Respiratory Rate Assessment from Photoplethysmographic Imaging

Walter Karlen[1,2], *Member, IEEE,* Ainara Garde[1], *Member, IEEE,* Dorothy Myers[1], Cornie Scheffer[2], *Member, IEEE,* J Mark Ansermino[1], and Guy A Dumont[1] *Fellow, IEEE*

***Abstract*— We present a study investigating the suitability of a respiratory rate estimation algorithm applied to photoplethysmographic imaging on a mobile phone. The algorithm consists of a cascade of previously developed signal processing methods to detect features and extract respiratory induced variations in photoplethysmogram signals to estimate respiratory rate. With custom-built software on an Android phone (Camera Oximeter), contact photoplethysmographic imaging videos were recorded using the integrated camera from 19 healthy adults breathing spontaneously at respiratory rates between 6 and 40 breaths/min. Capnometry was simultaneously recorded to obtain reference respiratory rates. Two hundred and ninety-eight Camera Oximeter recordings were available for analysis. The algorithm detected 22 recordings with poor photoplethysmogram quality and 46 recordings with insufficient respiratory information. Of the 232 remaining recordings, a root mean square error of 5.9 breaths/min and a median absolute error of 2.3 breaths/min was obtained. 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 higher than 20 breaths/min.**

I. Introduction

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 replaces 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]. Non-contact methods are based on the recording of skin color changes of subjects visible in the video, whereas contact methods are based on active illumination of the tissue and measurement of variation in the reflected light.

RR is an essential vital sign and important criterion for the diagnosis of pneumonia and other respiratory diseases [6]. Abnormal RR is an early sign of critical illness. Therefore, the ability to check multiple vital signs using a camera embedded into a mobile phone with no additional hardware is desirable.

The waveform obtained through photoplethysmographic imaging analysis is called the photoplethysmogram (PPG). The PPG signal represents 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.

Various approaches and algorithms have been proposed to extract RR from the PPG. These methods often target one or multiple respiratory induced variations (RIV) [7], [8]. RIV can be extracted using wavelet decomposition [8], digital filters [9], Fourier transforms [10], complex demodulation [11], and auto-regression [12]. In previous work, we have proposed multiple approaches to detect RR from PPG [13]–[15]. We have demonstrated that the RR obtained from all three RIV independently can be combined using a smart fusion approach [15]. This last approach was the most suitable when dealing with short PPG signals [13]. Other preparatory work consisted of the automatic detection of correct finger placement on the camera lens [16] and the automatic detection of the optimal region of interest (ROI) for photoplethysmographic imaging with the Camera Oximeter [17].

The combination of these novel algorithms provides a suitable method to extract the best possible PPG signals from recorded videos for subsequent analysis. The aim of this study was to estimate RR with these algorithms and assess the accuracy of the calculated RR.

II. Methods

A retrospective analysis of photoplethysmographic imaging recordings was performed in Matlab (Mathworks Inc, Natick, USA) to test our RR extraction algorithm 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, mean age 30.7 $\pm$ 8.7 years) 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

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

[1] Electrical & Computer Engineering in Medicine Group, Departments of Electrical & Computer Engineering and Anesthesiology, Pharmacology & Therapeutics, The University of British Columbia (UBC), 2332 Main Mall, Vancouver, BC, V6T 1Z4, Canada. Contact: walter.karlen@ieee.org

[2]Department of Mechanical and Mechatronic Engineering, University of Stellenbosch, 7600, South Africa.

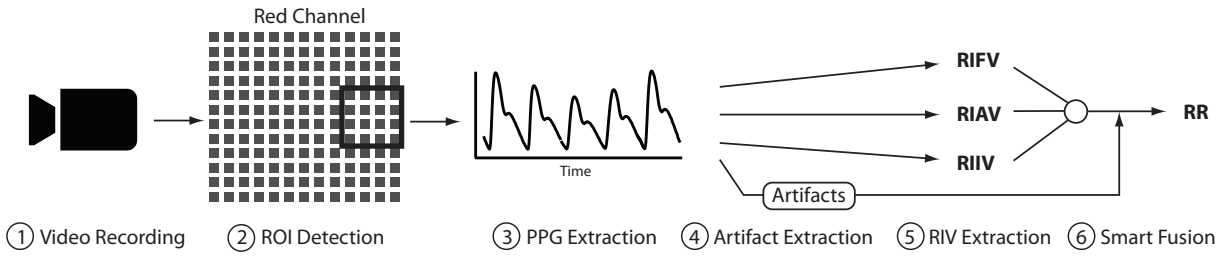

Fig. 1. Schematic of the algorithms used to extract respiratory rate (RR) from photoplethysmographic imaging with a mobile phone camera. 1) A video is recorded with the camera once a finger is correctly placed on the lens; 2) The optimal region of interest (ROI) is detected from the red channel of the video; 3) The photophlethysmogram (PPG) is extracted from this ROI, and position and amplitude of pulses calculated; 4) Artifacts are detected and labeled as such in the PPG; 5) Respiratory induced variations (RIV) are computed; 6) A Smart Fusion process merges the 3 RIV components, compares agreement and excludes artifacts for the calculation of RR.

imaging oximeter on a mobile phone (Camera Oximeter) [1]. After a health check, the subjects began by wearing two Phone Oximeters [18] on the non-dominant hand as reference oximeter measurements. Recording started at sea level with an inspired oxygen concentration (FiO_2) of 21%. The subjects then entered a normobaric hypoxia chamber with an 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%. At regular intervals during the experiment at given FiO_2 levels, 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 [16], OxiCam recorded a video file for 60 s. The video format was set to 240 x 320 pixels resolution (QVGA) and the frame rate (sampling rate) was 20 Hz. The white balance was set to incandescent, as this has been shown to be the optimal configuration for this type of camera [17]. 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. 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 (e.g. calibration process) in the capnogram were excluded. Also, recordings with RR lower than 6 breaths/min (containing episodes of apnea) and with RR higher than 40 breaths/min were considered outliers and excluded. All recording devices (Phone Oximeters, 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. Algorithm

Six steps were necessary to extract RR from the OxiCam video recordings (Figure 1).

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

2) ROI Detection: Optimal ROI for the red video channel was determined using the algorithm described in [17]. 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 poor and no further processing was conducted.

3) PPG Extraction: The PPG was extracted from the ROI using the incremental merge segmenting (IMS) algorithm [19]. The PPG signal was band-pass filtered using a 5th order Butterworth filter with cut-off frequencies at 0.08 Hz and 3 Hz. The IMS algorithm automatically detected pulse peaks and amplitudes.

4) Artifact Detection: Artifacts in the PPG signal were automatically identified within the IMS algorithm. Pulses with amplitudes exceeding lower and upper adaptive thresholds were labeled as artifacts [19]. In addition, artifacts were identified by scanning for abnormal pulse intervals outside of the normal range, defined from 230 to 2400 ms.

5) RIV Extraction: The 3 RIV components RIFV, RIIV, and RIAV were extracted from the PPG. The spectral power of each component was calculated and the frequency with the maximal power within the expected RR range was extracted. The RR range was adaptively determined using the heart rate. 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.

6) Smart Fusion of RR : RR estimation was determined by fusing the respiratory frequencies obtained from RIFV, RIIV, and RIAV by calculating their mean [15]. The quality of this fusion was evaluated by comparing the variance of 3 independent estimations. Estimations with variances higher than 16 breaths^2/min^2 were considered unreliable and discarded. In addition, estimations from windows containing 3 or more artifacts were also discarded. The median RR of the remaining estimations within a recording were calculated and used as the final RR estimation. If less than 4 non-discarded estimations were found throughout the recording, it was considered invalid and no RR was reported.

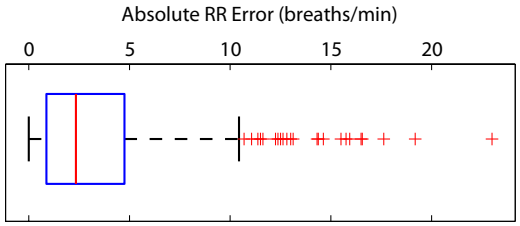

Fig. 2. Boxplot of the absolute error. The vertical solid red line is the median error, the box represents interquartile range and the crosses are outliers.

C. Analysis

Data obtained from the first subject (training subject) was used to adjust algorithm parameters such as IMS segment length and parameters for artifact threshold adaptation. The data from the remaining 18 subjects were used to test the algorithm and compare the performance. Bland-Altman plots [20] were created to compare the Camera Oximeter RR against the reference RR. 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^{co} are the reference and the camera oximeter observations respectively. The absolute error was displayed in a boxplot.

III. Results

A total of 350 Camera Oximeter recordings were obtained (18 from the training subject). Thirty-four were excluded because the reference RR was not available or outside the specified range. For testing the algorithm, 298 recordings remained. Of these, the ROI selection process discarded 22 cases that contained recordings of too poor quality to extract a reliable PPG. Further 46 cases were discarded during the Smart Fusion process as no RR could be reported (too many artifacts in the recording or no agreement between the 3 RIV). The remaining 232 cases were available for comparison with the reference RR. The RMSE was 5.9 breaths/min. The median absolute error was 2.34, range [0 23] (Figure 2). The bias was 3.2 breaths/min and the 95% confidence interval was [-7.3 13.2] (Figure 3).

IV. Discussion

There is a large interest in using mobile phones for health monitoring. In this study, we have investigated if contact photoplethysmographic imaging is suitable for calculating RR in adults. The algorithms used have been previously tested on PPG signals obtained from standard pulse oximetry and shown to be efficient for detecting pulses and artifacts [19], as well as for estimating RR [15]. In this study we applied these algorithms to the PPG obtained from a phone camera and observed a RMSE of 5.9 breaths/min and a median absolute error of 2.3 breaths/min. However, we observed greater errors at RRs above 20 breaths/min,

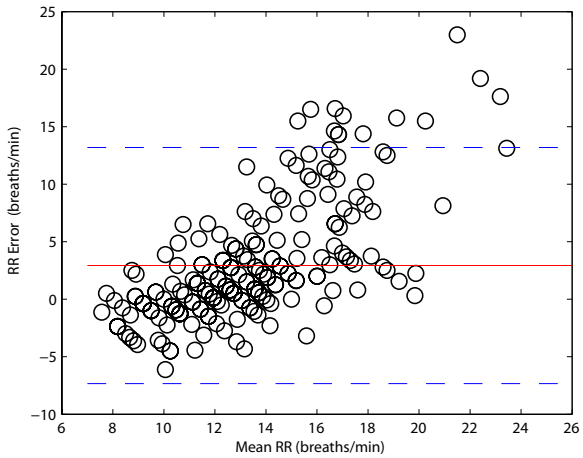

Fig. 3. Bland-Altman plot of the estimated RR from the Camera Oximeter PPG against the reference RR from counting breaths. The solid red line corresponds to the mean error (bias) and the dashed blue lines are ± 2 standard deviations (95% confidence interval).

with one observation having an error of 23 breaths/min at a reference RR of 33 breaths/min.

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 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. Analysis of our cases with high RR and large error revealed that many of these cases contain high power in lower frequencies (Figure 4). Since the algorithm prioritizes the frequency band with highest energy for selecting RR, the large RR error cannot be avoided. 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 largely, but the reference RR would be limited to a single value. The Camera Oximeter RR on the other hand, would calculate RR for 16 s windows and not all windows would necessarily contribute to the final calculation of RR due to the exclusion of measurements during the Smart Fusion (Figure 5). This limitation could be overcome by either comparing instantaneous RR or estimating RR with larger FFT windows (e.g. 60 s or 120 s as done in other studies [13]). However, increasing the window size would reduce the number of available recordings free from artifacts and would result in longer acquisition times. We conclude that high frequency and spontaneous breathing might be challenging to detect using a low-end camera on a mobile phone and recommend validating other algorithms with supplemental data containing breathing at higher RR. This limits the applicability to pneumonia screening where abnormally high RR needs to be reliably detected [6].

Our algorithm successfully eliminated unreliable record-

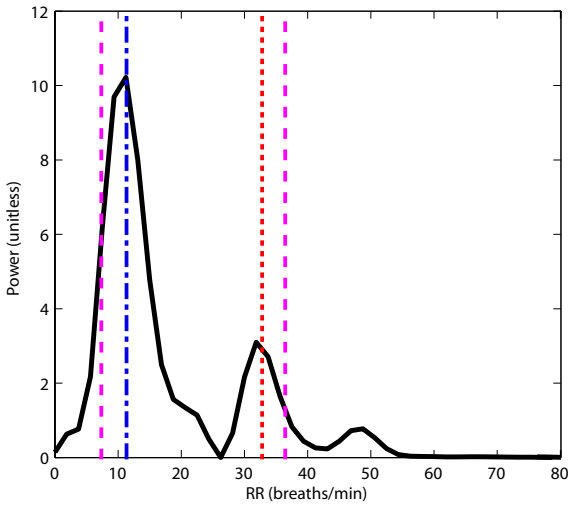

Fig. 4. Example of high power in a low respiratory rate (RR) band where reference RR is high (dotted line). The dashed lines represent the RR selection range, the dot-dash line is the selected RR by the algorithm.

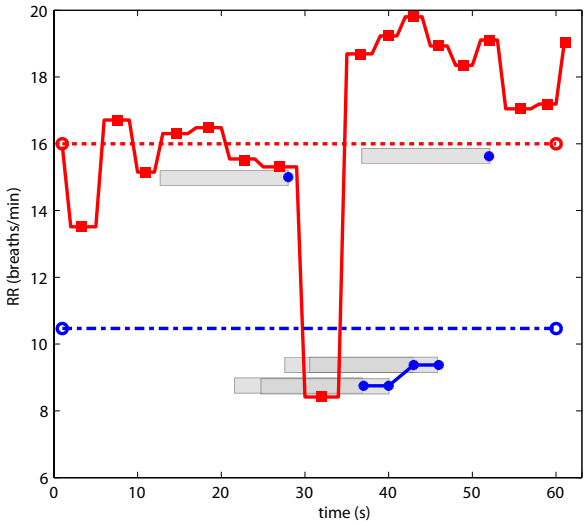

Fig. 5. Example of high variation in respiratory rate (RR) within 60 s. The estimated RR (dot-dash blue) is lower than the reference RR (dotted red) as valid estimation was primarily found for a window where RR was low. The window size was 16 s (grey boxes). Instantaneous RRs are shown as filled squares (reference) and circles (Camera Oximeter).

ings at 2 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 3 RIV. The elimination rate of 22.8% could be further reduced by implementing the complete algorithm into the mobile phone application and providing real-time feedback on acquisition status. If no RR after 60 s is obtained, the recording could be extended until the quality improves.

Acknowledgment

The authors thank the Global Engineering Team 2012 for designing the OxiCam application. Drs. Michael Koehle and Jim Rupert at UBC for kindly allowing the use of the hypoxia chamber for this study. Aryannah Umedaly and Drs. Heng Gan and Chris Petersen for assisting with the data collection.

References

[1] W. Karlen, J. Lim, J. M. Ansermino, G. A. Dumont, and C. Scheffer, "Design Challenges for Camera Oximetry on a Mobile Phone," in *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 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, 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, 2012.

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

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

[8] 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, 2004.

[9] K. Nakajima, T. Tamura, T. Ohta, H. Miike, and P. Oberg, "Photoplethysmographic measurement of heart and respiratory rates using digital filters," in *Proc. 15th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 1993, pp. 1006–7.

[10] 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.

[11] 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.

[12] 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.

[13] 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, 2014.

[14] 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. Conf.*, vol. 40. Zaragoza: IEEE, Sep. 2013, pp. 799–802.

[15] 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, 2013.

[16] 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.*, vol. 2013, Osaka, 2013, pp. 7480–3.

[17] 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.*, vol. 2013, Osaka, Jul. 2013, pp. 2263–6.

[18] 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 *Proc. Int. Conf. Heal. Informatics*, V. Traver, A. Fred, J. Filipe, and H. Gamboa, Eds. Rome, Italy: SciTePress, 2011, pp. 433–8.

[19] 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, 2012, pp. 3131–4.

[20] 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, 1986.