



## PAPER

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

Luca Iozza<sup>1</sup>, Jesús Lázaro<sup>2,3,4</sup>, Luca Cerina<sup>1</sup>, Davide Silvestri<sup>5</sup>, Luca Mainardi<sup>1</sup>, Pablo Laguna<sup>2,3</sup> and Eduardo Gil<sup>2,3</sup> <sup>1</sup> Dipartimento di Elettronica, Informatica e Bioingegneria, Politecnico di Milano, Milan, Italy<sup>2</sup> BSI CoS Group, Aragon Institute of Engineering Research (I3A), IIS Aragon, University of Zaragoza, Zaragoza, Spain<sup>3</sup> Centro de Investigación Biomedica en Red—Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Zaragoza, Spain<sup>4</sup> Department of Biomedical Engineering, University of Connecticut, Storrs, CT, United States of America<sup>5</sup> University of Bologna, Bologna, ItalyE-mail: [edugilh@unizar.es](mailto:edugilh@unizar.es)

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

RECEIVED  
1 May 2019REVISED  
27 August 2019ACCEPTED FOR PUBLICATION  
3 September 2019PUBLISHED  
27 September 2019

## 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 is evidence 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-continuous estimation of BR considering the effect of respiration on video-PPG (vPPG). **Approach:** The technology has been tested in two different experimental settings, including controlled BR and more challenging scenarios with spontaneous breathing patterns. 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 an additional source of information for the vPPG system. A combination of these methods has been exploited to estimate the respiratory rate every 10 seconds. **Main results:** According to the results, respiratory frequencies in the central range (0.2–0.4 Hz) may be estimated using the vPPG system with a low relative error,  $\epsilon_R < 2\%$  and interquartile range of the 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 an ubiquitous BR monitoring system. However, vPPG-based systems should consider the effect of the BR in the estimation, mainly in applications where the respiratory rate is out of the 0.2–0.4 Hz range.

## 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 (Fieselmann *et al* 1993, Cretikos *et al* 2008). Fieselmann showed that a BR greater than 27 breaths per minute is a predictor of cardiac arrest, and is even more informative compared to cardiac pulse and blood pressure changes. Moreover, Gravelyn and Weg (1980) demonstrated that BR is a specific marker of acute respiratory dysfunction such as hypoxemia (abnormally low level of O<sub>2</sub> in the blood) and hypercapnia (abnormal CO<sub>2</sub> in the blood).

Different methods have been exploited to monitor respiratory activity: airflow sensing technology uses a nasal sensor installed on a mouthpiece of a face mask 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 CO<sub>2</sub> than inspiratory air (Folke *et al* 2003); impedance plethysmography measures variations in transthoracic electrical impedance to derive ventilation (Cohen *et al* 1997); respiratory inductance plethysmography, which 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,

e.g. cumbersome for the patient, time consuming, or does not allow for long-term monitoring (Al-Khalidi *et al* 2011, Massaroni *et al* 2019). Nowadays there is still a lack of low-cost home care solutions that can be ubiquitous and allow for daily monitoring of respiratory activity.

In state-of-the-art methods, a step forward in the improvement of patient compliance has been obtained by the introduction of noncontact devices. There are clear advantages to noncontact 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 with respiration, basically chest and/or abdomen movement (Nakajima *et al* 2001, Bartula *et al* 2013, Zhao *et al* 2013, Al-Naji and Chahl 2016, Janssen *et al* 2016, Nam *et al* 2016, Regev and Wulich 2017, Rehoumaa *et al* 2018) or even shoulder movement (Shao *et al* 2014). In absence of other movement artefacts, these systems estimate the respiratory rate, and many of them have been validated while subjects are lying in bed; these subjects include neonates (Janssen *et al* 2016, Rehoumaa *et al* 2018), babies (Al-Naji and Chahl 2016, Regev and Wulich 2017), and adults (Nakajima *et al* 2001, Bartula *et al* 2013). Thermal cameras allow for respiratory rate estimation by monitoring dynamic thermal activity emitted from specific areas (usually nostrils) since the temperature of exhaled air is higher than that of inhaled air (Abbas *et al* 2011, Lewis *et al* 2011, Pereira *et al* 2017, Hochhausen *et al* 2018). In addition, some authors have combined both methods in order to increase the accuracy of respiratory rate estimation (Pereira *et al* 2016, Bennett *et al* 2017, Procházka *et al* 2017).

A contactless alternative is represented by the video-photoplethysmographic (vPPG) system. In the 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 as a method to derive respiratory features as BR. In Poh *et al* (2011), the BR was extracted as a component of the power spectrum of PRV derived from the vPPG signal. In Villarroel *et al* (2014) and Tarassenko *et al* (2014), a band pass filter tuned to the expected BR has been applied to the vPPG signal. Hence, the peak in the spectrum has been considered as the BR. Similarly, Al-Naji *et al* (2017) combined information of head movement and the 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 subregions. For each region, a set of weights was calculated using the chrominance method (already implemented in de Haan and Jeanne (2013)), which allows us to enhance the respiratory information and suppress motion distortions. Then the best set of weights was selected by the signal to noise ratio (SNR) quality parameter and used to exploit the breathing signal.

As shown in the literature, the 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 the PPG waveform (Garde *et al* 2013, Lázaro *et al* 2015). 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 the body to save energy by suppressing unnecessary heartbeats and increasing the efficiency of ventilation during the end of perfusion. Therefore, the respiratory component can be extracted from the 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, and a reduction in venous pressure. As second 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 BR.

Finally, as shown by Lázaro *et al* (2013), the sympathetic nervous system regulates the vascular resistance by means of vein 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 in the pulse width of the PPG signal. Therefore, the pulse width variability (PWV) is affected by respiratory modulation.

Focusing on noncontact respiration methods based on *photoplethysmography*, there is a need to characterize 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 since the physiological modulation of PPG by respiration depends on it. For instance, it is well known that RSA decreases at high respiratory rates (Sin *et al* 2010). Moreover, in the case of vPPG, this decrease comes with a low-passed shape in comparison to finger-PPG (fPPG), which increases the uncertainty in locating the pulse and therefore reduces the 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, e.g. during physical exercise, relax conditions or some clinical applications, among others. The aim of the present work is to test the feasibility of a contactless, camera-based method that estimates BR values based on vPPG signals while considering the effect of respiration on these vPPG signals. In addition, the added value of using movement information for the estimation will be evaluated. Compared to existing state-of-the-art methods, the proposed method's novelty lies in several aspects: providing an exhaustive analysis of the effect of respiratory rate in the estimation; implementation of a data fusion technique that combines breathing information derived from PPG waveform changes, PRV, PAV, PWV, and vertical head movements; assessment of PPG-derived respiratory (DR) modulation measured at two different body sites (finger versus head) using fPPG and vPPG technology, respectively; use of two different technological cameras in order to test the robustness of the implemented BR estimation. A preliminary version of this work has been reported (Iozza et al 2017).

## 2. Materials

### 2.1. Experimental protocols

Two experiments were conducted involving 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 artefacts, participants were instructed to move their head as little as possible. In both experiments the PPG signal was recorded from the index finger whereas the respiratory signal was recorded by a respiratory belt fixed to the subject's chest. Both signals were sampled at 256 Hz using the FlexComp Infinity™ 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  $1280 \times 720$  with a frame rate of 60 fps. The subject was informed to breathe at a constant rate, by using a set of breathing guidance videos showing a bar moving from 0.1 to 0.5 Hz at increments of 0.1 Hz. A 2 minute recording was made for each respiratory frequency.

**Experiment II:** In the second protocol a GigE Sony XCG-C30C camera was selected with a spatial resolution of  $659 \times 494$  pixels and 60 fps. The camera was equipped with a 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 min of spontaneous breathing, a period of apnea (whose duration depended on the subject's capacity), and 2 min of recovery.

The data in the second database were divided into two stages: *spontaneous breath* lasted for 110 s of the recordings and *recovery breath* lasted for 110 s after the apnea phase. Each stage was treated as independent, and the estimation of respiratory rate was conducted for each stage except for the apnea stage.

## 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 were processed offline to derive the vPPG signal according to the algorithm presented in Iozza et al (2016). Three regions of interest (ROI) were considered: forehead (ROI<sub>1</sub>), nose (ROI<sub>2</sub>), and cheek (ROI<sub>3</sub>). 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 the X–Y axis (horizontal and vertical movements, respectively) 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 the LKT feature tracker using a linear model for the prediction and the Y axis coordinate as observations of the LKT centroid tracker output.

Within each ROI a spatial average of pixel intensities of each channel (red, green, and blue) was 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 (finite impulse response filter using the *Hamming window*) with cut-off frequencies of  $f_{c1} = 0.1$  Hz and  $f_{c2} = 5$  Hz. To enhance the pulsatile component of the reflected light and reduce the motion noise, zero-phase component analysis (Bell and Sejnowski 1997) was applied to each ROI. As a result,  $M$  vPPG signals were obtained. 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:

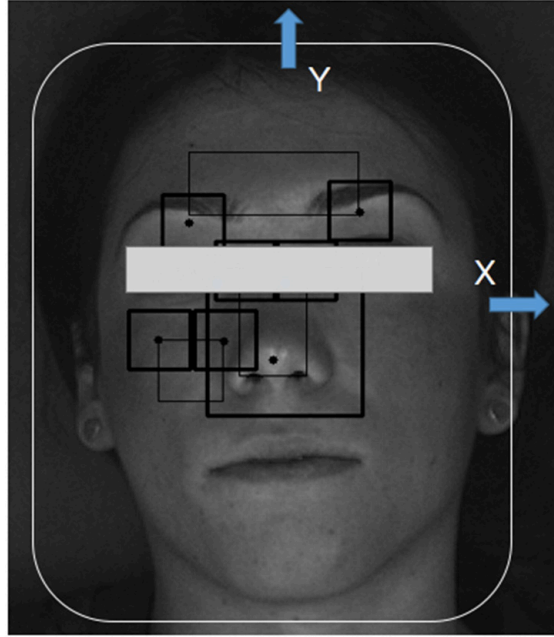


Figure 1. Tracking of the face according to horizontal ( $X$  axis) and vertical ( $Y$  axis) movement.

$$\text{SNR}_i = 10 * \log_{10} \left( \frac{\int_{f_1}^{f_2} \mathcal{P}_{\text{vPPG}_i}(f) df}{\int_{0.1}^{f_1} \mathcal{P}_{\text{vPPG}_i}(f) df + \int_{f_2}^5 \mathcal{P}_{\text{vPPG}_i}(f) df} \right), \quad (1)$$

where  $\mathcal{P}_{\text{vPPG}_i}(f)$  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  $f_p$  is the highest peak assumed to be the pulse frequency (measured in Hz). The signal with the highest SNR was selected as the target vPPG signal. Subsequently, the motion artefacts were automatically detected using the  $X$ – $Y$  coordinates of facial movement (as shown in figure 1). Within a sliding window 1 s long, the first derivatives of  $X$  and  $Y$  coordinates (defined as  $v_X$  and  $v_Y$ ) were calculated. If the mean of  $v_X$  or  $v_Y$  calculated inside the interval time was higher than a threshold (manually selected as 14 pixels  $s^{-1}$ ), the corresponding vPPG segment was labeled as artefactual and discarded from further analysis (see figure 2). Less than 5% of the total number of segments were excluded from the successive procedure.

### 3.2. Detection of fiducial points

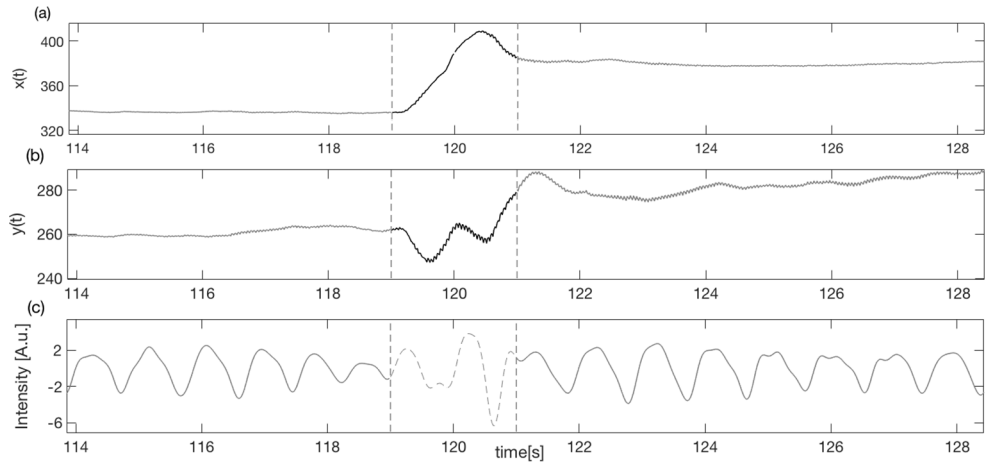
To derive the PWV, PRV, and PAV signals from both fPPG and vPPG, the method implemented by Lázaro *et al* (2013) was applied. Briefly, the systolic peaks  $n_{A_i}$  were identified by an automatic PPG pulse detector based on a low-pass filter differentiator and an adaptive time-varying threshold (Lázaro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

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

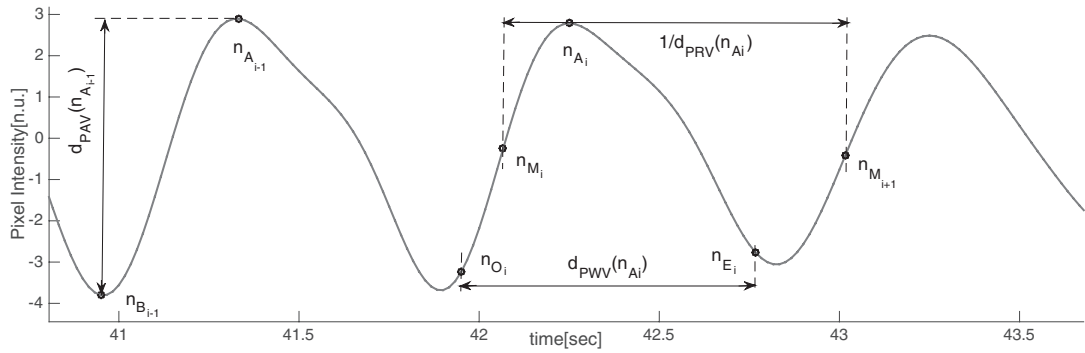
where  $s(n)$  is the PPG signal, and  $f_s$  the sampling frequency of the signal. Finally,  $n_{M_i}$  was detected as the half of the pulse amplitude of the cardiac pulse:

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

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



**Figure 2.** Example of artefact movement removal on vPPG signal (c) using the X–Y tracking signals (a) and (b) as references of noise.



**Figure 3.** DR signals defined by a set of fiducial points.

### 3.3. DR signals

#### 3.3.1. From PPG signals

Three DR signals were calculated using pulse-to-pulse methods based on PRV, PAV, and PWV. The DR signal based on PRV was obtained as a pulse-to-pulse series:

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

where  $u$  indicates that the signal is unevenly sampled, and  $n_{N_i}$  represents the arrival of the pulse after removing the ectopic and miss-detected pulses from  $n_{M_i}$ , by using the method proposed in Mateo and Laguna (2003). The other two derived signals, i.e. those based on PAV and PWV, were obtained as follows:

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

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

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. Nonrespiratory 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, the optimal range 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 the  $Y$  coordinate of the face movement returned from the LKT algorithm. Since respiration affects the vertical movement of the face, the  $y(t)$  signal derived from the  $Y$ -axis movement over

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

### 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)). The effect of including the tracking signal to increase the accuracy of the estimation is also considered (CvPDR+TR). Let us denote  $S_j(k, f)$  as the power spectral densities for the  $j$ th DR signal at every time instant  $k$ . From these spectra, an average is computed as follows:

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

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

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

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

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

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  $[0.15-0.5] \text{ Hz}$ :

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

If in the spectrum of each DR signal no peak has enough energy in the considered frequency band, the range is enlarged to  $[0.08-0.5] \text{ Hz}$  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  $[\hat{f}(k-1) - 0.075; \hat{f}(k-1) + 0.15] \text{ Hz}$ , where  $\hat{f}(k-1)$  is the previous estimated value. As can be imagined, the phase of initialization of the algorithm is decisive since subsequent respiratory rate estimations will be based on a window defined by the previous value  $\hat{f}(k-1)$ . Due to the morphology of the vPPG signal, for higher respiratory frequencies (e.g.  $f \geq 0.3 \text{ Hz}$ ) the modulation of the breath is less evident compared to the modulation presented on fPPG. To account for this, a modification of Lázaro's algorithm is proposed when the tracking signal is considered. Since the tracking DR signal is uncorrelated with the breath modulation on the 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 for  $d_{\text{PRV}}(n)$ ,  $d_{\text{PAV}}(n)$ ,  $d_{\text{PWV}}(n)$ , and  $d_{\text{TR}}(n)$  signals. Two  $\hat{f}$  values were estimated:  $\hat{f}_{\text{vPPG}}$  and  $\hat{f}_{\text{TR}}$ . Lastly, the initial respiratory frequency was established as follows:

**Table 1.** Percentage of use of each ROI according to the selected database.

| ROI      | Database I (%) | Database II (%) |
|----------|----------------|-----------------|
| Forehead | 60             | 61.11           |
| Nose     | 30             | 11.11           |
| Cheek    | 10             | 27.78           |

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

where  $\Delta = \hat{f}_{\text{TR}} - \hat{f}_{\text{vPPG}}$ . 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  $\hat{f}_{\text{TR}}$ .

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. Camera comparisons

To evaluate the performance of the two cameras on the extraction of the vPPG signal, the SNR of the vPPG signal of each subject was calculated according to the equation (1) and averaged for the whole population (as *mean*  $\pm$  *standard deviation*). The vPPG quality signal obtained from the commercial camera was significantly lower ( $9.06 \pm 3.84$  dB) compared to the one obtained from the industrial camera ( $12.20 \pm 4.73$  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 noticing that the *forehead* was the most selected region no matter the proposed experimental setting.

### 4.3. Database I

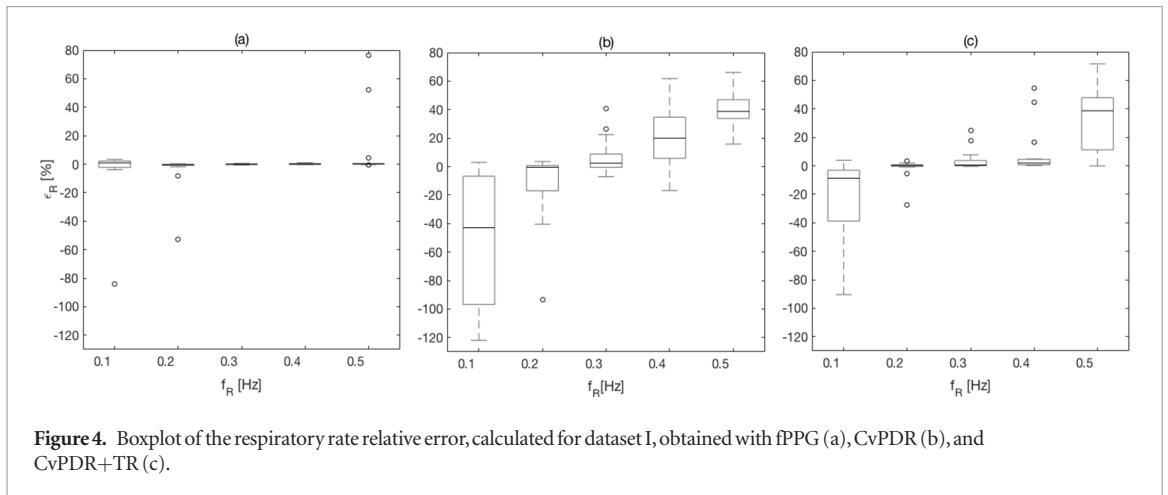
Results are shown in table 2, figures 4 and 5 for fPPG (a), CvPDR (b), and CvPDR+TR (c). Comparison of the methods shows different behaviors for different ranges:

- (i) The best results are obtained if the BR is in the range of [0.2–0.4] Hz. The median and interquartile ranges (IQRs) of  $\epsilon_R$  are below 5% for fPPG and CvPDR+TR, but are higher for CvPDR. In particular, for the fPPG there is only one estimation with a relative error higher than 10% (defined as  $n_{\text{OUT}_{10\%}}$ ) among all records within this range (0.2–0.4 Hz) for all subjects. For CvPDR,  $n_{\text{OUT}_{10\%}} = 22$  corresponds to 43.1% of the estimations (13 at  $f_R = 0.4$  Hz) while for CvPDR+TR  $n_{\text{OUT}_{10\%}} = 6$  (11.7%).
- (ii) In the boundary of the respiratory frequency range ( $f_R = 0.1$  Hz to  $f_R = 0.5$  Hz) there is a drastic drop in performances of vPPG-based systems compared to the fPPG system. At  $f_R = 0.1$  Hz, the decreased accuracy has a methodological reason. According to the developed method, the initialization should enlarge the 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 in performance in all respiratory frequencies for CvPDR when the  $d_{\text{TR}}(n)$  signal is taken into account. The large benefits of using the  $d_{\text{TR}}(n)$  signal are shown for  $f_R = 0.2$  and mainly  $f_R = 0.4$  Hz, where the median/IQR is clearly reduced.

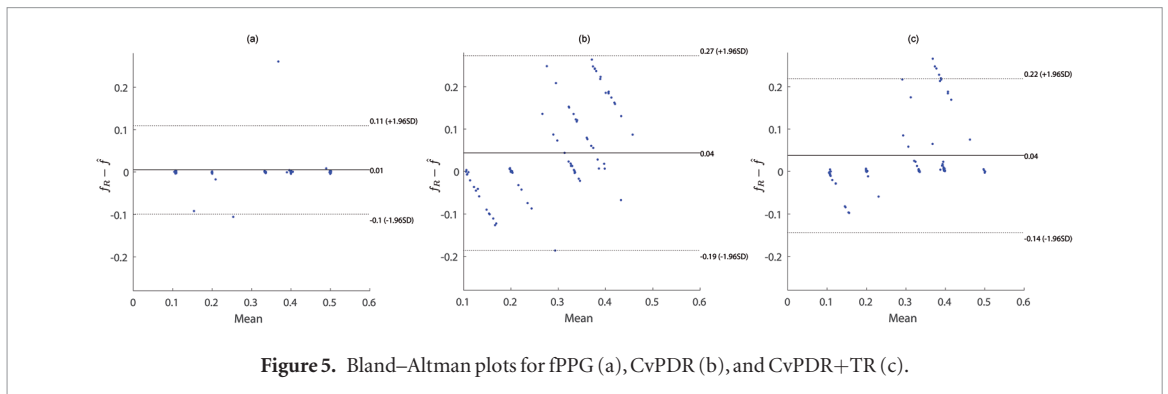
### 4.4. Database II

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 the 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, its median relative error is around 1%. The slight decrease in performance for the vPPG system during the recovery stage (median/IQR at 0.46/4.81%, see table 3) is due to the re-establishment of spontaneous



**Table 2.** Relative error  $\epsilon_R$  in database I, expressed as median and interquartile ranges (IQRs).

| $f_R$ (Hz) | fPPG             |         |                  | CvPDR            |         |                  | CvPDR+TR         |         |                  |
|------------|------------------|---------|------------------|------------------|---------|------------------|------------------|---------|------------------|
|            | $\epsilon_R$ (%) | IQR (%) | $n_{OUT_{10\%}}$ | $\epsilon_R$ (%) | IQR (%) | $n_{OUT_{10\%}}$ | $\epsilon_R$ (%) | 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               |



respiratory rate after a period of stress caused by apnea. Despite this, the fPPG system maintains a low median/IQR value (0.27/1.63%, see table 3).

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

An example of time frequency analysis is shown in figure 6, where the stages of spontaneous and recovery breath are shown for the fPPG (figure 6(a) and figure 6(b) respectively) and vPPG (figure 6(c) and figure 6(d), respectively). As can be seen, the two methods show similar behavior in the respiratory rate tracking over the time.

## 5. Discussion

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

The first step for vPPG extraction is face detection. The classical Viola–Jones detector (Viola and Jones 2001), which uses cascaded classifiers on Haar-like features, was used. This method provides real-time 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 (Liao et al 2016, Ranjan et al 2017). Therefore, this face detection method has some limitations when dealing with faces in a crowd, face rotation, inclined or angled faces, expression variations, and low image resolution. These should be considered in BR estimation.



**Table 3.** Relative error  $\epsilon_R$  in database II, expressed as median and IQRs.

| Phase  | fPPG             |         |                  | CvPDR            |         |                  | CvPDR+TR         |         |                  |
|--------|------------------|---------|------------------|------------------|---------|------------------|------------------|---------|------------------|
|        | $\epsilon_R$ (%) | 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                |

According to our results, the accuracy of 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 in 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 is a low-pass-filtered version of the fPPG signal. Therefore, RSA's impact on the face is attenuated more rapidly compared to the respiratory modulation found in peripheral zones such as the finger. Moreover, since the vPPG system works without contact between the sensor and the skin, it is more sensitive to artefact motion. Therefore, the presence of a high energy peak in the power spectrum not related to the respiration may confound the method that selects an erroneous respiratory frequency.

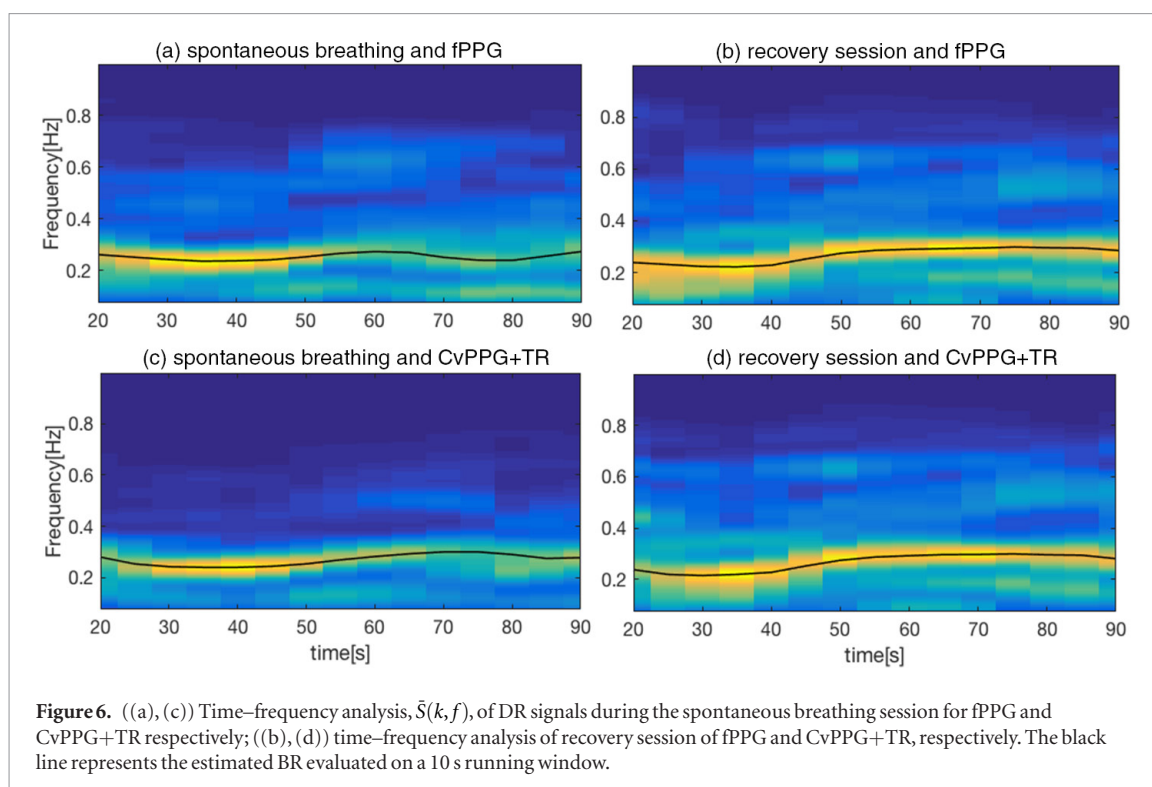
Bland–Altman plots show a positive bias in all considered methods, indicating an underestimation of the respiratory rate. However, a detailed analysis of figure 5 indicates an overestimation at low respiratory rates. Additionally, fPPG presents a lower standard deviation (SD) than vPPG-based methods, and the inclusion of the tracking signal reduces the SD of the error.

A review of methods for BR estimation based on electrocardiogram (ECG) and fPPG is presented in Charlton *et al* (2016). The performance of the 314 algorithms evaluated was assessed based on 2SD; the results ranged from 4.7 breaths 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 vPPG 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; this points out the relevance of considering the effect of respiratory rate in the modulation of the vPPG signal. The approach presented in Poh *et al* (2011) does not derive any DR signal, although BR is estimated from the heart rate variability spectrum with BRs in the range of 0.16–0.35 Hz. Their results showed a mean error and SD of 0.12 and 1.33 bpm, respectively. Using the same BR range, our results had a 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 is concerned with the limited number of recruited subjects, which prevents us from extrapolating their conclusions. Beside this, their method presents superior performances in higher respiratory frequencies estimation ( $f_R > 0.5$ ) compared to the present work.

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

In terms of selected ROIs, 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, and it allows us to reduce the camera noise level by averaging a greater number of pixels (Verkruyssen *et al* 2008). Another explanation related to physiology could be the higher skin pulsatility of the forehead. In the future, it would be desirable to develop a fusion technique including signals from different ROIs and 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 and perform vPPG signal extraction on the whole available face.

The performance of the vPPG system improves if the tracking signal is used in combination with pulse-to-pulse methods (see figures 4(b) and (c)). Note that a motion-based system is sensitive to subject motion while vPPG-based systems promise motion robustness. Thus, factors concerning expected BRs and movement should be considered in the vPPG system design depending on the final application. We should look for other sources of respiratory information to improve the BR estimation made by the algorithm. Therefore, as shown in Nam *et al*



(2016), in future works the optical reflectance induced by respiratory abdominal motion could be taken under consideration.

Finally, the proposed data fusion technique has shown some limitations in the estimation of BR with the vPPG system. Indeed, the method was tuned to present a higher accuracy estimation with respiratory frequencies in the center of the respiratory range; as a drawback, the relative error increases considerably at the boundary of the respiratory frequency range. Therefore, another aspect to consider is the use of machine learning tools, such as neural networks, which 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 the 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 an 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 the vPPG system with a low relative error,  $\epsilon_R < 2\%$  and IQR of the order  $\text{IQR} < 5\%$ . 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 the vPPG signal is affected by BR. Therefore, vPPG-based systems should consider the effect of the BR in the estimation, mainly in applications where the respiratory rate is out of the 0.2–0.4 Hz range.

## Acknowledgment

This work has been partially financed by Ministerio de Economía, Industria y Competitividad (MINECO) and by fondos FEDER through projects 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.

## ORCID iDs

Eduardo Gil  <https://orcid.org/0000-0001-7285-0715>

## 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** 1–17
- Al-Khalidi F, Saatchi R, Burke D, Elphick H and Tan S 2011 Respiration rate monitoring methods: a review *Pediatric Pulmonol.* **46** 523–9
- Al-Naji A and Chahl J 2016 Remote respiratory monitoring system based on developing motion magnification technique *Biomed. Signal Process. Control* **29** 1–10
- Al-Naji A, Gibson K and Chahl J 2017 Remote sensing of physiological signs using a machine vision system *J. Med. Eng. Technol.* **41** 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** 1273–85
- Bartula M, Tigges T and Muehlsteff J 2013 Camera-based system for contactless monitoring of respiration *35th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Osaka, Japan)* pp 2672–5
- Bell A J and Sejnowski T J 1997 The ‘independent components’ of natural scenes are edge filters *Vis. Res.* **37** 3327–38
- Bennett S L, Goubiran R and Knoefel F 2017 Comparison of motion-based analysis to thermal-based analysis of thermal video in the extraction of respiration patterns *39th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Jeju, Korea)* pp 3835–9
- 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 Rev. Biomed. Eng.* **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 *Physiol. Meas.* **37** 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 Trans. Biomed. Eng.* **44** 555–66
- Corbishley P and Rodriguez-Villegas E 2007 Breathing detection: towards a miniaturized, wearable, battery-operated monitoring system *IEEE Trans. Biomed. Eng.* **55** 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** 657–9
- de Haan G and Jeanne V 2013 Robust pulse rate from chrominance-based rPPG *IEEE Trans. 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. Int. Med.* **8** 354–60
- Folke M, Cernerud L, Ekström M and Hök B 2003 Critical review of non-invasive respiratory monitoring in medical care *Med. Biol. Eng. Comput.* **41** 377–83
- Garde A, Karlen W, Dehkordi P, Ansermino J and Dumont G 2013 Empirical mode decomposition for respiratory and heart rate estimation from the photoplethysmogram *Computing in Cardiology (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** 4941–57
- Gravelyn T R and Weg J G 1980 Respiratory rate as an indicator of acute respiratory dysfunction *JAMA* **244** 1123–5
- 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**
- Iozza 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** 1934–44
- Iozza L, Lázaro J, Gil E, Cerina L, Mainardi L and Laguna P 2017 Respiratory rate detection using a camera as contactless sensor *Computing in Cardiology (Rennes, France)*
- Janssen R, Wang W and Moo A 2016 Video-based respiration monitoring with automatic region of interest detection *Physiol. Meas.* **37** 100–14
- Lázaro J, Gil E, Bailón R, Mincholé A and Laguna P 2013 Deriving respiration from photoplethysmographic pulse width *Med. Biol. Eng. Comput.* **51** 233–42
- 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 Inf.* **18** 240–6
- 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** 2317–33
- 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** 877–87
- Liao S, Jain A K and Li S Z 2016 A fast and accurate unconstrained face detector *IEEE Trans. Pattern Anal. Mach. Intell.* **38** 211–23
- Lucas B D and Kanade T 1991 Detection and tracking of point features *Carnegie Mellon Technical Report* MU-CS-91-132
- Massaroni C, Nicolò A, Lo Presti D, Sacchetti M, Silvestri S and Schena E 2019 Contact-based methods for measuring respiratory rate *Sensors* **19** 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** 334–43
- Meredith D J, Clifton D, Charlton P, Brooks J, Pugh C W and Tarassenko L 2012 Photoplethysmographic derivation of respiratory rate: a review of relevant physiology *J. Med. Eng. Technol.* **36** 1–7
- 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** N21–8
- 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** 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 *39th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Jeju, Korea)* pp 3818–21
- Pereira C B, Yu X, Blazek V, Venema B and Leonhardt S 2016 Multisensor data fusion for enhanced respiratory rate estimation in thermal videos *38th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Orlando, USA)* pp 1381–4
- 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** 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**
- 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 Trans. Pattern Anal. Mach. Intell.* **41** 121–35

- Regev N and Wulich D 2017 A simple, remote, video based breathing monitor *39th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (Jeju, Korea)*
- Rehoumaa H, Noumeira R, Bouachirb W, Jouvetc P and Essourica S 2018 3D imaging system for respiratory monitoring in pediatric intensive care environment *Comput. Med. Imaging Graph.* **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** 2760–7
- 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.* H1588–99
- Tarassenko L, Villarroel M, Guazzi A and Jorge J 2014 Non-contact video-based vital sign monitoring using ambient light and auto-regressive models *Physiol. Meas.* **39** 807–31
- 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** 172–5
- Verkrusse W, Svaasand L O and Nelson J S 2008 Remote plethysmographic imaging using ambient light *Opt. Exp.* **16** 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 Technol. Lett.* **1** 87–91
- Viola P and Jones M 2001 Rapid object detection using a boosted cascade of simple features *Proc IEEE Conf. on Computer Vision and Pattern Recognition (Kauai, USA)* pp 511–7
- Yasuma F and Hayano J I 2004 Respiratory sinus arrhythmia *Chest* **125** 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** e71384