

Monitoring breathing rate by fusing the physiological impact of respiration on video-photoplethysmogram with head movements

Luca Iozza¹, Jesús Lázaro^{2,3,4}, Luca Cerina¹, Davide Silvestri⁵,
Luca Mainardi¹, Pablo Laguna^{2,3} and Eduardo Gil^{2,3}

¹Dipartimento di Elettronica, Informatica e Bioingegneria, Politecnico of Milan, Milan, Italy

²BSICoS Group, Aragon Institute of Engineering Research (I3A), IIS Aragon, University of Zaragoza, Zaragoza, Spain

³Centro de Investigacion Biomedica en Red - Bioingenieria, Biomateriales y Nanomedicina (CIBER-BBN), Zaragoza, Spain

⁴Department of Biomedical Engineering, University of Connecticut, Storrs, USA

⁵University of Bologna, Italy

E-mail: edugilh@unizar.es

Abstract.

Objective: The simple observation of breathing rate (BR) remains the first and often the most sensitive marker of acute respiratory dysfunction. In fact, there are evidences that drastic changes in BR are a predictive indicator of adverse events (i.e. cardiac arrest). The aim of this study is to develop a camera-based technology that may provide near-continuously estimation of BR considering the effect of respiration on videoPPG (vPPG).

Approach: The technology has been tested in two different experimental settings, including controlled BR and more challenging scenarios with spontaneous breathing pattern. Video data were processed offline to derive the vPPG signal. The method derives respiration from beat-to-beat PPG rate and morphology changes in amplitude and width driven by the physiological relationships between vPPG and respiration. Moreover, respiratory-induced head movements were used as additional source of information for the vPPG system. A combination of these methods has been exploited to estimate the respiratory rate every 10 second.

Main Results: According to the results, respiratory frequencies in the central range (0.2-0.4 Hz) may be estimated using vPPG system with low relative error, $\epsilon_R < 2\%$ and interquartile range of order $IQR < 5\%$. However, the vPPG system showed a drop in performance at respiratory range boundaries, around 0.1 Hz and 0.5 Hz.

Significance: This camera-based technology can be used as ubiquitous BR monitoring system. However, vPPG-based systems should consider the effect of the BR in the estimation, mainly in applications where respiratory rate is out of the range 0.2-0.4 Hz.

Keywords: v-PPG, contactless system, camera, breathing rate.

1. Introduction

Monitoring breathing rate (BR) is a strong and specific predictor of serious adverse events such as cardiac arrest and unplanned intensive care unit admission (Cretikos et al., 2008; Fieselmann et al., 1993). Fieselmann showed that a BR greater than 27 breaths per minute is a predictor of cardiac arrest, even more informative compared to cardiac pulse and blood pressure changes. Moreover Gravelyn and Weg (1980) demonstrated that the BR is a specific marker of acute respiratory dysfunction such as hypoxemia (abnormal low level of O₂ in the blood) and hypercapnia (abnormal CO₂ in the blood).

Different methods have been exploited to monitor respiratory activity: airflow sensing technology uses nasal sensor installed on a mouthpiece of facemask to measure the volume of air exhaled; microphones recording respiratory sounds (Corbishley and Rodriguez-Villegas, 2007); airflow can be detected because expiratory air is warmer, has higher humidity and contains more CO₂ than inspiratory air (Folke et al., 2003); impedance plethysmography measures a variation of transthoracic electrical impedance to derive ventilation (Cohen et al., 1997); respiratory inductance plethysmography includes an elastic respiratory band made of strain gauge that, by the respiratory-induced motion, changes the resistance of the conductor according to the respiration or based on the modulation of the cardiac activity (Charlton et al., 2017). Each method has its own disadvantages, i.e. cumbersome for the patient, time-consuming and does not allow long-term monitoring (AL-Khalidi et al., 2011; Massaroni et al., 2019). Nowadays there is still a lack of low cost home-care solution that can be ubiquitous and allow daily monitoring of respiratory activity.

In the state of art, a step forward in the improvement of patient compliance has been obtained by the introduction of non-contact devices. There are clear advantages to non-contact respiration monitoring methods. These include improved patient comfort (especially for long term monitoring) as the subject is not tied to an instrument and improved accuracy as distress caused by a contact device may alter the respiration rate (AL-Khalidi et al., 2011). Three main camera-based approaches have been proposed to monitor respiratory information based on *movement*, *thermal* and *photoplethysmography* principles. Movement based methods exploit image processing techniques for the detection of shifting associated to respiration, basically chest and/or abdomen movement (Rehoumaa et al., 2018; Regev and Wulich, 2017; Janssen et al., 2016; Al-Naji and Chahl, 2016; Nam et al., 2016; Bartula et al., 2013; Zhao et al., 2013; Nakajima et al., 2001) or even shoulder movement (Shao et al., 2014). In absence of other movement artifacts, these systems estimate the respiratory rate, and many of them have been validated while subjects lying in bed, including neonates (Rehoumaa et al., 2018; Janssen et al., 2016), babies (Al-Naji and Chahl, 2016; Regev and Wulich, 2017) and adults

(Bartula et al., 2013; Nakajima et al., 2001). Thermal cameras allow respiratory rate estimation by monitoring dynamic thermal activity emitted from specific areas (usually nostrils), as the temperature of exhaled air is higher than the inhaled air (Hochhausen et al., 2018; Pereira et al., 2017; Lewis et al., 2011; Abbas et al., 2011). In addition, some authors have combined both methods in order to increase the accuracy (Bennett et al., 2017; Procházka et al., 2017; Pereira et al., 2016).

A contactless alternative is represented by video-photoplethysmographic system (vPPG). In literature, the vPPG-based method has been widely used to monitor cardiac parameters such as heart rate and pulse rate variability (PRV). However it has been poorly studied in the state of art as method to derive respiratory features as BR. In Poh et al. (2011) the BR has been extracted as a component of power spectrum of PRV derived from vPPG signal. In Villarroel et al. (2014) and Tarassenko et al. (2014) a band pass filter tuned to the expected BR has been applied on vPPG signal. Hence the peak in the spectrum has been considered as BR. Similarly, Al-Naji et al. (2017) combined information of head movement and vPPG signal for respiratory rate estimation based on spectral analysis and filtering. The first pioneering work aimed at deriving a breathing signal from vPPG was performed by Gastel et al. (2016). The method firstly proposed to divide the face into different sub-regions. For each of them a set of weights has been calculated using the Chrominance method (already implemented in de Haan and Jeanne (2013)) that allows to enhance the respiratory information and suppress motion distortions. Then the best set of weights has been selected by the SNR quality parameter and has been used to exploit the breathing signal.

As shown in literature, contact PPG signal has been used to derive breathing information based on the physiological effects of respiration over it. It is known that respiration modulates pulse beat occurrence and PPG waveform (Lázaro et al., 2015; Garde et al., 2013). Indeed during inspiration pulse-to-pulse interval shortens while during expiration it is prolonged. The following phenomenon, called respiratory sinus arrhythmia (RSA), plays an important role in respiratory-circulatory interaction improving the efficiency of pulmonary gas exchange (Yasuma and Hayano, 2004). During the mechanism of alveolar ventilation/perfusion matching, RSA allows to save energy by suppressing unnecessary heartbeats and increasing the efficiency of the ventilation during the end of perfusion. Therefore the respiratory component can be extracted from pulse rate variability (PRV).

The mechanics of respiration, in particular inspiration, can lead to a reduction in tissue blood volume (reflected in a reduction of vPPG amplitude) mainly by two distinct mechanisms (Meredith et al., 2012): reduction in cardiac output; reduction of venous pressure. As second backward effect, the negative intrathoracic pressure during inspiration produces a reduction in venous pressure which pools the blood from the vascular bed within the tissue. Therefore the pulse amplitude variability (PAV) can be an indirect measure of breathing rate.

Finally, as showed by Lázaro et al. (2013), the sympathetic nervous system regulates

the vascular resistance by means of veins constriction, venous compliance and blood volume. During expiration the blood vessels are more flexible than during inspiration. This mechanism affects the pulse wave propagation velocity that is reflected on the pulse width of PPG signal. Therefore the pulse width variability (PWV) is affected by the respiratory modulation.

Focusing on non-contact respiration methods based on *photoplethysmography*, there is a need of characterizing the physiological effects that affect the estimation. Most of the studies in the field have been performed during spontaneous breathing at resting conditions without considering different BRs (Verkruysse et al., 2008; Poh et al., 2011; Villarroel et al., 2014; Tarassenko et al., 2014; Al-Naji et al., 2017). Our hypothesis is that these methods could be affected by the respiratory rate, as the physiological modulation of PPG by respiration depends on it. For instance, it's well known that RSA decreases at high respiratory rates (Sin et al., 2010) and in the case of vPPG this decrease comes together with a low-passed shape in comparison to fPPG which increases the uncertainty in locating the pulse, and so reducing RSA estimation. These effects should be carefully considered in BR estimation applications based on vPPG depending on the expected respiratory pattern. Mainly when BR is out of its typical values at rest, i.e. in sport, relax conditions or some clinical applications among others. The aim of the present work is to test the feasibility of contactless, camera-based method that estimates BR values based on vPPG signal considering the effect of respiration on vPPG signal. In addition, the added value of using movement information for the estimation will be evaluated. Compared to the state of art, the novelty lies in several aspects as: providing an exhaustive analysis of the effect of respiratory rate in the estimation; implementation of data fusion technique that combines breathing information derived from photoplethysmography (PPG) waveform changes, pulse rate variability (PRV), pulse amplitude variability (PAV) and pulse width variability (PWV), and by vertical head movements; assessment of PPG-derived respiratory modulation measured at two different body sites (finger vs head) using respectively finger PPG and vPPG technology; use of two different technological cameras in order to test the robustness of implemented BR estimation. A preliminary version of this work has been reported (Iozzia et al., 2017).

2. Materials

2.1. Experimental protocols

Two experiments were conducted respectively in 20 healthy subjects. In both experiments the subject was sitting on a chair, facing the camera device at a distance of approximately 0.5 – 1 m. To reduce motion artifacts, participants were instructed to move their head as less as possible. In both experiments the PPG signal was recorded from index finger, whereas the respiratory signal was recorded by a respiratory belt fixed

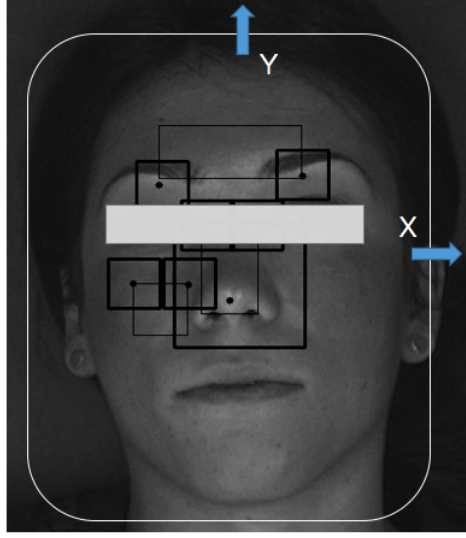


Figure 1: Tracking of the face according to horizontal (X axis) and vertical (Y axis) movement.

on the subject's chest. Both signals were sampled at 256 Hz using the FlexComp InfinitiTM biofeedback system (Thought Technology Ltd., Montreal, Canada). The study was conducted in accordance with the ethical principles of the Declaration of Helsinki.

Experiment I: Video recordings were performed by the Logitech C922 Pro Stream Webcam. The resolution was set to 1280×720 with a frame rate of 60 frames-per-second (fps). The subject was informed to breath at a constant rate, by using a set of videos showing a moving bar, from 0.1 to 0.5 Hz with an increment of 0.1 Hz. For each respiratory frequency a recording of 2 minutes was applied.

Experiment II: In the second protocol a GigE Sony XCG-C30C camera was selected with a spatial resolution of 659×494 pixels and 60 fps. The camera was equipped with 15 mm fixed focal length lens (Tamron 25 HB), which was used to magnify the region occupied by the participant's face. The protocol was the following: 2 minutes of spontaneous breathing, a period of apnea (whose duration depended on subject's capacity) and 2 minutes of recovery.

The data in the second database were divided into two stages: *spontaneous breath* lasted 110 s of recordings and *recovery breath* lasted 110 s after the apnea phase. Each stage was treated as independent, for which the estimation of respiratory rate was executed. The stage of apnea was not considered in this study.

3. Methods

3.1. Signal preprocessing

Videos were saved in H.264 format, with a compression rate factor close to 0 (so no effect related to data compression is expected) and processed offline to derive the vPPG signal according to the algorithm presented in [Iozzia et al. \(2016\)](#). Three regions of

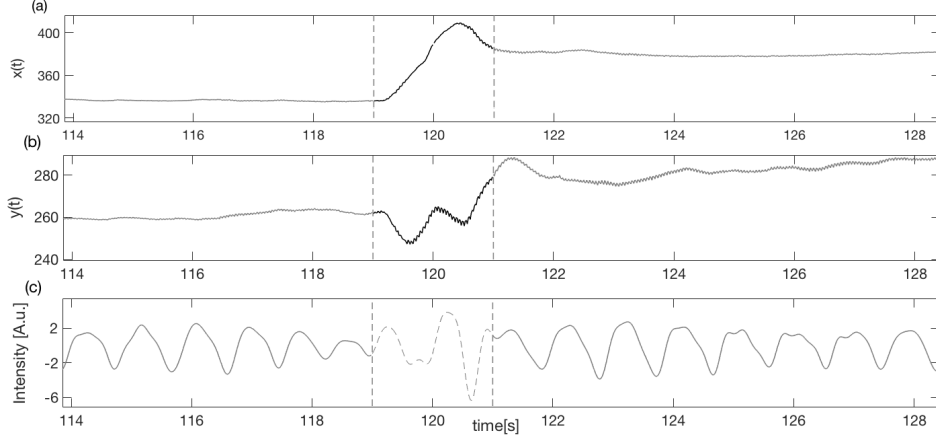


Figure 2: Example of artifact movement removal on vPPG signal (c) using the X-Y tracking signals (a)-(b) as reference of noise.

interests (ROI) were considered: forehead (ROI_1), nose (ROI_2) and cheek (ROI_3). ROI detection and tracking were developed in OpenCV: detections were performed by the Viola-Jones face detection algorithm (Viola and Jones, 2001) while ROI tracking along X-Y axis (respectively horizontal and vertical movements) was obtained by the Lucas-Kanade-Tomasi (LKT) motion flow tracking algorithm (Lucas and Kanade, 1991). A modification of LKT was presented in this work: a Kalman filter was implemented to smooth movements of the face tracked with LKT feature tracker using a linear model for the prediction and as observations the Y axis coordinate of the LKT centroid tracker output.

Within each ROI a spatial average of pixel intensities of each channel (red, green and blue) has been calculated for each frame to generate N raw signals (where N is the number of color channels). Considering M as the number of ROI , a total of $M * N$ raw traces was obtained. The time series were first detrended (Tarvainen et al., 2002) and band-pass filtered (FIR filter using *Hamming window*) with cut-off frequencies $f_{c1} = 0.1$ Hz and $f_{c2} = 5$ Hz. To enhance the pulsatile component of the reflected light and reduce the motion noise, the zero-phase component analysis (ZCA) method (Bell and Sejnowski, 1997) was applied in each ROI . As a result, M vPPG signals were obtained. Among them, the selection of the target vPPG signal was achieved by calculating the power spectral density (PSD) on the entire signal and by measuring the SNR using the following formula:

$$SNR_i = 10 * \log_{10} \left(\frac{\int_{f_1}^{f_2} \mathcal{P}_{vPPG_i}(f) df}{\int_{0.1}^{f_1} \mathcal{P}_{vPPG_i}(f) df + \int_{f_2}^5 \mathcal{P}_{vPPG_i}(f) df} \right) \quad (1)$$

where $\mathcal{P}_{vPPG_i}(f)$ with $i \in \{1, 2, \dots, M\}$ is the PSD of i^{th} vPPG signals, $f_1 = f_p - 0.15$ Hz, $f_2 = f_p + 0.15$ Hz and where f_p is the highest peak assumed to be the pulse frequency (measured in Hz). The signal with the highest SNR was selected as target

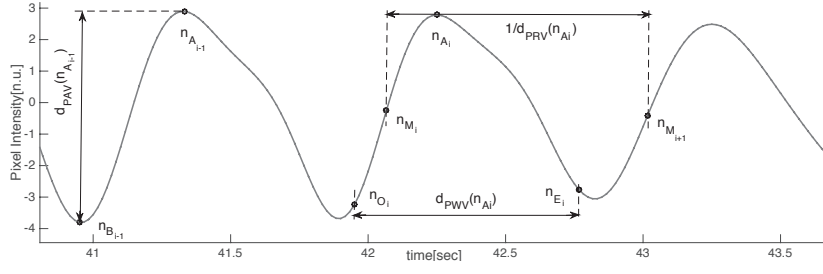


Figure 3: Derived respiratory signals defined by a set of fiducial points.

vPPG signal. Subsequently, the motion artifacts were automatically detected using the $X - Y$ coordinates of facial movement (as shown in Fig.1). Within a sliding window of 1 s length, the first derivative of X and Y coordinates (defined as v_X and v_Y) was calculated. If the mean of v_X or v_Y calculated inside the interval time was higher than a threshold (manually selected at 14 pixels/s), the corresponding vPPG segment was labeled as artifactual and discarded for further analysis (see Fig. 2). Less than 5% of the total number of segments was discarded to the successive procedure.

3.2. Detection of fiducial points

To derive Pulse Width Variability (PWV), Pulse Rate Variability (PRV) and Pulse Amplitude Variability (PAV) signals both from finger-PPG (fPPG) and vPPG the method implemented by [Lázaro et al. \(2013\)](#) was applied. Briefly the systolic peaks n_{A_i} were identified by an automatic PPG pulse detector based on a low-pass filter differentiator and an adaptive time-varying threshold ([Lázaro et al., 2014](#)). Next the diastolic points n_{B_i} were identified inside a temporal window previous to n_{A_i} :

$$n_{B_i} = \arg \min_n \{s(n)\}, n \in [n_{A_i} - 0.3f_s, n_{A_i}], \quad (2)$$

where $s(n)$ is the PPG signal and f_s the sampling frequency of the signal. Finally n_{M_i} was detected as the half of the rising edge of the cardiac pulse:

$$n_{M_i} = \arg \min_n \left\{ \left| s(n) - \frac{s(n_{A_i}) + s(n_{B_i})}{2} \right| \right\}, n \in [n_{B_i}, n_{A_i}], \quad (3)$$

The width of each pulse wave was measured by considering the onset n_{O_i} and the end of the wave n_{E_i} , while the distances between $n_{M_{i+1}}$ and n_{M_i} were used to estimate the pulse rate (see Fig. 3). More details are shown in the original work by [Lázaro et al. \(2013\)](#). Although PWV and PRV seem similar in this figure both DR signals are considered as its physiological origin is completely different. PRV is reflecting the RSA, i.e. the effect of respiration in the generation of beats at the sinoatrial node, whilst PWV is reflecting the effect of pulse wave propagation, which is also affected by the respiratory modulation ([Lázaro et al., 2013](#)).

3.3. Derived respiratory signals

3.3.1. From PPG signals Three derived respiratory (DR) signals were calculated using pulse-to-pulse methods PRV, PAV and PWV. The DR signal based on PRV was obtained as pulse-to-pulse series:

$$d_{PRV}^u(n) = \sum_i f_s \frac{1}{n_{N_i} - n_{N_{i-1}}} \delta(n - n_{A_i}) \quad (4)$$

where 'u' indicates that the signal is unevenly sampled, and n_{N_i} represents the arrival of pulse after removing the ectopic and missed pulses from n_{M_i} by using the method proposed by [Mateo and Laguna \(2003\)](#). The other two derived signals based on PAV and PWV were obtained as follows:

$$d_{PAV}^u(n) = \sum_i [s(n_{A_i}) - s(n_{B_i})] \delta(n - n_{A_i}) \quad (5)$$

$$d_{PWV}^u(n) = \sum_i \frac{1}{f_s} [n_{E_i} - n_{O_i}] \delta(n - n_{A_i}) \quad (6)$$

An outlier rejection rule based on median absolute deviation was applied ([Bailón et al., 2006](#)) and the signals were resampled at 4 Hz using the cubic spline interpolation. Non-respiratory frequencies should be removed from respiratory signals to avoid erroneously identifying spurious frequency content as the BR. There is no consensus on the optimal range of plausible respiratory frequencies. Furthermore, it may need to be adjusted according to the patient population or the specific application [Charlton et al. \(2017\)](#). A band-pass filter was applied with the low and high cutoff frequencies 0.075 Hz and 1 Hz, respectively.

3.3.2. From tracking signals A fourth DR signal was obtained by Y coordinate of the face movement returned from LKT algorithm. Since respiration affects the vertical movement of the face, the $y(t)$ signal derived from the Y-axis movement along the time was used as a source to detect breath. The signal was first resampled at 4 Hz and then a high-pass filter with a cutoff frequency of 0.075 Hz was applied to remove the signal trend. However the subject may show various involuntarily head movements not correlated to the breath such as *yaw movement* (turning the head), *rolling movement* (tilting the head) and *pitching movement* (looking up and down).

To detect possible abrupt changes in the signal due to these movements, the speed of the two motion signals $v_X(n)$ and $v_Y(n)$ was used. Defining a sliding temporal window of 2 s, if both the speed along $v_X(n)$ and $v_Y(n)$ signals was above the threshold $T_1 = 14$ pixels/s (empirically chosen) the corresponding segment of the signal was excluded. Since the *pitching* movement will be the most disturbing source of noise, a second threshold $T_2 = 30$ pixels/s was set to reject episodes with high vertical head movements by controlling the speed $v_Y(n)$. Finally, the motion signal $d_{TR}(n)$ was obtained as band pass filtered version of $y(n)$ with the cutoff frequencies of $f_{c1} = 0.075$ Hz and $f_{c2} = 1$ Hz.

3.4. Respiratory frequency estimation

The respiratory rate has been estimated from DR signals on a time interval of 30 s shifted every 10 s by adapting Lázaro's method (Lázaro et al., 2015).

Briefly the method uses a combination of PWV, PAV and PRV (denoted as combined vPPG derived respiration, CvPDR) and, in addition, the effect of including the tracking signal to increase the accuracy of the estimation is considered (CvPDR+TR). Let's indicate $S_j(k, f)$ the power spectral densities for the j -th DR signal at every time instant k . From these spectra an average is computed as:

$$\bar{S}(k, f) = \sum_j \chi_j^A(k) \chi_j^B(k) S_j(k, f) \quad (7)$$

where χ_j^A and χ_j^B are two weighting factors used to restrict the inclusion in the average $\bar{S}(f)$ to: a) peaked spectra; b) spectra whose peakness is not much lower than the peakness of the other spectra at that time instant. For this reason the following quantity $P_j(k)$ has been calculated:

$$P_j(k) = \frac{\int_{f_p(k,j)-0.05Hz}^{f_p(k,j)+0.05Hz} S_j(k, f) df}{\int_{0.15Hz}^{0.5Hz} S_j(k, f) df} \times 100 \quad (8)$$

where $P_j(k)$ is the percentage of $S_j(k, f)$, related to the j -th DR signal, contained in the interval centered around the highest peak $f_p(k, j)$. Peaked spectra are defined as those which have $P_j(k)$ greater than a threshold defined by $\xi = 40\%$, thus:

$$\chi_j^A(k) = \begin{cases} 1, & P_j(k) \geq \xi \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

while the condition $\chi_j^B(k)$ selects the $S_j(k, f)$ whose $P_j(k)$ is not less than $\lambda = 30\%$ of the maximum $P_j(k)$:

$$\chi_j^B(k) = \begin{cases} 1, & P_j(k) \geq \max_j \{P_j(k)\} - \lambda \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

The respiratory frequency $\hat{f}(0)$ is calculated as the frequency that corresponds to the maximum of $\bar{S}(0, f)$ in the frequency band $[0.15 - 0.5]$ Hz:

$$\hat{f}(0) = \arg \max_{f \in [0.15-0.5]Hz} \{\bar{S}(0, f)\} \quad (11)$$

If in the spectrum of each DR signal no peak has enough energy in the considered frequency band, the range is enlarged to $[0.08 - 0.5]$ Hz and a new search is started.

After initialization of $\hat{f}(0)$, the frequency band for the searching of $\hat{f}(k)$ was defined as $[\hat{f}(k-1) - 0.075; \hat{f}(k-1) + 0.15]$ Hz where $\hat{f}(k-1)$ is the previous estimated value.

As it can be imagined, the phase of initialization of the algorithm is decisive, since next respiratory rate estimations will be based on a window defined by the previous value $\hat{f}(k-1)$. Due to the morphology of vPPG signal, for higher respiratory frequencies (v.g. $f \geq 0.3$ Hz), the modulation of the breath is less evident compared to the modulation presented on fPPG. To account for this aspect, a modification of Lázaro's algorithm was proposed when tracking signal is considered. Since the tracking derived respiratory signal is uncorrelated to the breath modulation on vPPG signal, it could be possible to increase the robustness of the respiratory rate estimation even for high respiratory rates. Therefore during the initialization, the procedure described to estimate the respiratory rate was repeated separately to $d_{PRV}(n)$, $d_{PAV}(n)$, $d_{PWV}(n)$ and to $d_{TR}(n)$ signal. Two \hat{f} were estimated: \hat{f}_{vPPG} and \hat{f}_{TR} and the initial respiratory frequency was established as:

$$\hat{f}(0) = \begin{cases} \hat{f}_{TR}, & \text{if } \Delta \geq 0.3 \\ \hat{f}_{vPPG}, & \text{otherwise} \end{cases} \quad (12)$$

where $\Delta = \hat{f}_{TR} - \hat{f}_{vPPG}$. If the tracking signal was detecting a higher respiratory frequency compared to the one measured by pulse-to-pulse methods, $\hat{f}(0)$ was forced to start at \hat{f}_{TR} .

Finally, the results are expressed as relative error calculated according to the following equation:

$$\epsilon_R = \frac{f_R - \hat{f}}{f_R} \times 100 \quad (13)$$

where f_R is the respiratory frequency obtained from the respiratory belt used as a reference for the analysis of fPPG and vPPG.

4. Results

4.1. Cameras comparison

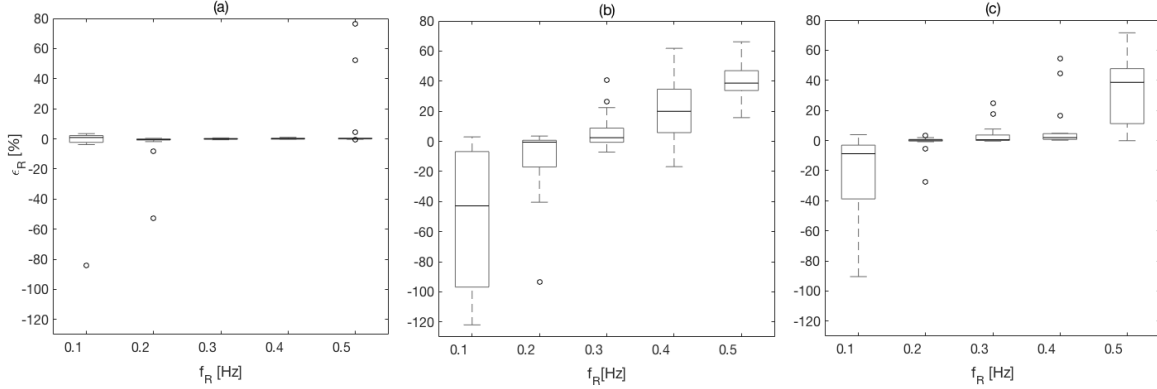
To evaluate the performance of the two cameras on the extraction of vPPG signal, the SNR of the vPPG signal of each subject has been calculated according to the Eq. 1 and averaged for the whole population (as *mean ± standard deviation*). The vPPG quality signal obtained from the commercial camera is significantly lower (9.06 ± 3.84 dB) compared to the one obtained from the industrial one (12.20 ± 4.73 dB).

4.2. ROI selection

As stated before, the *ROI* used for breath extraction was selected according to the highest *SNR*. In Table 1 it is worth to notice that the *forehead* was the most selected region independently of proposed experimental setting.

Table 1: Percentage of use of each *ROI* according to selected Database.

ROI	Database I [%]	Database II [%]
<i>forehead</i>	60	61.11
<i>nose</i>	30	11.11
<i>cheek</i>	10	27.78

**Figure 4:** Boxplot of the respiratory rate relative error, calculated for the dataset I, obtained with fPPG (a), CvPDR (b) and CvPDR+TR (c) respectively**Table 2:** Relative error, ϵ_R expressed as median and interquartile ranges (IQR) in Database I.

f_R [Hz]	fPPG			CvPDR			CvPDR+TR		
	ϵ_R [%]	IQR[%]	$n_{OUT_{10\%}}$	ϵ_R [%]	IQR[%]	$n_{OUT_{10\%}}$	ϵ_R [%]	IQR[%]	$n_{OUT_{10\%}}$
0.1	0.89	4.56	1	-37.78	89.30	11	-8.80	35.69	6
0.2	-0.59	0.75	1	-0.56	17.70	5	0.26	1.11	1
0.3	-0.14	0.51	0	2.43	9.41	4	0.49	3.62	2
0.4	0.12	0.63	0	19.93	28.82	13	1.94	3.88	3
0.5	0.28	0.47	2	38.73	13.23	17	38.65	36.55	13

4.3. Database I

Results are shown in Table 2, Fig. 4 and Fig. 5 for fPPG (a), CvPDR (b) and CvPDR+TR (c). The comparison of the methods shows different behavior for different ranges:

- (i) The best results are obtained if the BR is in the range $[0.2-0.4]$ Hz. The median and the interquartile ranges (IQR) of ϵ_R are below the 5% for fPPG and CvPDR+TR, while they are higher for CvPDR. In particular for the fPPG there is only one estimation with a relative error higher than 10% (defined as $n_{OUT_{10\%}}$) among all

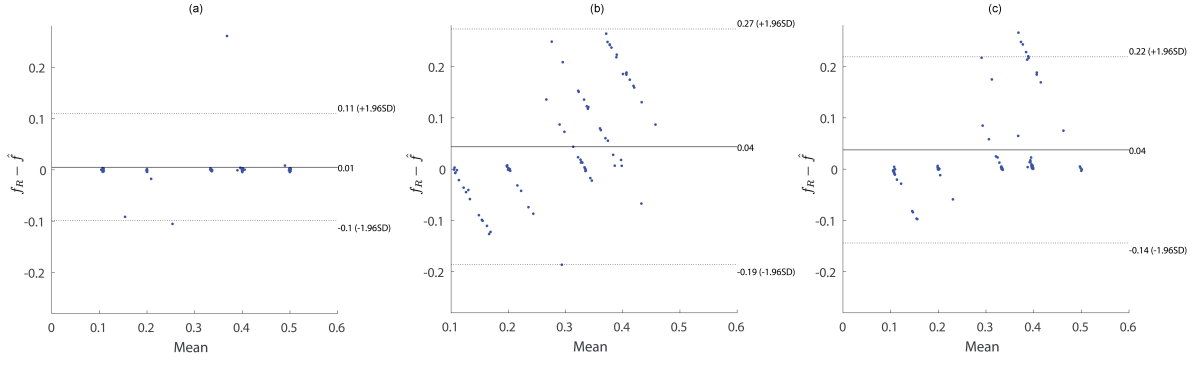


Figure 5: Bland Altman plots for fPPG (a), CvPDR (b) and CvPDR+TR (c).

records within this range (0.2-0.4 Hz) for all subject, for CvPDR $n_{OUT_{10\%}} = 22$ corresponding to 43.1% of the estimations (13 at $f_R = 0.4$ Hz) while the system CvPDR+TR presents $n_{OUT_{10\%}} = 6$ (11.7%).

- (ii) In the boundary of respiratory frequency range ($f_R = 0.1$ Hz and $f_R = 0.5$ Hz) there is a drastic drop in performances of vPPG-based systems compared to the fPPG one. At $f_R = 0.1$ Hz the decreased accuracy has a methodological aspect. According to the developed method, the initialization should enlarge respiratory frequency range to $[0.08 - 0.5]$ Hz. However this condition is not always satisfied leading to an overestimation of f_R .

Note the increase of performance in all respiratory frequencies for CvPDR when $d_{TR}(n)$ signal is taken into account. The large benefits of using $d_{TR}(n)$ signal are shown for $f_R = 0.2$ and mainly $f_R = 0.4$ Hz, where the median/IQR was clearly reduced.

4.4. Database II

The validation of the method for spontaneous breathing was carried out on the second database. As before, the results on respiratory rate obtained by vPPG-based systems were compared to the results obtained by fPPG system. It is worth noting (see Table 3) that the results are comparable looking at the median/IQR between fPPG and vPPG when tracking information is considered. Although results based only on vPPG (CvPDR) are worse than fPPG median relative error is around 1%. The slight decrease of performance for vPPG system during the recovery stage (median/IQR at

Table 3: Relative error, ϵ_R expressed as median and interquartile ranges (IQR) in Database II.

	fPPG			CvPDR			CvPDR+TR		
Phase	$\epsilon_R[\%]$	$IQR[\%]$	$n_{OUT_{10\%}}$	$\epsilon_R[\%]$	$IQR[\%]$	$n_{OUT_{10\%}}$	$\epsilon_R[\%]$	$IQR[\%]$	$n_{OUT_{10\%}}$
Spont.	0.07	1.46	2	1.10	6.36	5	0.43	2.11	2
Recov.	0.27	1.63	1	1.01	10.12	4	0.46	4.81	1

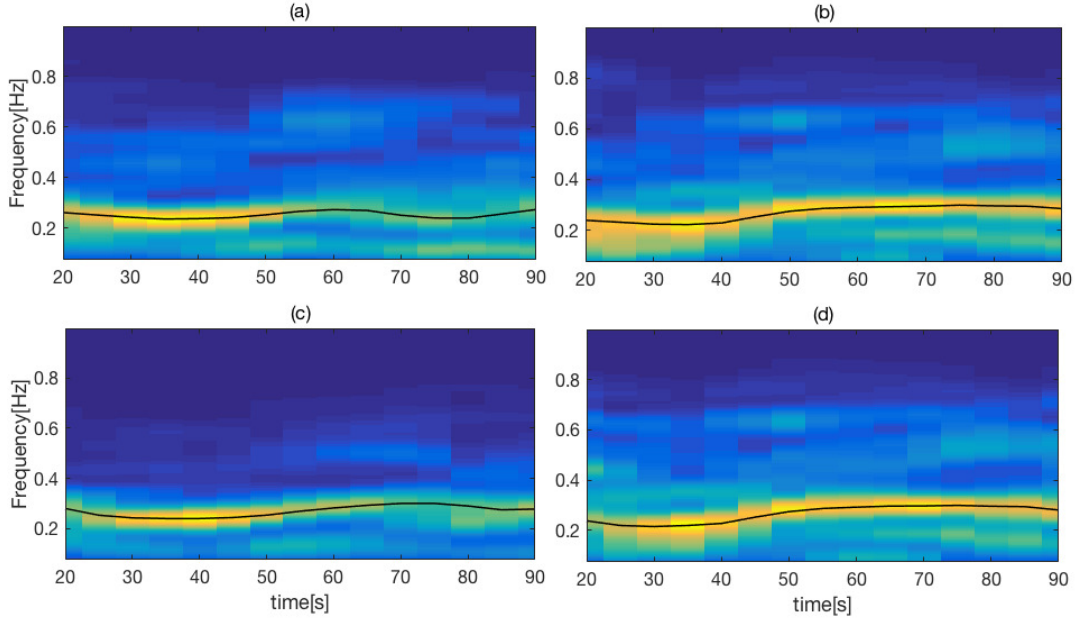


Figure 6: (a-c) represent the time-frequency analysis, $\bar{S}(k, f)$, of DR signals during spontaneous breathing session for fPPG and CvPPG+TR respectively; (b-d) represent the time-frequency analysis of recovery session respectively of fPPG and CvPPG+TR systems; the black line represents the estimated BR evaluated on a 10 s running window.

0.46/4.81%, see Table 3) is due to the reestablishment of spontaneous respiratory rate after a period of stress caused by apnea. Despite this, the fPPG system provides a low value maintaining low median/IQR value (0.27/1.63%, see Table 3).

Moreover as confirmed by the Database I, both the systems (vPPG and fPPG) failed to detect the respiratory rate of the subject 01 whose frequency was up to 0.4 Hz both in spontaneous and recovery phase.

An example of time frequency analysis is shown in Fig. 6, where the two stages of spontaneous and recovery breath were inserted respectively for the fPPG (Fig. 6(a)) and vPPG (Fig. 6(b)). As it can be seen, the two system show a similar behavior in the respiratory rate tracking over the time.

5. Discussion

The present work demonstrated the feasibility of camera-based contactless system to estimate instantaneous BR on a group of healthy subjects during two experimental settings: controlled BR and spontaneous breathing.

The first step for vPPG extraction is face detection. The classical Viola-Jones face detector (Viola and Jones, 2001) which uses cascaded classifiers on Haar-like features was used. This method provides realtime face detection, but works best for full, frontal, and well lit faces. It has been shown that in unconstrained face detection, features like Haar wavelets do not capture the discriminative facial information at different illumination

variations or poses ([Ranjan et al., 2017](#); [Liao et al., 2016](#)). So this face detection method presents some limitations such as faces in a crowd, face rotation, inclined or angled faces, expression variations and low image resolution that should be considered for BR estimation in these scenarios.

According to our results, the accuracy in respiratory frequency estimation is almost comparable between vPPG-based methods and fPPG systems. However, when concentrating on respiratory frequencies at the boundaries of the typical respiratory frequency ranges, the fPPG presents generally better results compared to the vPPG in terms of extraction of breath information from the variability of PPG amplitude, width and pulse occurrence. A possible explanation for the drastically reduced performances at very high respiratory frequencies can be given by the physiological nature of the two signals. The vPPG signal looks a low-pass filtered version of fPPG signal. Therefore RSA impact on the face is attenuated more rapidly compared to the respiratory modulation found in peripheral zones as at the finger. Secondly since the vPPG system works without the contact between the sensor and the skin, it is more sensible to artifact motion. Therefore the presence of high energy peak in the power spectrum not related to the respiration may confound the method that selects an erroneous respiratory frequency.

Bland Altman plots show a positive bias in all considered methods, indicating an underestimation of the respiratory rate. However a detailed analysis of [Fig. 5](#) indicate an overestimation at low respiratory rates. On the other hand, fPPG presents lower SD than vPPG based methods and the inclusion of the tracking signal reduce the SD of the error.

A review of methods for BR estimation based on ECG and fPPG is presented in ([Charlton et al., 2016](#)). The performance of the 314 algorithms evaluated was assessed based on 2SD whose results range from 4.7 breath per minute (bpm) at best to 50.1 bpm at worst. Our fPPG method obtained a 2SD of 6.38 bpm, being outperformed only by the best fPPG method whose 2SD was 6.2 bpm. Focusing on vPPG based methods, a mean absolute error (MAE) of 0.541 bpm for video PPG was reported in [Al-Naji et al. \(2017\)](#) using a protocol with respiratory rate ranging from 0.2 to 0.3 Hz. Constraining the analysis to this respiratory rate, our results are similar (MAE=0.537 bpm) for CvPPG+TR. However when all respiratory rates in our study are considered (0.1 to 0.6 HZ) the MAE increases up to 2.98 bpm, pointing out the relevance of considering the effect of respiratory rate in the modulation of vPPG signal. The approach presented in [Poh et al. \(2011\)](#) does not derive any DR signal, although BR is estimated from the HRV spectrum with BRs in the range of 0.16 to 0.35 Hz. Their results show a mean error and SD of 0.12 and 1.33 bpm, respectively. Our results using the same BR range show higher mean error (0.29 bpm) but lower SD (1.27 bpm). Again, the performance decreases (mean error and SD of 2.25 and 5.55 bpm respectively) when the entire BR range in our study is considered. To the best of our knowledge, only the work of [Gastel et al. \(2016\)](#) exploited the different ways respiration modulates vPPG to track instantaneous respiratory rate covering a huge range of breathing frequencies. Despite their innovative

work, the main limitation concerns with the limited number of recruited subjects which prevents to extrapolate the conclusions there referred. Beside this aspect, their method presents superior performances in higher respiratory frequencies estimation ($f_R > 0.5$) compared to the present work.

The second experimental setup aimed at testing the system with a industrial camera including a phase of spontaneous breathing and another of irregular breathing pattern provoked by recovery after apnea. The proposed method is able to estimate the respiratory rate with a IQR relative error lower than 5% both for fPPG and vPPG, corroborating the good performances obtained in the estimation of respiratory frequencies at central respiratory range during the first protocol. Interestingly, although the commercial camera produces a signal quality lower than the industrial one, the accuracy in respiratory rate estimation is still comparable with fPPG system.

In terms of selected ROI, the forehead appears to be the region with the highest SNR. A possible reason may be the region size: compared to the nose and cheek, the forehead is the largest area that allows to reduce the camera noise level averaging on a greater number of pixels (Verkruysse et al., 2008). Another explanation related to physiology could be the higher skin pulsatility of forehead. In the future, it would be desirable the development of fusion technique including signals from different ROIs adding a weighting factor depending on the SNR. Another possibility may be the use of skin segmentation techniques that avoid the selection of the best ROI performing vPPG signal extraction on the whole available face.

The performance of vPPG system improves if the tracking signal is used in combination with pulse-to-pulse methods (see Fig. 4(b),(c)). Note that a motion based system is fragile to subject motion, while vPPG-based systems may still promise motion robustness. Thus factors concerning expected breathing rates and movement should be considered in the video PPG system design depending on the final application. This aspect suggests to look for other sources of respiratory information to improve the BR estimation made by the combination algorithm. Therefore, as shown in Nam et al. (2016), in the future works, the optical reflectance induced by the respiratory abdominal motion could be taken under consideration.

Finally the proposed data fusion technique has shown some limitations in the estimation of BR with vPPG system. Indeed, the method was tuned to present higher accuracy estimation to respiratory frequencies in the center of respiratory range; as a drawback the relative error increases considerably at the boundary of respiratory frequency range. Therefore another aspect is to consider the use of machine learning tools, as neural networks techniques, that could result in a more robust combination of input information.

6. Conclusion

The present work has shown the possibility to estimate instantaneous BR from vPPG. The method derives respiration from beat-to-beat PPG rate and morphology changes

in amplitude and width driven by the physiological relationships between vPPG and respiration. Moreover, respiratory-induced head movements were used as additional source of information for the vPPG system. Results of a combination of these methods show that respiratory frequencies in the central range (0.2-0.4 Hz) may be estimated using vPPG system with low relative error, $\epsilon_R < 2\%$ and interquartile range of order $IQR < 5\%$. However, the vPPG system showed a drop in performance at respiratory range boundaries, around 0.1 Hz and 0.5 Hz, which shows that respiratory modulation of vPPG signal is affected by BR. Therefore vPPG-based systems should consider the effect of the BR in the estimation, mainly in applications where respiratory rate is out of the range 0.2-0.4 Hz.

Acknowledgment

This work has been partially financed by Ministerio de Economía, Industria y Competitividad (MINECO) and by fondos FEDER through the project RTI2018-097723-B-I00 and DPI2016-75458-R, and by CIBER in Bioengineering, Biomaterials & Nanomedicine (CIBER-BBN) through Instituto de Salud Carlos III. This work has received funding from the European Unions Framework Programme for Research and Innovation Horizon 2020 (2014-2020) under the Marie Skłodowska-Curie Grant Agreement No. 745755 .

References

- Abbas, A. K., Heimann, K., Jergus, K., Orlikowsky, T. and Leonhardt, S. (2011), ‘Neonatal non-contact respiratory monitoring based on real-time infrared thermography’, *Biomed. Eng. Online* **10**(93), 1–17.
- AL-Khalidi, F., Saatchi, R., Burke, D., Elphick, H. and Tan, S. (2011), ‘Respiration rate monitoring methods: A review’, *Pediatric Pulmonology* **46**(6), 523–529.
- Al-Naji, A. and Chahl, J. (2016), ‘Remote respiratory monitoring system based on developing motion magnification technique’, *Biomedical Signal Processing and Control* **29**, 1–10.
- Al-Naji, A., Gibson, K. and Chahl, J. (2017), ‘Remote sensing of physiological signs using a machine vision system’, *Journal of Medical Engineering & Technology* **41**(5), 396–405.
- Bailón, R., Sörnmo, L. and Laguna, P. (2006), ‘A robust method for ECG-based estimation of the respiratory frequency during stress testing’, *IEEE Trans. Biomed. Eng.* **53**, 1273–1285.
- Bartula, M., Tigges, T. and Muehlsteff, J. (2013), Camera-based system for contactless monitoring of respiration, *in* ‘35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan’, pp. 2672–2675.

- Bell, A. J. and Sejnowski, T. J. (1997), ‘The ‘Independent Components’ of natural scenes are edge filters’, *Vision Res.* **37**, 3327–3338.
- Bennett, S. L., Goubran, R. and Knoefel, F. (2017), Comparison of motion-based analysis to thermal-based analysis of thermal video in the extraction of respiration patterns, in ‘39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju , Korea’, pp. 3835–3839.
- Charlton, P. H., Birrenkott, D. A., Bonnici, T., Pimentel, M. A., Johnson, A. E., Alastruey, J., Tarassenko, L., Watkinson, P. J., Beale, R. and Clifton, D. A. (2017), ‘Breathing rate estimation from the electrocardiogram and photoplethysmogram: A review’, *IEEE reviews in biomedical engineering* **11**, 2–20.
- 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’, *Physiological measurement* **37**(4), 610.
- Cohen, K. P., Ladd, W. M., Beams, D. M., Sheers, W. S., Radwin, R. G., Tompkins, W. J. and Webster, J. G. (1997), ‘Comparison of impedance and inductance ventilation sensors on adults during breathing, motion, and simulated airway obstruction’, *IEEE Transactions on Biomedical Engineering* **44**(7), 555–566.
- Corbishley, P. and Rodriguez-Villegas, E. (2007), ‘Breathing detection: towards a miniaturized, wearable, battery-operated monitoring system’, *IEEE Transactions on Biomedical Engineering* **55**(1), 196–204.
- Cretikos, M. A., Bellomo, R., Hillman, K., Chen, J., Finfer, S. and Flabouris, A. (2008), ‘Respiratory rate: the neglected vital sign’, *MJA* **188**(11), 657–659.
- de Haan, G. and Jeanne, V. (2013), ‘Robust pulse rate from chrominance-based rPPG’, *IEEE Trans. On Biomed. Eng.* **60**, 2878–86.
- 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. Intern. Med.* **8**(7), 354–360.
- Folke, M., Cernerud, L., Ekström, M. and Hök, B. (2003), ‘Critical review of non-invasive respiratory monitoring in medical care’, *Medical and Biological Engineering and Computing* **41**(4), 377–383.
- Garde, A., Karlen, W., Dehkordi, P., Ansermino, J. and Dumont, G. (2013), Empirical mode decomposition for respiratory and heart rate estimation from the photoplethysmogram, in ‘Computing in Cardiology 2013, Zaragoza, Spain’, pp. 799–802.
- Gastel, M. V., Stuijk, S. and Haan, G. D. (2016), ‘Robust respiration detection from remote photoplethysmography’, *Biomed. Opt. Express* **7**(12), 4941–4957.
- Gravelyn, T. R. and Weg, J. G. (1980), ‘Respiratory rate as an indicator of acute respiratory dysfunction’, *JAMA* **244**(10), 1123–1125.
- Hochhausen, N., Pereira, C. B., Leonhardt, S., Rossaint, R. and Czaplik, M. (2018),

- ‘Estimating respiratory rate in post-anesthesia care unit patients using infrared thermography: An observational study’, *Sensors* **18**(5).
- Iozzia, L., Cerina, L. and Mainardi, L. (2016), ‘Relationships between heart-rate variability and pulse-rate variability obtained from video-PPG signal using ZCA’, *Physiol. Meas.* **37**, 1934–1944.
- Iozzia, L., Lázaro, J., Gil, E., Cerina, L., Mainardi, L. and Laguna, P. (2017), ‘Respiratory rate detection using a camera as contactless sensor’, in ‘Computing in Cardiology 2017, Rennes, France’, Vol. 44.
- Janssen, R., Wang, W. and Moo, A. (2016), ‘Video-based respiration monitoring with automatic region of interest detection’, *Physiological Measurement* **37**, 100–114.
- Lázaro, J., Gil, E., Bailón, R., Mincholé, A. and Laguna, P. (2013), ‘Deriving respiration from photoplethysmographic pulse width’, *Medical & Biological Engineering & Computing* **51**(1), 233–242.
- Lázaro, J., Gil, E., Vergara, J. M. and Laguna, P. (2014), ‘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.* **18**, 240–246.
- Lázaro, J., Nam, Y., Gil, E., Laguna, P. and Chon, K. H. (2015), ‘Respiratory rate derived from smartphone-camera-acquired pulse photoplethysmographic signals’, *Physiol. Meas.* **36**(11), 2317–2333.
- Lewis, G. F., Gatto, R. G. and Porges, S. W. (2011), ‘A novel method for extracting respiration rate and relative tidal volume from infrared thermography’, *Psychophysiology* **48**(7), 877–887.
- Liao, S., Jain, A. K. and Li, S. Z. (2016), ‘A fast and accurate unconstrained face detector’, *IEEE Transactions on Pattern Analysis Machine Intelligence* **38**(02), 211–223.
- Lucas, B. D. and Kanade, T. (1991), ‘Detection and tracking of point features’, *Technical Report MU-CS-91-132, Carnegie Mellon University*.
- Massaroni, C., Nicolò, A., Lo Presti, D., Sacchetti, M., Silvestri, S. and Schena, E. (2019), ‘Contact-based methods for measuring respiratory rate’, *Sensors* **19**(4), 908.
- Mateo, J. and Laguna, P. (2003), ‘Analysis of heart rate variability in the presence of ectopic beats using heart timing signal’, *IEEE Trans. Biomed. Eng.* **50**, 334–343.
- 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’, *Journal of Medical Engineering and Technology* **36**, 1–7.
- Nakajima, K., Matsumoto, Y. and Tamura, T. (2001), ‘Development of real-time image sequence analysis for evaluating posture change and respiratory rate of a subject in bed’, *Physiol Meas.* **22**(3), N21–N28.
- Nam, Y., Kong, Y., Reyes, B., Reljin, N. and Chon, K. H. (2016), ‘Monitoring of Heart and Breathing Rates Using Dual Cameras on a Smartphone’, *Plos One* **11**(3), 1–15.

- Pereira, C. B., Heimann, K., Venema, B., Blazek, V., Czaplik, M. and Leonhardt, S. (2017), Estimation of respiratory rate from thermal videos of preterm infants, *in* ‘39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea’, pp. 3818–3821.
- Pereira, C. B., Yu, X., Blazek, V., Venema, B. and Leonhardt, S. (2016), Multisensor data fusion for enhanced respiratory rate estimation in thermal videos, *in* ‘38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, USA’, pp. 1381–1384.
- Poh, M.-Z., McDuff, D. J. and Picard, R. W. (2011), ‘Advancements in Noncontact, Multiparameter Physiological Measurements Using a Webcam’, *IEEE Trans. Biomed.Eng.* **58**(1), 7–11.
- Procházka, A., Charvátová, H., Vysata, O., Kopal, J. and Chambers, J. (2017), ‘Breathing analysis using thermal and depth imaging camera video records’, *Sensors* **17**(6).
- Ranjan, R., Patel, V. M. and Chellappa, R. (2017), ‘Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition’, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **41**(1), 121–135.
- Regev, N. and Wulich, D. (2017), A simple, remote, video based breathing monitor, *in* ‘39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea’, pp. 1788–1791.
- Rehoumaa, H., Noumeira, R., Bouachirb, W., Jouvetc, P. and Essourica, S. (2018), ‘3d imaging system for respiratory monitoring in pediatric intensive care environment’, *Computerized Medical Imaging and Graphics* **70**, 17–28.
- Shao, D., Yang, Y., Liu, C., Tsow, F., Yu, H. and Tao, N. (2014), ‘Noncontact Monitoring Breathing Pattern, Exhalation Flow Rate and Pulse Transit Time’, *IEEE Trans. Biomed.Eng.* **61**(11), 2760–2767.
- Sin, P. Y. W., Galletly, D. C. and Tzeng, Y. C. (2010), ‘Influence of breathing frequency on the pattern of respiratory sinus influence of breathing frequency on the pattern of respiratory sinus’, *Am J Physiol Heart Circ Physiol* pp. H1588–H1599.
- Tarassenko, L., Villarroel, M., Guazzi, A. and Jorge, J. (2014), ‘Non-contact video-based vital sign monitoring using ambient light and auto-regressive models’, *Physiological Measurement* pp. 807–831.
- Tarvainen, M. P., Ranta-Aho, P. O. and Karjalainen, P. A. (2002), ‘An advanced detrending method with application to HRV analysis’, *IEEE Trans. Biomed. Eng.* **49**, 172–175.
- Verkruysse, W., Svaasand, L. O. and Nelson, J. S. (2008), ‘Remote plethysmographic imaging using ambient light’, *Opt. Exp.* **16**(26), 21434–45.
- Villarroel, M., Guazzi, A., Jorge, J., Davis, S., Watkinson, P., Green, G., Shenvi, A., McCormick, K. and Tarassenko, L. (2014), ‘Continuous non-contact vital sign

- monitoring in neonatal intensive care unit', *Healthcare Technology Letters* **1**(3), 87–91.
- Viola, P. and Jones, M. (2001), 'Rapid object detection using a boosted cascade of simple features', *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Kauai, USA* pp. 511–517.
- Yasuma, F. and Hayano, J.-I. (2004), 'Respiratory sinus arrhythmia', *CHEST* **125**(2), 683–90.
- Zhao, F., Li, M., Qian, Y. and Tsien, J. Z. (2013), 'Remote measurements of heart and respiration rates for telemedicine', *PLoS ONE* **8**(10).