ETH zürich

ETH Library

Estimation of Respiratory Rate from Photoplethysmographic Imaging Videos Compared to Pulse Oximetry

Journal Article

Author(s):
Karlen, Walter; Garde, Ainara; Myers, Dorothy; Scheffer, Cornie; Ansermino, J. Mark; Dumont, Guy

Publication date:
2015-07

Permanent link:
https://doi.org/10.3929/ethz-b-000100726

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
IEEE Journal of Biomedical and Health Informatics 19(4), https://doi.org/10.1109/JBHI.2015.2429746

This page was generated automatically upon download from the ETH Zurich Research Collection.
For more information, please consult the Terms of use.

Estimation of Respiratory Rate from Photoplethysmographic Imaging Videos Compared to Pulse Oximetry

Walter Karlen, *Senior Member, IEEE,* Ainara Garde, *Member, IEEE,* Dorothy Myers,
Cornie Scheffer, *Member, IEEE,* J Mark Ansermino, and Guy A Dumont *Fellow, IEEE*

This is the accepted version of the publication:
Karlen, W., Garde, A., Myers, D., Scheffer, C., Ansermino, J. M., Dumont, G. A. (2015). Estimation of Respiratory Rate From Photoplethysmographic Imaging Videos Compared to Pulse Oximetry. IEEE Journal of Biomedical and Health Informatics, 19(4), 1331–1338. http://doi.org/10.1109/JBHI.2015.2429746

***Abstract*—We present a study evaluating two respiratory rate estimation algorithms using videos obtained from placing a finger on the camera lens of a mobile phone. The two algorithms, based on Smart Fusion and Empirical Mode Decomposition (EMD), consist of previously developed signal processing methods to detect features and extract respiratory induced variations in photoplethysmographic signals to estimate respiratory rate. With custom-built software on an Android phone, photoplethysmographic imaging videos were recorded from 19 healthy adults while breathing spontaneously at respiratory rates between 6 to 32 breaths/min. Signals from two pulse oximeters were simultaneously recorded to compare the algorithms' performance using mobile phone data and clinical data. Capnometry was recorded to obtain reference respiratory rates. Two hundred seventy-two recordings were analyzed. The Smart Fusion algorithm reported 39 recordings with insufficient respiratory information from the photoplethysmographic imaging data. Of the 232 remaining recordings, a root mean square error (RMSE) of 6 breaths/min was obtained. The RMSE for the pulse oximeter data was lower at 2.3 breaths/min. RMSE for the EMD method was higher throughout all data sources as, unlike the Smart Fusion, the EMD method did not screen for inconsistent results. The study showed that it is feasible to estimate respiratory rates by placing a finger on a mobile phone camera, but that it becomes increasingly challenging at respiratory rates greater than 20 breaths/min, independent of data source or algorithm tested.**

***Index Terms*—Pulse Oximetry, Respiratory Rate, Photoplethysmographic Imaging, Vital Signs from Video, Empirical Mode Decomposition**

I. Introduction

This work was supported by the Swiss National Science Foundation and Grand Challenges Canada.

W. Karlen is with the Department of Health Sciences and Technology, ETH Zurich, 8092 Zurich, Switzerland. He was with the Electrical & Computer Engineering in Medicine Group, The University of British Columbia (UBC), Canada and the Department of Mechanical and Mechatronic Engineering, University of Stellenbosch, South Africa. Contact: walter.karlen@ieee.org

A. Garde, D. Meyers, G.A. Dumonat and J.M. Ansermino are with the Electrical & Computer Engineering in Medicine Group, Departments of Electrical & Computer Engineering and Anesthesiology, Pharmacology & Therapeutics, UBC, Vancouver, BC, V6T 1Z4, Canada.

C. Scheffer, deceased, was with the Department of Mechanical and Mechatronic Engineering, University of Stellenbosch, 7600 Stellenbosch, South Africa.

CAMERAS embedded on mobile phones allow the monitoring of vital signs based on changes in the recorded light intensity variations [1]. This so-called photoplethysmographic imaging could possibly substitute for traditional pulse oximetry by using an imaging array instead of a single photo detector. While the primary research focus was on the estimation of heart rate using this technique, it has been shown that respiratory rate (RR) can be extracted using non-contact [2]–[4] and contact methods [5], [6]. Non-contact methods are based on the recording of skin color changes visible in the video taken of subjects, whereas contact methods are based on active illumination of the tissue and measurement of variation in the reflected light. Instead of using these photoplethysmographic techniques, cameras of mobile phones have also used motion tracking of chest to estimate RR [7].

RR is an essential vital sign and important criterion for the diagnosis of pneumonia and other respiratory diseases [8]. Abnormal RR is an early sign of critical illness. Therefore, the ability to check multiple vital signs using a camera embedded in a mobile phone, with no additional hardware, would provide a significant advantage for the diagnosis of a number of important clinical conditions.

The waveform obtained through photoplethysmography analysis is called the photoplethysmogram (PPG) and imaging PPG (iPPG) when using an imaging sensor. The PPG and iPPG signals represent blood volume changes in tissue and is modulated by both heart rate and respiration. Respiration modulates the PPG waveform in three ways: 1) The respiratory induced frequency variation (RIFV) - A periodic change in heart rate that is caused by an autonomic nervous system response. The heart rate synchronizes with the respiratory cycle; this is also known as respiratory sinus arrhythmia. 2) Respiratory induced intensity variation (RIIV) - A change in the baseline signal that is caused by a variation of perfusion due to intrathoracic pressure variation. 3) Respiratory induced amplitude variation (RIAV) - A change in pulse strength that is caused by a decrease in cardiac output due to reduced ventricular filling during inspiration.

A number of approaches and algorithms have been proposed to extract RR from the PPG waveform. These methods often target one or multiple respiratory induced variations (RIV)

[9], [10]. RIV can be extracted using wavelet decomposition [10], digital filters [11] , Fourier transforms [12], complex demodulation [13], and auto-regression [14]. In previous work, we have proposed multiple approaches to detect RR from the PPG signal [15]–[17]. We have demonstrated that the RR obtained from all three RIV independently can be combined using a Smart Fusion algorithm [17]. This last approach was the most suitable when dealing with short PPG signals [15]. In [16] we have used Empirical Mode decomposition (EMD) to extract linear (RIIV) and non-linear (RIFV) respiratory components on the PPG signal. EMD was not sensitive to the non-linear RIAV component. EMD was introduced by Huan et al. and is particularly suitable for analyzing non-linear and non-stationary signals [18]. It is an adaptive decomposition technique based on the local properties of the signal to derive its basis functions. It has been successfully applied to reduce motion artifacts in the PPG signal [19], extract RR from electrocardiograms (ECG) [20], and decompose respiratory sounds [21].

In previous studies, we developed algorithms to extract the optimal iPPG signals from the videos recorded from the Camera Oximeter. We implemented an algorithm that automatically detects the correct finger placement on the camera lens [22] and another algorithm that automatically identifies the optimal region of interest (ROI) from photoplethysmographic imaging [23]. The combination of these novel algorithms provides a suitable method to extract the optimal iPPG signals for subsequent analysis. Therefore, the aim of this study was to estimate RR with the Smart Fusion and EMD-based algorithms and assess their accuracy using the reference RR. We have compared the performance of the two algorithms using iPPG recorded from the Camera Oximeter, and PPG signals recorded from commercially available pulse oximeters. A preliminary version of the Smart Fusion comparison has been previously reported [6].

II. Methods

A retrospective analysis of photoplethysmographic imaging recordings was performed in Matlab (Mathworks Inc, Natick, USA) to test and compare our RR extraction algorithms for embedded use on a mobile phone.

A. Experimental Setup

After obtaining Health Canada and institutional ethics approval and written informed consent, 19 healthy non-smoking subjects (10 males, 9 females, median age 28 (range 19 to 50) years, median Fitzpatrick Skin Phototype III (range II to V) [24]) with no history of cardio-respiratory disease were recruited for a controlled hypoxia study. The primary aim of this study was to calibrate a photoplethysmographic imaging oximeter on a mobile phone (Camera Oximeter) at various oxygenation levels [1]. The setup of this study was ideal for our experiments since a secondary effect of exposure to hypoxic air is an involuntary physiological response accompanied by an increase in RR. This allowed us to collect data at higher RRs, which would only typically be available with controlled breathing (e.g. by using a metronome). After a health check, two Phone Oximeters [25] were placed on the non-dominant hand as reference oximeter measurements. These reference devices collected clinical PPG from two independent, Federal Drug Administration (FDA) approved sensors to compare the performance of RR extraction algorithms between high end medical devices and mobile phone based data. The Phone Oximeters connected an iPod touch 4th generation (Apple Inc, Cuppertino, USA) to either a Nonin (PON) Xpod (Nonin Medical Inc., Plymouth, USA) or to a Masimo (POM) low-power module (Masimo Corp., Irvine, USA) pulse oximeter. Both systems were configured to record the PPG signal with 16 bit resolution and with a basic bandpass filter enabled. The PPG signal of PON was recorded at a sampling rate of 75 Hz and POM at a sampling rate of 62.5 Hz.

Recording started at sea level with an inspired oxygen concentration (FiO_2) of 21%. The subjects then entered a normobaric hypoxia chamber with the FiO_2 set to 12%. The FiO_2 was then increased step-wise to 17% and then reduced back to 12%. When an FiO_2 of 12% was attained, the subjects exited the chamber and were monitored again at an FiO_2 of 21% (Figure 1). At regular intervals during the experiment at given FiO_2 levels (21%, 12%, 13%, 14.5%, 16%, 17% and reverse), the subjects performed a recording with the Camera Oximeter. The recording consisted of placing the camera of a low-cost mobile phone (Samsung Galaxy Ace) on the index finger of the dominant hand. A custom software application (OxiCam) was launched. Once the finger was correctly detected using an automatic algorithm previously developed and validated [22], OxiCam recorded a video file for 60 s. The video format was set to 240 x 320 pixels resolution (QVGA) at a frame rate (sampling rate) of 20 Hz. The white balance was set to "incandescent", as this has been shown to be the optimal configuration for this type of camera [23]. During the Camera Oximeter recording, respiratory activity was recorded using a face mask connected to a Datex-Ohmeda S5 Collect capnography device recording flow, CO_2 and O_2 at 100 Hz. The reference RR was extracted from the capnogram using an automated algorithm counting the number of breaths within the 60 s recordings. This count was manually validated by an expert using the flow signal. Recordings with incorrect counts due to artifacts in the capnogram (e.g. calibration process) were excluded. Recordings with RR lower than 6 breaths/min (containing episodes of apnea) and with RR higher than 40 breaths/min were considered outliers for this data set and therefore excluded. All recording devices (POM, PON, Camera Oximeter and Datex-Ohmeda) were synchronized to a common time server and a marker was pressed on the Datex-Ohmeda and Phone Oximeter systems at the beginning of each Camera Oximeter recording to verify synchronization.

B. iPPG Signal Extraction Algorithm

Three steps were necessary to obtain a iPPG waveform from the Camera Oximeter video recordings (Figure 2).

1) Video Recordings: The video recordings from the hypoxia experiments were imported to Matlab in an RGB format and paired with the reference RR obtained from capnometry.

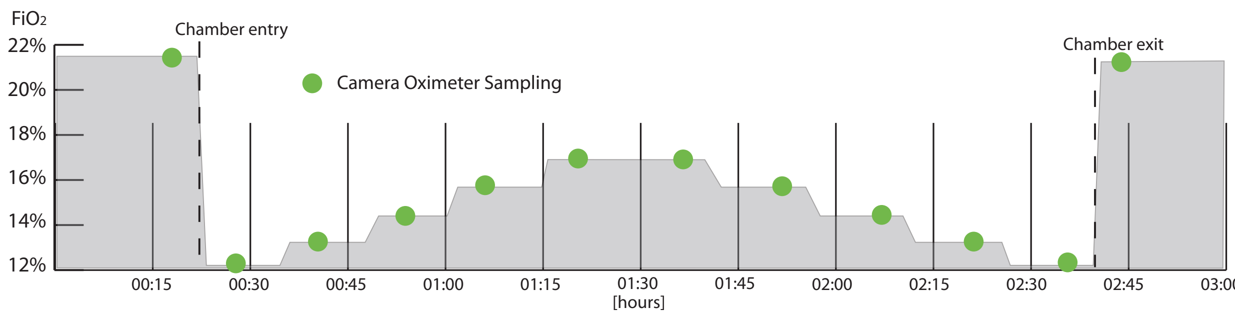

Fig. 1. Experimental setup with changing FiO_2 input. The green dots represent time points when sampling with the Camera Oximeter was performedand reference respiratory rate was measured using a capnometer. Reference pulse oximetry was recorded throughout the entire experiment.

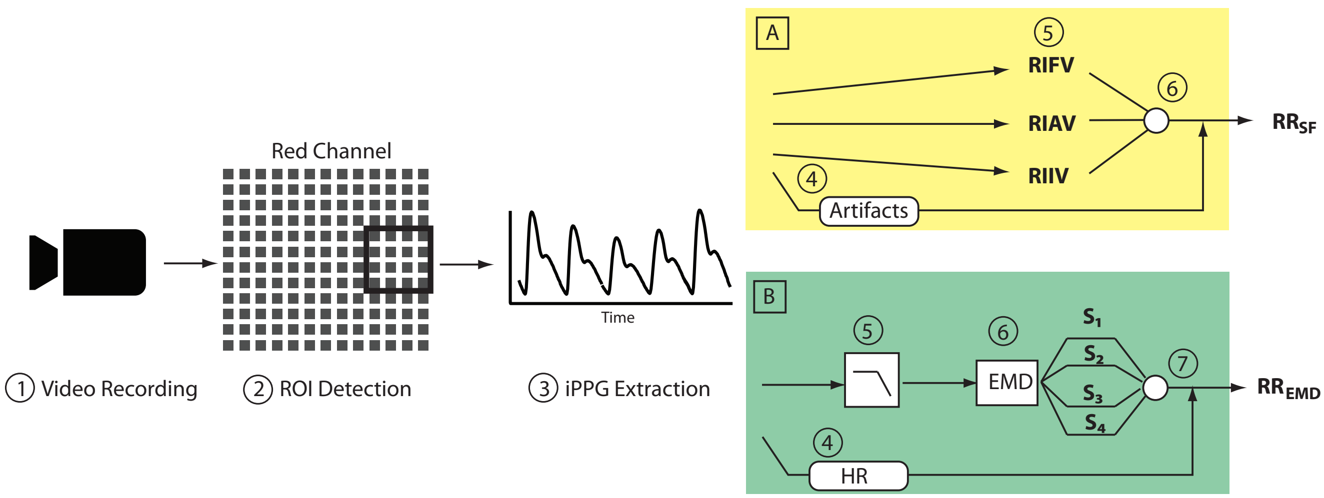

Fig. 2. Schematic of the algorithms used to extract respiratory rate (RR) from photoplethysmographic imaging with a mobile phone camera. 1) A video is recorded with the camera once a finger is correctly placed on the lens; 2) The optimal region of interest (ROI) is detected from the red channel of the video; 3) The imaging photoplethysmogram (iPPG) is extracted from this ROI, and the position and amplitude of pulses calculated; A4) Artifacts are detected and labeled as such in the PPG signal ; A5) Respiratory induced variations (RIV) are computed; A6) A Smart Fusion process merges the 3 RIV components, compares agreement and excludes artifacts for the calculation of RR; B4) calculation of heart rate (HR); B5) Low-pass filter; B6) Empirical mode decomposition (EMD) and calculation of maximal energy and frequencies for each decomposed signal; and B7) Frequency with higher energy below 0.3 times the HR range are selected as for RR.

2) ROI Detection: Optimal ROI for the red video channel was determined using the algorithm described in [23]. The ROI area was increased to 40 pixels^2 to render the selection process more efficient. If no suitable ROI was found, the recording was considered to be of poor quality and no further processing was conducted.

3) PPG Extraction: The iPPG signal was extracted from the ROI by averaging all pixel intensities within the ROI. The iPPG signal was band-pass filtered using a 5th order Butterworth filter with cut-off frequencies at 0.08 Hz and 3 Hz. Then the incremental merge segmenting (IMS) algorithm was used to validate the iPPG [26]. The IMS divided the iPPG into small segments of fixed length. Segments were then iteratively combined to larger segments if the lines through the endpoints of adjacent segments shared similar slopes. This resulted in a discretization of the iPPG waveform into main features such as diastole, systole and, if present, dichrotic notch. From these features, the IMS algorithm automatically detected pulse peaks, amplitudes and artifacts.

Two distinct approaches were used to extract RR from the iPPG signal using the Smart Fusion (Figure 2 A) and the EMD-based (Figure 2 B) algorithms.

C. Smart Fusion Algorithm

1) Artifact Detection: Artifacts in the iPPG signal were automatically identified within the IMS algorithm. Pulses with amplitudes exceeding the lower and upper adaptive thresholds were labeled as artifacts [26]. In addition, artifacts were identified by scanning for abnormal pulse intervals outside of the normal range, defined from 230 to 2400 ms.

2) RIV Extraction: The 3 RIV components RIFV, RIIV, and RIAV were extracted from the PPG signal. The spectral power of each component was calculated and the frequency with the maximal power within the expected RR range was extracted. The expected RR range was adaptively determined using the heart rate, such as RR_{min} = HR/11 and RR_{max}=HR/2.2 where HR is the heart rate in beats/min. These rules were empirically obtained by analyzing the range of ratios between HR and RR in the CapnoBase dataset used in [17]. The spectral power was calculated using a Fast Fourier Transform (FFT) with a sliding window of 16 s length and 3 s time steps. Consequently, 3 independent RIV estimations were obtained 15 times in each 60 s recording.

3) Smart Fusion of RR: RR estimation was determined by fusing the respiratory frequencies obtained from RIFV, RIIV, and RIAV by calculating their mean [17], such as

$$RR_F = (RR_{RIFV} + RR_{RIAV} + RR_{RIIV})/3.$$

The quality of this fusion was evaluated by comparing the variance of 3 independent estimations. RR_F estimations with variances higher than 16 breaths^2/min^2 were considered unreliable and discarded. In addition, estimations from the 16 s sliding windows containing 3 or more artifacts were also discarded. The median RR of the remaining RR_F estimations within a recording was calculated and used as the final RR estimation (RR_{SF}). If less than 4 of 15 estimations remained throughout the 60 s recording, the recording was considered invalid and no RR_{SF} was reported.

D. EMD-based Algorithm

EMD decomposes non-stationary and non-linear signals into a set of mono-component signals called intrinsic mode functions (IMF). Each IMF function represents an oscillation mode embedded in the signal and must satisfy two conditions; 1) the number of extrema and zero crossings must be either equal or differ by one; and 2) the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. In a process that is not dissimilar to wavelet analysis, EMD decomposes the signal into different resolution scales. However, in EMD the basis functions are directly extracted from the data, while in wavelet analysis, a pre-designed mother wavelet, which determines the basis functions for the different scales, is selected before the analysis. Therefore, EMD can better represent the local characteristics of a signal and adapt to oscillation patterns over time. Given a PPG waveform $x(t)$ the EMD decomposition can be performed by:

1) Extracting the envelopes defined by the local maxima and minima separately. All the local maxima are connected by a cubic spline line to create the upper envelope (e_{max}). The same process is followed with the local minima to create the lower envelope (e_{min}).
2) Computing their mean designated as $m_1 = (e_{max} + e_{min})/2$.
3) Extracting the first component $h_1 = x(t) - m_1$.
4) If h_1 does not satisfy the IMF conditions, repeat steps (1 to 3) treating h_1 as data, $h_{11} = h1 - m_{11}$.
5) Repeating this $sifting$ process k times until h_{1k} is an IMF, that is $h_{1k} = h_{1(1-k)} - m_{1k}$. Then $c_1 = h_{1k}$ is designated as the first IMF of the data.
6) Separating c_1 (which contains the finest scale or highest frequency component of the data) from the rest of the data $r_1 = x(t) - c_1$, and apply the same $sifting$ process to the residue r_1.

The sifting process is stopped following a predetermined criteria, such as when either the IMF or the residue, becomes so small that it is less than the predetermined value, or when the residue becomes a monotonic function from which no more IMF can be extracted.

EMD-based RR estimation was performed through the following steps (Figure 2 B):

1) Calculation of Heart Rate: The spectral power of the iPPG signal was computed from a FFT, using a sliding window of 30 s with 10 s time steps. The frequency with the maximal power corresponded to the cardiac frequency and was therefore extracted as heart rate.

2) Filtering of Heart Rate: Each 30 s iPPG sliding window segment was filtered by a low pass filter (cut-off frequency of 1 Hz) to remove the cardiac component.

3) EMD: The filtered segments were decomposed into four intrinsic mode functions and their residual signal.

4) RR Selection: The spectral power of each decomposed component was calculated, and the frequency peak with the highest power in the spectral domain was extracted for each intrinsic mode function and from the residual signal. The frequency peaks located close to the heart rate ($f \geq f_{HR} - f_{HR} * 0.3$) were not considered. The frequency peak with the highest power among all intrinsic mode functions and residual signal reflected the predominant low frequency modulation of the PPG signal and was extracted as the RR estimation (RR_{EMD}).

E. Analysis

Data obtained from the first subject (training subject) was used to adjust the algorithm parameters such as IMS segment length and parameters for artifact threshold adaptation as well as EMD-based configuration. The data from the remaining 18 subjects were used to test the algorithm and compare the performance. Bland-Altman plots [27] were created to compare the three PPG data sources against the reference RR extracted from the capnogram for the Smart Fusion and EMD-based algorithms. Bias and limits of agreement for repeated measurements were calculated using the pairwise method for which the within-subject and between-subject mean square errors were calculated using one way anova [28]. In addition, the unnormalized root mean square error (RMSE) (breaths/min) was calculated, such as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i^{ref} - x_i^{co})^2}$$

where n is the number of observations and x^{ref} and x^{ox} are the reference and the oximeter observations respectively.

III. Results

A total of 350 Camera Oximeter recordings were obtained (18 from the training subject). Sixty were excluded because the reference RR was not available or outside the specified range or no PPG from POM or PON was available. There were 272 recordings remaining to test the algorithms. The median heart rate was 75 (range 43 to 102) beats/min and the median reference RR obtained by capnometry was 14 (range 6 to 32) breaths/min. Of these, the ROI selection process of the Camera Oximeter discarded 34 cases that contained recordings of too poor quality to extract a reliable PPG. The remaining 238 cases were available for comparison with the reference RR (Figure 3). An example of extracted PPG waveforms can be observed in Figure 4.

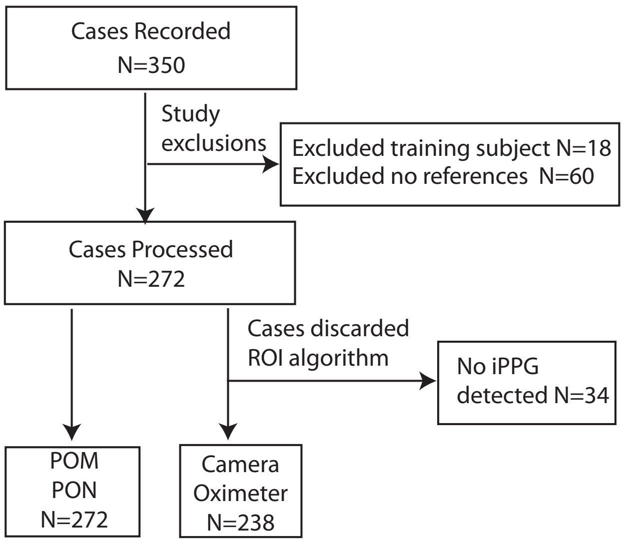

Fig. 3. Flowchart describing how recorded videos were excluded and removed from analysis.

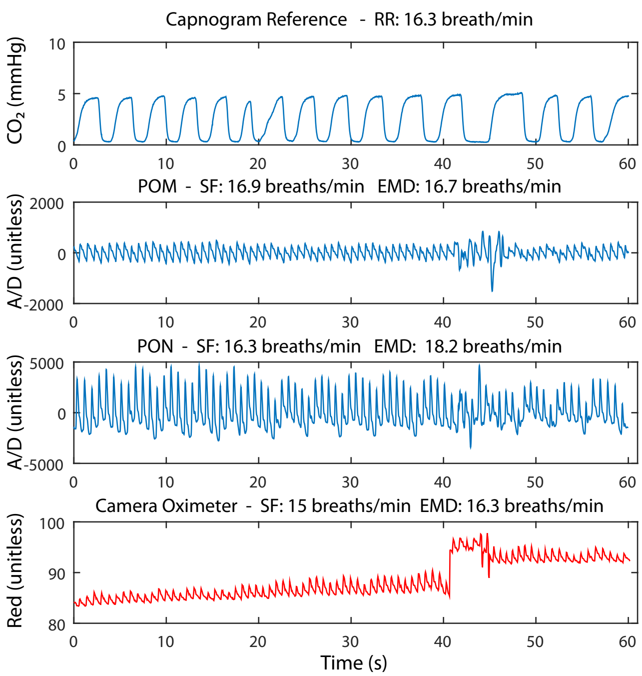

Fig. 4. Comparison of raw waveforms for one case. Reference capnogram waveform (top); waveforms from medical grade pulse oximeters from Masimo (POM) and Nonin (PON); waveform obtained through ROI process of red channel from Camera Oximeter (bottom).

A. Estimation Error

The RMSE, bias and limits of agreement are shown in the Bland-Altman plots for both algorithms and each data source of Figure 5 . The largest RMSE and bias was obtained for the POM - EMD combination. The Smart Fusion process did not report RR for 99 (POM), 55 (PON), and 39 (Camera Oximeter) cases due to too many artifacts in the recording or no agreement between the 3 RIV found for the entire recording (yellow small circles in Figure 5). RR_{EMD} was reported for all 272 analyzed cases. For comparison purposes, the same cases that the Smart Fusion eliminated were also depicted in yellow for the EMD-based results.

The eliminated estimations by the Smart Fusion algorithm had significantly larger absolute errors for all data sources (Student's t-test $p<0.01$). No such difference could be observed when comparing errors of the EMD-based algorithm.

RR_{SF} estimation had significantly lower absolute errors when using pulse oximeter PPG compared to the Camera Oximeter (Student's t-test $p<0.01$). Consequently the RMSE was also lower for PON (2.27 breaths/min) and POM (2.29 breaths/min) than for the Camera Oximeter (6.01 breaths/min) estimation.

B. Computational Efficiency

The proposed algorithms were sufficiently efficient to be implemented on a mobile phone app. Steps 1 to 3 of Figure 2 were implemented for data collection in OxiCam and ran without effort on a low range Android phone (Android v2.3, single-core CPU 800 MHz, 278 MB RAM). The calculation of RR (Figure 2, sub-calculation A and B) was not implemented on the app. Sub-calculation A4 to A6 (Smart Fusion) required the resampling of the PPG pulse signals to 4 Hz and processing of 3 FFTs of 16 s windows. As the FFT calculation is a native function in mobile operating systems (mostly for voice/music operations), this is no significant computational effort. Sub-calculation B (EMD) is computationally more extensive. However, compared to other proposed techniques in the literature, EMD is a very simple and easy to implement method. It relies on a sifting algorithm which can introduce delays depending on the complexity of the signal.

IV. Discussion

In this study, we showed that it is feasible to estimate RR by placing an adult finger on a mobile phone camera. This contact photoplethysmographic imaging method became increasingly challenging at respiratory rates higher than 20 breaths/min. The algorithms used have been previously tested against the same dataset of PPG signals obtained from standard pulse oximetry [29] and shown to be efficient for detecting pulses and artifacts [26], as well as for estimating RR [16], [17]. The studied algorithms showed to be computationally efficient, facilitating real-time processing on a mobile phone. We applied these algorithms to the iPPG obtained from a phone camera and the PPGs of two standard pulse oximeters. For the Smart Fusion algorithm we observed a RMSE of 6 breaths/min for the Camera Oximeter iPPG compared to 2.3 breaths/min obtained by the pulse oximeter PPGs. Independent of data source, we observed greater errors for the Smart Fusion at RRs above 20 breaths/min, with one iPPG observation having an error of 23 breaths/min at a reference RR of 33 breaths/min. Our dataset contains too few recordings (36 cases) with RRs above 20 breaths/min to perform a conclusive analysis. However, such large errors have not been previously observed when studying RR estimation using photoplethysmographic imaging. Poh et al. reported an RMSE of 1.3 breaths/min for non-contact RR estimation, but this low number was obtained from only 12 samples whose reference RR only ranged from 19 to 21 breaths/min [3]. Similarly, the study in [5] analyzed

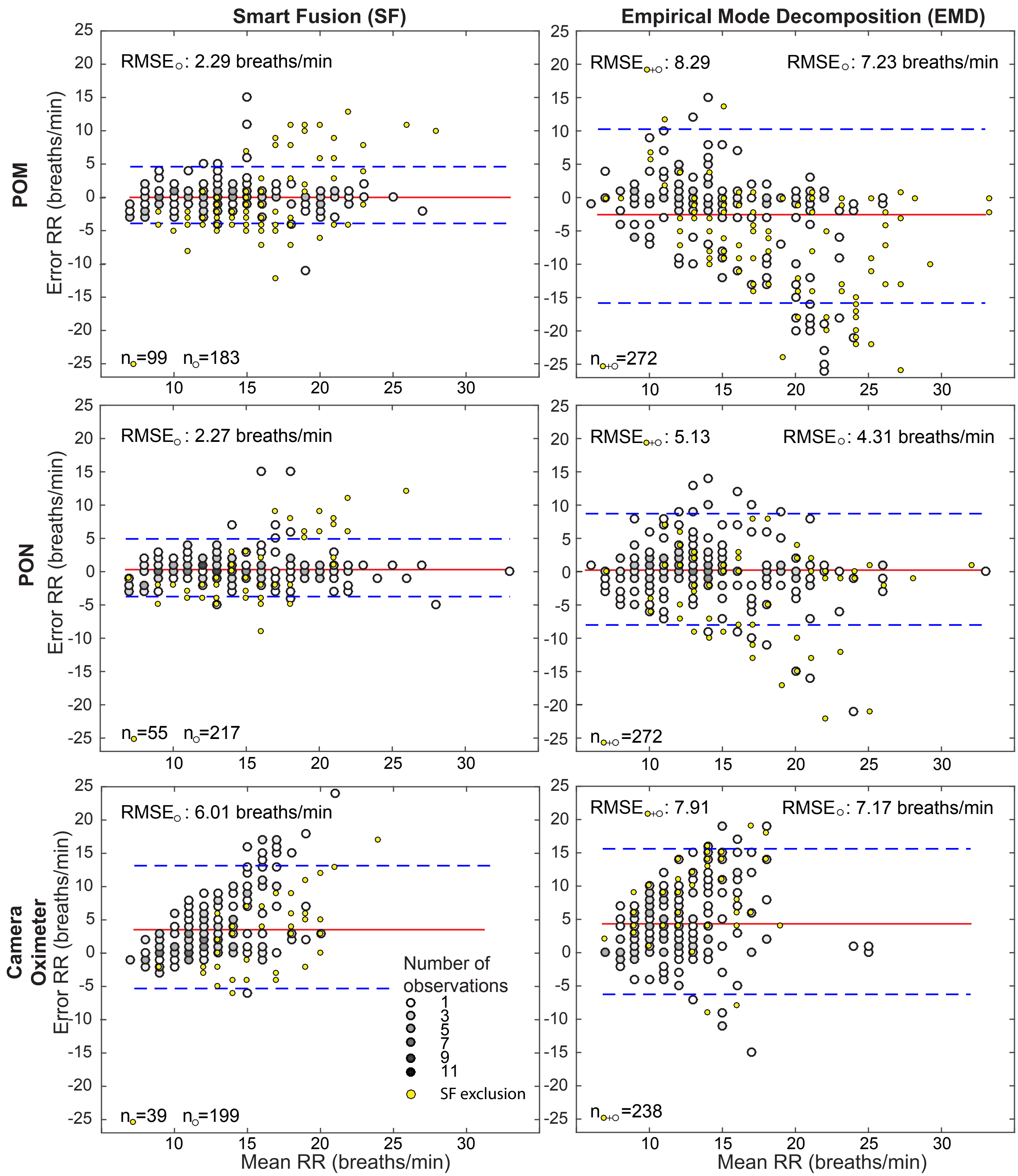

Fig. 5. Bland-Altman plots of the estimated RR using the Smart Fusion (SF, left) and Empirical Mode Decomposition (EMD, right) based methods on the PPG obtained from the Masimo (POM, top) and Nonin (PON, middle) pulse oximeters, as well as the iPPG from the Camera Oximeter (bottom). The root mean square error (RMSE) is calculated against the reference RR obtained from counting breaths in the capnogram. The solid red line corresponds to the mean error (bias) and the dashed blue lines are 98% limits of agreement. The smaller yellow dots correspond to the cases eliminated by the SF algorithm.

a single subject that was breathing to a metronome in the 12 to 24 breaths/min range, forcing regular breathing over 2 min. This study did not directly report accuracy for RR, but graphically displayed good agreement.

The Smart Fusion algorithm eliminated unreliable recordings during 2 distinct processing steps. The ROI selection process eliminated recordings with poor quality signals due to low signal to noise ratios. The Smart Fusion process eliminated further recordings that had too many artifacts or were too challenging to obtain consistent RR with the 3 RIV. While the Smart Fusion successfully eliminated almost all estimations with larger errors for the POM and PON, the data from the Camera Oximeter could not be entirely cleaned. The medical devices produced cleaner and higher quality signals, with a higher signal-to-noise ratio and at triple sampling rates compared to the photoplethysmographic imaging with a low-cost phone camera (Figure 4). It is therefore not surprising that additional confounding factors were present in the iPPG signal of the Camera Oximeter. For example, movement artifacts lead to large baseline shifts in the Camera Oximeter signal, whereas in the POM and PON it was successfully suppressed (Figure 4). Therefore we suspect that internal dynamics of the camera also play a role in the poorer performance. Another confounding factor is the presence of low frequency (0̃.1 Hz) components, such as Mayer waves [30]. In our previous work the presence of this high power component contributed to estimation errors [17] that coincided throughout all respiratory induced variations and could not be eliminated by the Smart Fusion. Since the algorithm prioritizes the frequency band with highest energy for selecting RR, the large RR error cannot be avoided. Analysis of our cases with high RR and large error revealed that many of these cases also contained high power in lower frequencies (Figure 6). However, in the current work we introduced an adaptive threshold on the RR limits based on HR. This approach allowed us to reduce the influence of Mayer waves on the computation of the Smart Fusion RR which was efficient on the PPG signals, but less so on the iPPG. Therefore we suspect that factors specific to the hardware and not physiology such as Mayer waves were influencing the iPPG. Interestingly, the EMD-based algorithm that used a longer window of 30 s was less susceptible to this error, but did overestimate the RR more frequently, which can be seen as an increase in the negative errors in the Bland-Altman plots. This can be partially explained that the EMD primarily captures the RIIV of a PPG signal, but is also sensitive to RIFV. We observed that often RIFV had higher power at harmonics rather than at the RR frequency which lead to overestimations when the power of the correct RIIV estimation was weak. Furthermore, the EMD-based algorithm did not rely on elimination of RR estimations that contained artifacts or mismatches between different respiratory components. Consequently, the RMSE was higher compared to the Smart Fusion algorithm.

The number of cases for which the Smart Fusion does not provide a reading can have an impact on the usability of the system. While the exclusion of cases due to poor iPPG signal is independent of processing method, the Smart Fusion shows an additional large elimination rate (POM 36.4%; PON 20.2%; Camera Oximeter 26.8%). The elimination rate correlates with the RMSE obtained with the EMD algorithm. Together with the observed significantly lower absolute errors, this would suggest that indeed the poor performing estimations are eliminated. To improve usability, maintain reliability, and reduce the elimination rate the complete algorithm could be implemented onto the mobile phone application by providing

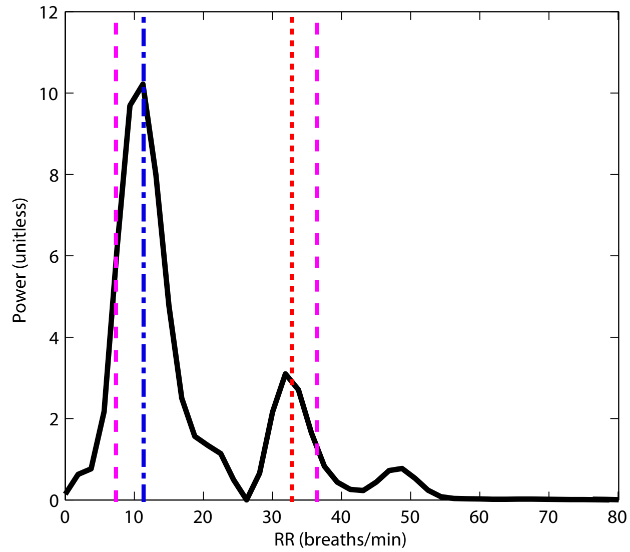

Fig. 6. Example of high power in a low respiratory rate (RR) band where reference RR is high (dotted line). The dashed lines represent the RR selection range, the dot-dash line is the selected RR by the algorithm.

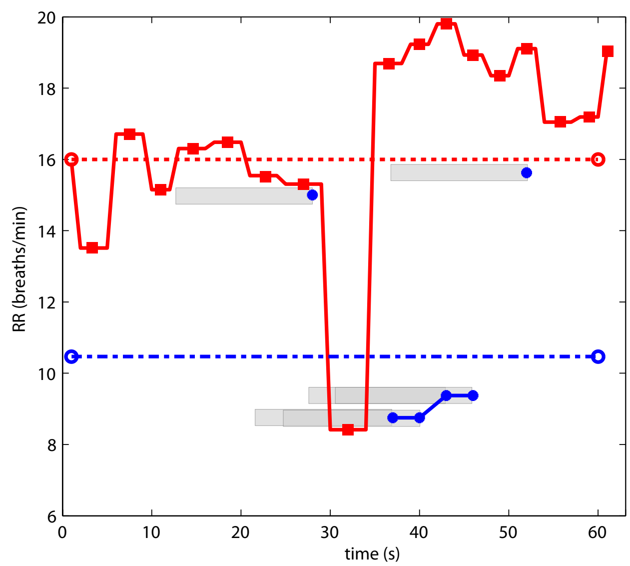

Fig. 7. Example of high variation in respiratory rate (RR) within 60 s. The estimated RR (dot-dash blue) is lower than the reference RR (dotted red) as valid estimation was primarily found for a window where RR was low. The window size was 16 s (grey boxes). Instantaneous RRs are shown as filled squares (reference) and circles (Camera Oximeter).

real-time feedback on acquisition. If no RR is obtained after 60 s, the recording could be extended until the quality of the PPG signal improves.

A further confounding factor is the large variability in the RR throughout the measured 60 s. Our experiment allowed the subjects to breathe freely. The spontaneous RR could therefore vary greatly, but the reference RR would be limited to a single value. The Camera Oximeter on the other hand, would calculate RR for 16 s windows and not all windows would not necessarily contribute to the final calculation of RR due to the exclusion of measurements during the Smart Fusion (Figure 7). This limitation could be overcome by either comparing to the instantaneous RR or estimating RR with larger FFT windows. In the EMD-based method we used 30 s windows, but could not achieve a reduction in error with this window size. The EMD-based method used windows of 60 s or 120 s in previous studies [15] that showed better results. Therefore larger window sized are recommended. However, increasing the window size would reduce the number of available recordings free from artifacts and would result in longer acquisition times making the Smart Fusion approach less practical.

A limitation of our study was is that only healthy adults were evaluated. Elderly, severely sick patients and children present additional challenges that have not been investigated in this study. For example, the presence of respiratory sinus arrhythmia is reduced. This could prevent the Smart Fusion algorithm of finding an agreement between the three RR estimations. We also recommend validating other photoplethysmographic imaging algorithms with supplemental data containing breathing at higher RR. Our current results indicate that RR estimation from iPPG would not be applicable to diagnostic screening in children who have faster breathing rates.

We conclude that high frequency and spontaneous breathing might be challenging to detect using a low-end camera on a mobile phone. However, the observed low RMSE obtained with the Smart Fusion algorithm on the pulse oximeter data is very promising and such implementation will be pursued.

Acknowledgment

The authors thank the Global Engineering Team 2012 for designing the OxiCam application. We thank Drs. Michael Koehle and Jim Rupert and members of their labs at UBC for kindly allowing the use of the hypoxia chamber for this study. Aryannah Umedaly, Drs. Heng Gan and Chris Petersen, and the Pediatric Anesthesia Research Team assisted with the data collection. We are grateful to all subjects who accepted to participate in this study.

References

[1] W. Karlen, J. Lim, J. M. Ansermino, G. Dumont, and C. Scheffer, "Design challenges for camera oximetry on a mobile phone," in *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Jan. 2012, pp. 2448–51.

[2] W. Verkruysse and L. Svaasand, "Remote plethysmographic imaging using ambient light," *Opt. Express*, vol. 16, pp. 21 434–45, 2008.

[3] M.-Z. Poh, D. J. McDuff, and R. W. Picard, "Advancements in noncontact, multiparameter physiological measurements using a webcam." *IEEE Trans. Biomed. Eng.*, vol. 58, no. 1, pp. 7–11, Jan. 2011.

[4] M. Bartula, T. Tigges, and J. Muehlsteff, "Camera-based System for Contactless Monitoring of Respiration," in *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2013, pp. 2672–5.

[5] C. G. Scully, J. Lee, J. Meyer, A. M. Gorbach, D. Granquist-Fraser, Y. Mendelson, and K. H. Chon, "Physiological parameter monitoring from optical recordings with a mobile phone." *IEEE Trans. Biomed. Eng.*, vol. 59, no. 2, pp. 303–6, Feb. 2012.

[6] W. Karlen, A. Garde, D. Myers, C. Scheffer, J. M. Ansermino, and G. A. Dumont, "Respiratory Rate Assessment from Photoplethysmographic Imaging," in *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2014, pp. 5397–400.

[7] H.-y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian video magnification for revealing subtle changes in the world," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 1–8, Jul. 2012.

[8] WHO, *Pocket Book of Hospital Care for Children - Guidelines for the Management of Common Illnesses with Limited Resources*, Geneva, CH, 2005.

[9] A. Johansson, "Neural network for photoplethysmographic respiratory rate monitoring," *Med. Biol. Eng. Comput.*, vol. 41, no. 3, pp. 242–8, May 2003.

[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. Monit. Comput.*, vol. 18, no. 5-6, pp. 309–12, Dec. 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. Monit. Comput.*, 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, Aug. 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] A. Garde, W. Karlen, J. M. Ansermino, and G. A. Dumont, "Estimating respiratory and heart rates from the correntropy spectral density of the photoplethysmogram." *PLoS One*, vol. 9, no. 1, p. e86427, Jan. 2014.

[16] A. Garde, W. Karlen, P. Dehkordi, J. M. Ansermino, and G. A. Dumont, "Empirical mode decomposition for respiratory and heart rate estimation from the photoplethysmogram," in *Comput. Cardiol.*, A. Murray, Ed., vol. 40. Zaragoza: IEEE, Sep. 2013, pp. 799–802.

[17] W. Karlen, S. Raman, J. M. Ansermino, and G. A. Dumont, "Multiparameter respiratory rate estimation from the photoplethysmogram." *IEEE Trans. Biomed. Eng.*, vol. 60, no. 7, pp. 1946–53, Jul. 2013.

[18] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. London*, vol. 454, no. 1971, pp. 903–95, 1998.

[19] Q. Wang, P. Yang, and Y. Zhang, "Artifact reduction based on Empirical Mode Decomposition (EMD) in photoplethysmography for pulse rate detection." in *Proc. IEEE Conf. Eng. Med. Biol.*, vol. 2010, Jan. 2010, pp. 959–62.

[20] R. Balocchi, D. Menicucci, E. Santarcangelo, L. Sebastiani, a. Gemignani, B. Ghelarducci, and M. Varanini, "Deriving the respiratory sinus arrhythmia from the heartbeat time series using empirical mode decomposition," *Chaos, Solitons & Fractals*, vol. 20, pp. 171–7, 2004.

[21] M. Lozano, J. A. Fiz, and R. Jané, "Estimation of Instantaneous Frequency from Empirical Mode Decomposition on Respiratory Sounds Analysis," in *Proc. IEEE Conf. Eng. Med. Biol.*, 2013, pp. 981–4.

[22] W. Karlen, J. Lim, J. M. Ansermino, G. A. Dumont, and C. Scheffer, "Recognition of correct finger placement for photoplethysmographic imaging," in *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Osaka, Jul. 2013, pp. 7480–3.

[23] W. Karlen, J. M. Ansermino, G. A. Dumont, and C. Scheffer, "Detection of the optimal region of interest for camera oximetry," in *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Osaka, Jul. 2013, pp. 2263–6.

[24] T. Fitzpatrick, S.-r. S. Types, and I. T. Vi, "The validity and practicality of sun-reactive skin types I through VI," *Arch. Dermatol.*, vol. 124, no. 6, pp. 869–71, 1988.

[25] J. Hudson, S. M. Nguku, J. Sleiman, W. Karlen, G. A. Dumont, C. Petersen, C. B. Warriner, and J. M. Ansermino, "Usability testing of a prototype Phone Oximeter with healthcare providers in high- and low-medical resource environments," *Anaesthesia*, vol. 67, no. 9, pp. 957–67, 2012.

[26] W. Karlen, J. M. Ansermino, and G. A. Dumont, "Adaptive Pulse Segmentation and Artifact Detection in Photoplethysmography for Mobile Applications," in *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, San Diego, Aug. 2012, pp. 3131–4.

[27] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement." *Lancet*, vol. 1, no. 8476, pp. 307–10, Feb. 1986.

[28] ——, "Measuring agreement in method comparison studies." *Stat. Methods Med. Res.*, vol. 8, no. 2, pp. 135–60, 1999.

[29] 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 *Annu. Meet. Soc. Technol. Anesth.*, West Palm Beach, 2010, p. 25. [Online]. Available: www.capnobase.org

[30] C. Julien, "The enigma of Mayer waves: Facts and models." *Cardiovascular research*, vol. 70, no. 1, pp. 12–21, Apr. 2006.