Smartphone-Camera-Acquired Pulse Photoplethysmographic Signal for Deriving Respiratory Rate

Jesús Lázaro*, Yunyoung Nam, Eduardo Gil, Pablo Laguna, and Ki H. Chon

***Abstract*—A method for deriving respiratory rate from smartphone-camera-acquired pulse photoplethysmographic (SCPPG) signal is presented. It combines information from three derived respiration signals based on pulse width, amplitude, and rate variability (PWV, PAV, and PRV). Evaluation is performed over a database containing SCPPG signals recorded from 10 healthy subjects during controlled respiration experiments at rates from 0.2 to 0.6 Hz with a step of 0.1 Hz, using iPhone 4S device. Results suggest that habitual spontaneous respiratory rates (0.2–0.4 Hz) can be estimated from SCPPG signals by PWV and by PRV with low relative error (median of order 0.5% and IQR of order 2%). PWV method maintained its performance at rates up to 0.5 Hz, and the accuracy can be improved by combining it with other methods such as PRV and PAV.**

I. Introduction

Respiratory information is usually monitored by using cumbersome devices which are unmanageable in certain situations such as stress test or sleep studies, and which may interfere with natural breathing. Thus, several algorithms for deriving respiratory information from non-invasive devices have been presented. Some of them are based on pulse photoplethysmographic (PPG) signal [1], [2].

Smartphone devices can record PPG signals with a camera and a flash light [3]. Smartphones are very interesting devices in ambulatory scenarios, since the spectacular improvements of their computational power and their wireless communications make the transference of information really simple. Obtaining respiratory rate from smartphone devices may open several applications, such as anxiety, fatigue, or stress level monitoring. However, Smartphone-camera-PPG (SCPPG) signals are in general noisier than traditional PPG signals and their sampling rate is lower.

In this paper, the PPG-based methods for deriving respiratory rate presented in [2] are adapted to SCPPG signals. They are based on respiration-related modulations in pulse width, amplitude, and rate variability (PWV, PAV, and PRV).

This work was supported in part by the US Army Medical Research and Materiel Command (USAMRMC) under Grant No. W81XWH-12-1-0541, by Universidad de Zaragoza under fellowship PIFUZ-2011-TEC-A-003, by Ministerio de Economía y Competitividad (MINECO), FEDER; under projects TEC2010-21703-C03-02 and FIS-PI12/00514, by CIBER de Bioingeniería, Biomateriales y Nanomedicina through Instituto de Salud Carlos III, and by Grupo Consolidado BSICoS from DGA (Aragón) and European Social Fund (EU).

J. Lázaro, E. Gil and P. Laguna are with the BSICoS Group, I3A, University of Zaragoza, Zaragoza, Spain, and CIBER de Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, Spain (*corresponding author e-mail: jlazarop@unizar.es).

Y. Nam is with the Department of Computer Science and Engineering, Soonchunhyang University, Asan 336-745, Korea.

K. H. Chon is with the Department of Biomedical Engineering, Worcester Polytechnic Institute, MA 01609 USA.

II. Materials and Methods

A. Data and signal preprocessing

The database includes SCPPG recorded by an iPhone 4S device, from 10 healthy subjects during controlled respiration experiments. Subjects were instructed to breathe at a constant rate according to a timed beeping sound. The data were collected for respiratory rates ranging from 0.2 to 0.6 Hz at an increment of 0.1 Hz.

The SCPPG signals were extracted from a 50x50 pixel average of a region on the green video signal at each frame as in [3]. Subsequently, signals were interpolated by cubic splines at a sampling rate of $F_s = 100$ Hz. Then, registers where divided into 60s-length data segments that are shifted every 10 seconds. The baseline contamination was removed with a high-pass filter with a cutoff frequency of 0.3 Hz, and high frequency noise was considerably attenuated by a low-pass filter with a cutoff frequency of 35 Hz. Segments with 30% or more of the time with artifacts according to artifact detector [4] were discarded.

B. Significant point detection

SCPPG pulses apex points n_{A_i} were detected by an automatic PPG pulse detector based on a low-pass-differentiator filter and a time-varying threshold [5]. Then, baseline point of the i^{th} SCPPG pulse n_{B_i} was defined as the minimum of SCPPG signal within interval $[n_{\mathrm{A}_i} - 0.4F_s, n_{\mathrm{A}_i}]$. The middle point n_{M_i} is defined as the point where SCPPG has reached half of the maximum pulse amplitude. Further details are given in [5]. In order to measure the width of the pulses, pulse wave onset n_{O_i} and end n_{E_i} points were detected by using the algorithm described in [2], which is based on the derivative of the preprocessed signal.

C. Derived respiration signals

Three derived respiration (DR) signals were extracted from SCPPG signal, based on PRV, PAV, and PWV:

$$d^u_{\mathrm{PRV}}(n) = \sum_i F_s \frac{1}{n_{\mathrm{M}_i} - n_{\mathrm{M}_{i-1}}} \delta\left(n - n_{\mathrm{M}_i}\right) \quad (1)$$

$$d^u_{\mathrm{PAV}}(n) = \sum_i \left[x(n_{\mathrm{A}_i}) - x(n_{\mathrm{B}_i})\right] \delta\left(n - n_{\mathrm{A}_i}\right) \quad (2)$$

$$d^u_{\mathrm{PWV}}(n) = \sum_i \frac{1}{F_s} \left[n_{\mathrm{E}_i} - n_{\mathrm{O}_i}\right] \delta\left(n - n_{\mathrm{A}_i}\right) \quad (3)$$

where $x(n)$ denotes the preprocessed SCPPG signal, and superscript "u" denotes that the signal is unevenly sampled.

A 4-Hz evenly sampled version of each DR signal by cubic spline interpolation was obtained. The resulting signals are denoted without the superscript "u".

978-1-4799-3969-5/14/$31.00

D. Respiratory rate estimation

1) From one DR signal: The power spectrum density (PSD) of the j^{th} DR signal $S_j(f)$ was estimated by applying a modified periodogram using a Hamming window. Then, the respiratory rate $\hat{f}$ is estimated as the frequency at where the absolute maximum of the PSD is located, within the studied band $[0.15, 0.7]$ Hz.

2) From combination of the three DR signals (P_{COMB}): Information from the three DR signals was combined to increase the robustness of the respiratory rate estimation, by an adaptation of the algorithm presented in [2]. The PSDs are "peak-conditioned averaged"; only those $S_j(f)$ which are peaked enough take part in the averaging. In this paper, "peaked" denotes that a certain percentage (ξ) of PSD must be contained in an interval around its highest peak. In mathematical terms, "peakness" of a PSD is defined as follows:

$$P_j = \frac{\int_{f_p(j)-0.05\text{Hz}}^{f_p(j)+0.05\text{Hz}} S_j(f)df}{\int_{0.15\text{Hz}}^{0.7\text{Hz}} S_j(f)df} \tag{4}$$

where $f_p(j)$ denotes the highest peak within the studied band $[0.15, 0.7]$ Hz in the PSD of the j^{th} DR signal.

In order to select those spectra that are sufficiently "peaked", two different criteria where established: χ^{A} and χ^{B}. On one hand, χ^{A} lets those spectra whose "peakness" is greater than a fixed value take part in the average as shown in (5). On the other hand, χ^{B} compares the spectra of different DR signals, letting those spectra more peaked take part in the average, although all of them have passed the χ^{A} criterion as shown in (6).

$$\chi_j^{\text{A}} = \begin{cases} 1, & P_j \geq \xi \\ 0, & otherwise \end{cases} \tag{5}$$

$$\chi_j^{\text{B}} = \begin{cases} 1, & P_j \geq \max\limits_j \{P_j\} - \lambda \\ 0, & otherwise \end{cases} \tag{6}$$

The "peak-conditioned" average is computed as follows:

$$\bar{S}(f) = \sum_j \chi_j^{\text{A}} \chi_j^{\text{B}} S_j(f). \tag{7}$$

Finally, $\hat{f}$ is estimated as follows:

$$\hat{f} = \underset{f\in[0.15,0.7]}{\arg\max} \left\{\bar{S}(f)\right\}. \tag{8}$$

III. Results

Relative error e_{R} with respect to the real respiratory rate (f_{R}) was obtained for each studied DR signal and combination as defined in (9). Medians and interquartile ranges (IQR) obtained for e_{R} from different DR signals and combination for each f_{R} are shown in Table I.

$$e_{\text{R}} = \frac{\hat{f} - f_{\text{R}}}{f_{\text{R}}} \times 100. \tag{9}$$

Note that aliasing problems may affect methods, since respiratory information is obtained only at pulse occurrence. For this reason, fragments associated to a respiratory rate higher than half mean pulse rate were excluded from the study.

TABLE I
OBTAINED MEDIANS (MED) AND IQRS FOR e_{R} FROM DIFFERENT DR SIGNALS AND COMBINATION FOR EACH f_{R}.

Method		f_{R} (Hz)				
		0.2	**0.3**	**0.4**	**0.5**	**0.6**
$d_{\text{PRV}}(n)$	Med (%)	0.10	-0.39	-0.15	-0.39	-18.29
	IQR (%)	1.95	1.63	2.14	10.30	47.36
$d_{\text{PAV}}(n)$	Med (%)	0.10	-0.07	-0.63	-28.13	-60.37
	IQR (%)	3.91	3.58	31.80	56.64	69.82
$d_{\text{PWV}}(n)$	Med (%)	0.10	0.26	-0.15	-0.20	-0.31
	IQR (%)	1.95	1.30	1.95	0.78	56.23
P_{COMB}	Med (%)	0.10	-0.07	-0.15	-0.20	-0.39
	IQR (%)	1.83	0.98	0.49	0.39	66.24

IV. Discussion and conclusions

Deriving information from SCPPG signals represent a challenging situation, since their low sampling rate and the ambient-light noise affect considerably their quality.

Those fragments associated with an f_{R} higher than half mean pulse rate were excluded because methods would track an alias in such situations. This remains a limitation because in a real situation it cannot be determined whether f_{R} is above half mean pulse rate or not. However, a high f_{R} with a low pulse rate does not represent a realistic physiological situation. In such situations when the autonomic nervous system requires a high respiratory rate, it also requires a high heart rate, e.g., during exercise.

In general, methods obtained better results for normal ranges of spontaneous f_{R} (0.2–0.4 Hz) than for higher rates (0.5 and 0.6 Hz) in e_{R} terms. A possible reason for this observation is that respiration-induced modulations on which DR signals are based (PRV, PAV and PWV) may have a less strong effect at high respiratory rates. In case of pulse rate, it is known that respiratory sinus arrhythmia (which modulates the heart rate) is reduced at high respiratory rates. However, PWV method and combination P_{COMB} maintained their performance up to $f_{\text{R}} = 0.5$ Hz. Although methods were not tested in a spontaneous breathing situation, PRV, PAV and PWV respiratory modulations were previously studied with PPG signals during spontaneous breathing experiments [2].

Results suggest that normal ranges of spontaneous f_{R} (0.2–0.4 Hz) can be estimated from SCPPG signals by the method based on PWV and the one based on PRV, with low e_{R} (median of order 0.5% and IQR of order 2%). PWV method maintained its performance at rates up to 0.5 Hz, and the accuracy can be improved by combining it with other methods such as PRV and PAV.

References

[1] K. H. Chon, S. Dash, and K. Ju, "Estimation of respiratory rate from photoplethysmogram data using timefrequency spectral estimation," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 8, pp. 2054–2063, 2009.

[2] J. Lázaro, E. Gil, R. Bailón, A. Mincholé, and P. Laguna, "Deriving respiration from photoplethysmographic pulse width," *Med. Biol. Eng. Comput.*, vol. 51, no. 1-2, pp. 233–242, 2013.

[3] 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. 303–306, 2012.

[4] E. Gil, J. M. Vergara, and P. Laguna, "Detection of decreases in the amplitude fluctuation of pulse photoplethysmography signal as indication of obstructive sleep apnea syndrome in children," *Biomed. Signal Process. Control*, vol. 3, no. 3, pp. 267–277, 2008.

[5] J. Lázaro, E. Gil, J. M. Vergara, and P. Laguna, "Pulse rate variability analysis for discrimination of sleep-apnea-related decreases in the amplitude fluctuations of pulse photoplethysmographic signal in children," *IEEE J. Biomed. Health Inform.*, vol. 18, no. 1, pp. 240–246, 2014.