasal/oral airflow waveform, we labeled the apnea/hypopnea events as the artifact on nasal/oral airflow waveform.

A recent study (Charlton et al., 2016) presented a very complete assessment of RR estimation using PPG. A wide range of available techniques for estimation of respiratory-induced variations from PPG, estimation of RR from respiratory-induced variations, and fusion of RR estimates were identified and then more than 300 algorithms were implemented by assembling all possible combinations of available techniques. The algorithms were applied to the Vortal data set of 39 resting adults, with RR ranging between 5 and 32 bpm and ranked based on 2SD. The first ten top-ranked algorithms had the 2SD values ranging from 6.2 to 7.9. To be able to compare the performance of our algorithm with the ten top-ranked algorithms, we applied our proposed method on the same data set. For peak conditioned algorithm, the value of 2SD was estimated as 6.83 bpm. It ranked our algorithm among the top ten algorithms; however, the performance of our algorithm slightly decreased when RR was lower than 8 breaths/min. In addition, for the top ranked algorithms, the value of CP$_2$ was reported as 71.5 while we obtained a CP$_2$ of 75 applying our proposed algorithm.

It is important to note that all top ranked algorithm reported in Charlton et al. (2016) estimated RR using 32-s windows while our method can estimate RR instantaneously. It suggests that our algorithm shows better performance compared to methods that extract IRR based on time-frequency analysis (Li et al., 2010; Addison et al., 2012).

In Charlton et al. (2016), the methods for extracting RR from ECG were assessed as well. The findings of that study showed that algorithms performed better when using ECG than PPG. The best algorithm had 95% LOAs of −4.7 to 4.7 bpm and a bias of 0.0 bpm when using the ECG.

In Charlton et al. (2016), the performance of thoracic Impedance Pneumography (IP) were assessed as well providing a bias of −0.2 bpm with 95% LOAs of −5.6 to 5.2 bpm. Thoracic IP is a commonly-used technique for continuous monitoring of RR that measures changes in the electrical impedance of the person's chest during respiration. Our results showed that the performance of our algorithm is comparable with the performance of thoracic IP.

Several studies based on the continuous wavelet transform (CWT) (Clifton et al., 2007; Addison et al., 2012), the short-time Fourier transform (STFT) (Shelley et al., 2006), and empirical mode decomposition (EMD) (Garde et al., 2014) have been proposed to detect RR from PPG. The results of a study conducted by Thakur et al. (2013) to compare SST to CWT, STFT and EMD showed the superior precision of SST at identifying components of complicated oscillatory signals. Moreover, the study showed that time-varying instantaneous frequencies could be clearly distinguished in the SST while there is much more smearing and distortion in the CWT and STFT.

In this study, the SST algorithm has been implemented with O(nlog(n)) computations per scale and peak-conditioned fusion algorithm has been implemented with O(n) computations where n is the number of samples. However, there are some fast algorithms to reduce the number of computations of SST implementation to O(n) per scale Rioul and Duhamel (1992).

This study introduces a new method to estimate IRR from pulse oximetry. This would expand the functionality of a

conventional pulse oximetry beyond the measurement of HR and SpO_2 to measure the respiratory rate continuously and instantly in the clinical setting and at home. Importantly, these are all achievable with a simple, cheap, single-sensor solution.

AUTHOR CONTRIBUTIONS

All authors made substantial contributions to the conception and design of the paper. PD contributed to data acquisition and processed the obtained data, designed, developed and evaluated the algorithm, and drafted the review. AG contributed to data acquisition and the design of algorithm and revised the paper critically for content. BM contributed to the design of algorithm

and revised the paper critically for content. JA and GD revised the paper critically for content. AG and BM have contributed equally to this work. All authors approved the final submission of the document and agreed to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

REFERENCES

Abid, A. Z., Gdeisat, M. A., Burton, D. R., and Lalor, M. J. (2007). Ridge extraction algorithms for one-dimensional continuous wavelet transform: a comparison. *J. Phys.* 76, 1–8. doi: 10.1088/1742-6596/76/1/012045

Addison, P., and Watson, J. (2004). Secondary transform decoupling of shifted non stationary signal modulation components: applications to photoplethysmography. *Int. J. Wavelets Multiresolut. Inform. Process.* 2, 43–57. doi: 10.1142/S0219691304000329

Addison, P. S., Watson, J. N., Mestek, M. L., and Mecca, R. S. (2012). Developing an algorithm for pulse oximetry derived respiratory rate RR(oxi): a healthy volunteer study. *J. Clin. Monit. Comput.* 26, 45–51. doi: 10.1007/s10877-011-9332-y

Barnhart, H. X., Haber, M. J., and Lin, L. I. (2007). An overview on assessing agreement with continuous measurements. *J. Biopharmaceut. Stat.* 17, 529–569. doi: 10.1080/10543400701376480

Charlton, P. H., Bonnici, T., Tarassenko, L., Clifton, D. A., Beale, R., and Watkinson, P. J. (2016). An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram. *Physiol. Meas.* 37, 610–626. doi: 10.1088/0967-3334/37/4/610

Chon, K. H., Dash, S., and Ju, K. (2009). Estimation of respiratory rate from photoplethysmogram data using time-frequency spectral estimation. *IEEE Trans. Biomed. Eng.* 56, 2054–2063. doi: 10.1109/TBME.2009.2019766

Clifton, D., Douglas, G. J., Addison, P. S., and Watson, J. N. (2007). Measurement of respiratory rate from the photoplethysmogram in chest clinic patients. *J. Clin. Monit. Comput.* 21, 55–61. doi: 10.1007/s10877-006-9059-3

Daubechies, I., Lu, J., and Wu, H. T. (2011). Synchrosqueezed wavelet transforms: an empirical mode decomposition-like tool. *Appl. Comput. Harmon. Anal.* 30, 243–261. doi: 10.1016/j.acha.2010.08.002

Dehkordi, P., Garde, A., Karlen, W., Petersen, C. L., Wensley, D., Dumont, G. A., et al. (2016). Evaluation of cardiac modulation in children in response to apnea/hypopnea using the Phone Oximeter[TM]. *Physiol. Meas.* 37, 187–202. doi: 10.1088/0967-3334/37/2/187

Fieselmann, J. F., Hendryx, M. S., Helms, C. M., and Wakefield, D. S. (1993). Respiratory rate predicts cardiopulmonary arrest for internal medicine inpatients. *J. Gen. Int. Med.* 8, 354–360. doi: 10.1007/BF02600071

Fleming, S., Thompson, M., Stevens, R., Heneghan, C., Plüddemann, A., Maconochie, I., et al. (2011). Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies. *Lancet* 377, 1011–1019. doi: 10.1016/S0140-6736(10)62226-X

Garde, A., Karlen, W., Ansermino, J. M., and Dumont, G. A. (2014). Estimating respiratory and heart rates from the correntropy spectral density of the photoplethysmogram. *PLoS ONE* 9:e86427. doi: 10.1371/journal.pone.0086427

Garde, A., Karlen, W., Dehkordi, P., Ansermino, J., and Dumont, G. (2013). Empirical mode decomposition for respiratory and heart rate estimation from the photoplethysmogram. *Comput. Cardiol.* 40, 799–802.

Karlen, W., Dumont, G., Petersen, C., Gow, J., Lim, J., Sleiman, J., et al. (2011). "Human-centered phone oximeter interface design for the operating room," in *Proceedings of the International Conference on Health Informatics* (Rome), 433–438.

Karlen, W., Raman, S., Ansermino, J. M., and Dumont, G. (2013). Multi-parameter respiratory rate estimation from the photoplethysmogram. *IEEE Trans. Biomed. Eng.* 60, 1946–1953. doi: 10.1109/TBME.2013.2246160

Karlen, W., Turner, M., Cooke, E., Dumont, G. A., and Ansermino, J. M. (2010). "Capnobase: signal database and tools to collect, share and annotate respiratory signals," in *Annual Meeting of the Society for Technology in Anesthesia* (West Palm Beach, FL), 48.

Lázaro Plaza, J. (2015). *Non-invasive Techniques for Respiratory Information Extraction Based on Pulse Photoplethysmogram and Electrocardiogram.* Ph.D. thesis, University of Zaragoza.

Li, J., Jin, J., Chen, X., Sun, W., and Guo, P. (2010). Comparison of respiratory-induced variations in photoplethysmographic signals. *Physiol. Meas.* 31, 415–425. doi: 10.1088/0967-3334/31/3/009

Lovett, P. B., Buchwald, J. M., Stürmann, K., and Bijur, P. (2005). The vexatious vital: neither clinical measurements by nurses nor an electronic monitor provides accurate measurements of respiratory rate in triage. *Ann. Emergency Med.* 45, 68–76. doi: 10.1016/j.annemergmed.2004.06.016

Meredith, D. J., Clifton, D., Charlton, P., Brooks, J., Pugh, C. W., and Tarassenko, L. (2012). Photoplethysmographic derivation of respiratory rate: a review of relevant physiology. *J. Med. Eng. Technol.* 36, 1–7. doi: 10.3109/03091902.2011.638965

Nakajima, K., Tamura, T., and Miike, H. (1996). Monitoring of heart and respiratory rates by photoplethysmography using a digital filtering technique. *Med. Eng. Phys.* 18, 365–372. doi: 10.1016/1350-4533(95)00066-6

Rioul, O., and Duhamel, P. (1992). Fast algorithms for discrete and continuous wavelet transforms. *IEEE Trans. Inform. Theory* 38, 569–586. doi: 10.1109/18.119724

Shelley, K. H. (2007). Photoplethysmography: beyond the calculation of arterial oxygen saturation and heart rate. *Anesth. Analg.* 105(Suppl. 6):S31–S36. doi: 10.1213/01.ane.0000269512.82836.c9

Shelley, K. H., Awad, A. A., Stout, R. G., and Silverman, D. G. (2006). The use of joint time frequency analysis to quantify the effect of ventilation on the pulse oximeter waveform. *J. Clin. Monit. Comput.* 20, 81–87. doi: 10.1007/s10877-006-9010-7

Subbe, C. P., Davies, R. G., Williams, E., Rutherford, P., and Gemmell, L. (2003). Effect of introducing the Modified Early Warning score on clinical outcomes, cardio-pulmonary arrests and intensive care utilisation in acute medical admissions. *Anaesthesia* 58, 797–802. doi: 10.1046/j.1365-2044.2003.03258.x

Thakur, G., Brevdo, E., Fuckar, N. S., and Wu, H. T. (2013). The synchrosqueezing algorithm for time-varying spectral analysis: robustness properties and new paleoclimate applications. *Signal Process.* 93, 1079–1094. doi: 10.1016/j.sigpro.2012.11.029

Thayer, J., Iii Sollers, J., Ruiz-Padial, E., and Vila, J. (2002). Estimating respiratory frequency from autoregressive spectral analysis of heart period. *IEEE Eng. Med. Biol. Mag.* 21, 41–45. doi: 10.1109/MEMB.2002.10 32638

Webster, G. W., Webster, J. G., and Webster, R. E. (1997). *Design of Pulse Oximeters*. Abingdon: Taylor & Francis Group.

WHO (2013). *Pocket Book of Hospital Care for Children: Guidelines for the Management of Common Childhood Illnesses*. Geneva: World Health Organization.

Zou, G. Y. (2013). Confidence interval estimation for the bland-altman limits of agreement with multiple observations per individual. *Stat. Methods Med. Res.* 22, 630–642. doi: 10.1177/0962280211402548