Camera-Based Method for Respiratory Rhythm Extraction From a Lateral Perspective

MARC MATEU-MATEUS, FEDERICO GUEDE-FERNÁNDEZ, MIGUEL ÁNGEL GARCÍA-GONZÁLEZ, JUAN JOSÉ RAMOS-CASTRO, (Member, IEEE), AND MIREYA FERNÁNDEZ-CHIMENO, (Member, IEEE)

Department of Electronic Engineering, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain

Corresponding author: Marc Mateu-Mateus (marc.mateu.mateus@upc.edu)

This work was supported in part by the Ministerio de Economía, Industria y Competitividad (MINECO) under Project DEP2015-68538-C2-2-R, and in part by the Universitat Politècnica de Catalunya.

ABSTRACT

This work proposes a new method based on computer vision algorithms to measure the respiratory rhythm of a subject from a lateral perspective. The proposed algorithm consists on tracking the motion of the intercostal and abdominal muscles by the means of dense optical flow, being the novelty of the proposed method the extraction of the respiratory signal from the phase of the optical flow, while extracting at the same time a quality index from the modulus. 15 healthy subjects were measured while seated, and 4 tests were performed for each subject involving different scenarios. The algorithm has been validated using a commercial wearable thorax inductive plethysmograph system. The instantaneous frequency for the constant frequency respiratory tests, and the breath to breath analysis and instantaneous frequency of the free breathing test have been computed to assess the performance and error of the proposed method for the respiratory acquisition. Finally, a statistical analysis has been performed to assess the accuracy and performance of the quality index. The results of the study show a high agreement between methods in the 0.1 Hz and 0.3 Hz test. For the Free breathing test, both the cycle by cycle and Instantaneous frequency results show a low error between methods with high sensitivity in the cycle detection. The hypothesis that the modulus of the optical flow could be used as a quality index has been corroborated, with very good statistical results. Moreover, due to the simplicity of the proposed algorithm, the proposed method can perform in real-time while measuring respiratory rhythm and assessing the quality of the acquired signal. Further studies taking into account external vibrations have to be performed, to assert that the proposed method can be used in demanding conditions.

INDEX TERMS

Non-contact, respiration, camera-based, lateral perspective, computer vision.

I. INTRODUCTION

Respiratory signal extraction is now a key topic in our society, ranging from home monitoring to autonomous vehicles, from assessing respiratory rate to detecting drowsiness while driving. Although there are multiple methods to obtain the respiratory signal, ranging from pneumotachography [1] or hot-wire anemometers [2], to more advanced methods based on respiratory inductive plethysmography [3], [4], nowadays, methods based on contactless measurement have gained popularity as they are truly unobtrusive, thus allowing the acquisition of physiological variables without disturbing the subject in any form. As technology evolves, the feasibility of measuring physiological variables using such methods, mixed with the technological advancements in computing hardware, makes research in these area a hot topic. Some examples that can be found in the literature of these methods comprise: continuous wave doppler radar [5], [6], depth-based cameras [7]–[9] or even regular consumer-grade cameras [10]–[13] to acquire respiration.

From all the different available methods, those based on computer vision analysis have been gaining popularity in the literature, as the hardware required to perform the measurements is constantly getting cheaper and they do not interfere on other measurements in any way. There are multiple computer vision approaches on how to perform a measurement of the respiration of a subject, for example: [10] uses the head movements obtained through the averaging of the red
channel, [11] uses the motion of the upper torso and head obtained through an IR camera and [12] uses a combination of the RGB colours extracted from a ROI at the pit of the neck. Although all the aforementioned methods yield good results in terms of respiratory acquisition, there are some downsides to these methods. The method proposed in [10], if there is no compensation by the tracking algorithm of small head motions, the estimated respiration will be inaccurate. The algorithm proposed by [11], even though being more robust to small movements is only valid for subjects lying on a bed, and because the camera is placed above the subject any occlusion of the camera could lead to critical errors in the respiratory acquisition. On the other hand the algorithm proposed by [12], as it measures the respiration at the pit of the neck is more robust to small movements and occlusions, but as a downside it requires a manual ROI selection from the subject.

Apart from the aforementioned examples, and although there are plenty of studies in the literature that focus on detecting respiration through cameras, there are very few that measure breathing from a lateral perspective [13], [14]. For instance, [14] even though it measures breathing from a lateral perspective in adult population, the study is focused on obtaining the variation of wall volume during tidal breathing, hence no breath to breath variability analysis is performed. For this reason the study in [14] cannot be compared with the proposed method. At the best knowledge of these authors, studies in the literature that measure respiratory rhythm using this same measuring location in adult population, have not been found.

Measuring from a lateral perspective solves an issue that is not usually addressed in the studies involving video-based methods, which is the privacy of the subject. As the camera is no longer pointing at the face [10] or upper body [12] of the subject, the potential privacy issues of being recorded by a camera are greatly minimized as no facial region can be seen.

The aim of this work is to present a novel method to measure the respiratory rhythm on healthy subjects while sitting from a lateral perspective. The proposed method acquires the movement of the intercostal and abdominal muscles by the means of an RGB camera, pointed at the side of the thorax of the subject while being seated. By tracking the evolution of the aforementioned muscles by the means of the dense optical flow algorithm, the respiration of the subject can be retrieved.

To prevent errors in the respiratory rhythm extraction due to involuntary movements or occlusions, a method to evaluate the quality of the signal has also been presented. This method is based on using the modulus of the optical flow as an indicator of the quality of the signal, being good quality when the source of the signal is only respiratory activity, and bad quality when the obtained signal is embedded with errors from movements or noise.

In order to validate the proposed method, the obtained respiratory signal has been compared with a commercial wearable inductive plethysmography system (RespiBand system from BioSignalsPlux\textsuperscript{TM} [15]) that has been used as a reference system. Four tests have been performed to validate the method, two of them at a controlled respiratory frequency, another with free breathing and a final test designed for the quality indicator assessment obtained from the modulus of the optical flow. Both the video feed from the camera and the reference system were acquired simultaneously using the same computer. A synchronization between the respiratory signal obtained with the proposed method and the reference method was made upon processing the video data.

A possible field of application of the proposed method, would be monitoring the respiratory signal on those activities that require the subject to be seated, for example: at home, office or even while driving.

II. MATERIALS AND METHODS
A. RESPIRATORY MECHANICS
Breathing is a fundamental physiological function of the body, that involves many processes in order to let air in and out of the lungs in two very distinct phases: inspiration and expiration. During the inspiration, fresh air is inhaled into the lungs. This is achieved by the contraction of the muscles in the chest (the fibres of the diaphragm muscle shorten and the external intercostal muscles contract) which enlarges the thoracic cage, producing a piston like movement that pulls the lungs down, thus produces a negative pressure inside the lungs which favours the flow of air from the atmosphere to the lungs. On the other hand during expiration, air containing the CO\textsubscript{2} exchanged at the lungs is exhaled to the atmosphere due to the relaxation of the muscles in the chest (the diaphragm and the external intercostal muscles relax) returning to its initial state, thus releasing the captured air [16], [17].

Although the muscles in the chest are the main responsible for the expansion of the thoracic cage, some studies in the literature [14], [18] have demonstrated that the chest wall mechanics is not only confined to the chest motion but also to the abdomen as well. In [16], [18] it is shown that the chest wall has at least two degrees of freedom regarding the movement (expansion and contraction) during breathing, corresponding to the upper thorax (rib cage) and the abdomen. This is important, because the concept of multiple movements related to the respiratory effort, led [14] to demonstrate the viability of performing a measurement of lateral chest wall volume by observing the evolution of the chest and abdomen via optical markers.

B. PROPOSED ALGORITHM
The proposed algorithm takes advantage of optical flow techniques to measure the respiratory signal by tracking the movements of the intercostal muscles, the rib cage and abdominal region, from a lateral perspective while seated. The proposed algorithm was build solely based on the OpenCV (Version: 3.4) library. In order to process the video feed once it has been acquired, a custom code has been written in Python (Version 3.7 with Cpp bindings like cython).
Algorithm 1 Proposed Algorithm

1: procedure RESPIRATORY SIGNAL FROM LATERAL CAMERA
2:    frame ← getGreyFrame()
3:    (X, Y) ← calcOpticalFlow(frame, preFrame)
4:    (r, θ) ← cart2polar(X, Y)
5:    mod ← mean(r)
6:    ph ← mean(θ)
7:    quality_index ← calcSignSkew(mod)
8:    raw_respiration ← ph
9:    preFrame ← frame

The proposed algorithm (Alg. 1) is shown as pseudo-code to clarify how the different components interact with each other.

The proposed method has very few steps, with a very simple and yet effective architecture. A clear depiction of the involved steps is necessary to understand how the raw respiratory signal and the quality index are obtained, as they are respectively derived from the phase and the modulus of the optical flow. The simplicity of the proposed method allows to infer the possible bottlenecks in the performance and latency of the respiratory signal extraction. As the only algorithm that has a high computational cost is the optical flow computation, choosing an efficient implementation would translate into real-time performance of the whole method.

The following subsections present a more detailed explanation of the used optical flow algorithm, the respiratory signal obtention and the quality index computation.

1) OPTICAL FLOW

Once the frame is acquired from the camera, it is necessary to convert the frame to the grey-scale colour space, in case an RGB camera is used, as it reduces the computational cost. To compute the optical flow, two frames are needed, the current frame and the previous frame. If the algorithm does not have a previous frame, a copy of the current frame is used instead.

Dense Optical Flow algorithms compute the displacement (motion) between two consecutive images by tracking the image features in a pixel by pixel basis. The Optical Flow algorithm used in this work is the Dense Inverse Search (DIS) by Kroeger et. al [19] which is part of the OpenCV library.

The main characteristic and the first step of the DIS optical flow is the computation of the correspondences between features of two consecutive frames by the means of the inverse flow is the computation of the correspondences between features of two consecutive frames by the means of the inverse flow. By Kroeger et. al [19], the inverse Optical Flow (motion) between two consecutive images by tracking the image features in a pixel by pixel basis. The Optical Flow algorithm used in this work is the Dense Inverse Search (DIS) by Kroeger et. al [19] which is part of the OpenCV library.

The main characteristic and the first step of the DIS optical flow is the computation of the correspondences between features of two consecutive frames by the means of the inverse search method: for a given template patch \(T\) of size \(θ_{px} × θ_{px}\) centred at the location \(x = (x, y)^T\) inside the reference image \(I_r\), find the correspondence between \(I_r\) and the next frame \(I_{r+1}\) for the same patch \(T\) using the gradient descent algorithm. To perform the matching, the goal is to minimize the sum of squared distances (Eq. 1) over each patch to find a wrapping vector \(u = (u, v)^T\).

The optimization of the algorithm is performed using the inverse Lucas-Kanade algorithm [20], and as proposed in [20] and by the authors of the DIS algorithm [19], the inverse objective function (Eq. 2) can be optimized to estimate the update vector \(Δu\) around the current \(u\).

\[
\text{u} = \text{argmin}_u \sum_x [I_{r+1}(x + u') - T(x)]^2 \quad (1)
\]

\[
Δu = \text{argmin}_{Δu} \sum_x [T(x - Δu) - I_{r+1}(x + u)]^2 \quad (2)
\]

where \(T\) is a template patch, \(I_r\) is the reference image (previous frame), \(I_{r+1}\) is the new frame acquired from the camera, \(u = (u, v)^T\) is the wrapping vector representing the increment of the original \(x\) to the new location, and finally, \(Δu\) is the increment of \(u\) given \(u ← u + Δu\) that updates the wrapping from the previous frames.

The second step in the computation of the DIS optical flow defined in [19], is the “densification” of computed correspondences from the previous step. This step consists on building the dense flow field \(U_x\) (Eq. 3) for each pixel \(x\), by applying a weighted averaging to the displacement estimation of all patches from the previous step between the reference image (previous frame) and the new frame.

\[
d_i(x) = I_{r+1}(x + u_i) - T(x)
\]

\[
Z = \sum_i \max(1, ||d_i(x)||_2) \lambda_{i,x}
\]

\[
U_i(x) = \frac{1}{Z} \sum_i \frac{\lambda_{i,x}}{\max(1, ||d_i(x)||_2)} \cdot u_i \quad (3)
\]

where the \(λ_{i,x}\) is the indicator of the patch for a given location in the reference image, being \(λ_{i,x} = 1\) if the patch \(i\) has the same location as \(x\) in the reference image, \(||.||_2\) represents the euclidean norm, \(d_i(x)\) is the intensity difference at a given pixel between the wrapped image \(I_{r+1}\) and the template patch \(T\), \(u_i\) is the displacement for that given patch, and finally \(Z\) represents the normalization applied to compute the dense flow field.

The full mathematical definition of the DIS algorithm, by Kroeger et. al, and the performance optimisations undertaken to work with real-time constraints, can be found in [19].

2) SIGNAL EXTRACTION

Once the optical flow by the means of the DIS algorithm has been computed the algorithm returns two matrices, one for the \(x\) coordinates \((X)\) and one for the \(y\) coordinates \((Y)\) of the optical flow field, which contain respectively the \(x\) and \(y\) components of the new position of the features tracked by the dense optical flow. To obtain the information needed in order to acquire the respiratory signal and to compute the quality index, the phase and the modulus of the optical flow has to be computed.

a: RAW RESPIRATORY SIGNAL

In order to obtain the raw respiratory signal from the optical flow, the phase of the optical flow is used as it can be seen in the following equations:

\[
θ = \text{atan2}(X/Y) \quad (4)
\]
operation is to compute the modulus of the flow field, and the second is to normalize the resulting signal by computing the third moment estimator (Skewness) of a sliding window.

The modulus of the flow field (Eq. 6) is computed by the means of the euclidean distance ($l_2$ norm) taking into account all the vectors of the flow field. Once the modulus of each vector has been computed, an averaging of all the values is performed to obtain the sample that will conform the modulus signal ($S(k)$ on Eq. 7).

$$M = \sqrt{X^2 + Y^2} \quad (6)$$

$$S(k) = \frac{1}{N} \sum_{i=1}^{N} M_i \quad (7)$$

where $M$ is the matrix containing all the modulus from the vectors in the flow field, $X$ and $Y$ are matrix containing all the $x$ and $y$ coordinates of the flow field respectively. $S(k)$ is the modulus signal for the $k$th frame, and finally $N$ is the total number of points of the flow field.

Once the modulus has been obtained, there are three steps involved to compute the skewness and the final quality index: the sliding window conformation, where a sliding window of $n$ samples has been used, the skewness calculation (Eq. 8) and the recursive filtering (Eq. 9). An analysis of the influence of the sample length of the sliding window is performed in the results section to assess which number of samples yields better results.

$$\bar{s} = E[S] \approx \frac{1}{n} \sum_{i=1}^{n} s_i$$

$$m_3(k) = E \left[ \left( \frac{S - \mu}{\sigma} \right)^3 \right] \approx \frac{1}{n} \sum_{i=1}^{n} (s_i - \bar{s})^3$$

$$Q(k) = 0.9 Q(k - 1) + 0.1 m_3(k) \quad (8)$$

where $S$ is the sliding window containing the modulus signal, $\bar{s}$ is the mean of the sliding window, $n$ is the number of samples of the sliding window computed as the length of the window in seconds multiplied by the sampling frequency of the method, $m_3(k)$ is the $k$th sample of the skewness of the sliding window (raw quality index), and finally $Q(k)$ is the $k$th sample of the quality index where the initial $Q$ sample is defined as $Q(0) = 0$.

The first step involves the conformation of the sliding window from the raw modulus samples. A first in first out (FIFO) approach has been assumed as it is efficient and easier to implement. Following the classical FIFO window, whenever a new sample is obtained from the modulus, the oldest sample in the sliding window is discarded.

The second step consists on computing the skewness (3rd statistical moment) of the sliding window (Eq. 8). This method has been chosen as it quantifies the symmetry of the statistic inside the window. In Fig. 2 an example of three snapshots of a 30 second window can be appreciated. Fig. 2a corresponds to regular breathing, Fig. 2b corresponds to the transition between the end of the free breathing period and the start of reading out loud a text, and finally, Fig. 2c

\[ FIGURE 1. \] 120 px cropped region of the inspiration (a) and expiration (b) frames at the thorax with the produced optical flow (yellow lines). Images extracted using an example code of the OpenCV library.

\[ \text{where } \theta \text{ is the matrix of angles (phase) comprised between } (-\pi, \pi], \text{ X and Y are matrices containing all the } x \text{ and } y \text{ coordinates of the flow field respectively. } R(\theta) \text{ is the raw respiration sample of the } k \text{th frame, and finally } N \text{ is the total number of points in the flow field.} \]

The phase of the flow field is computed using the arc-tangent formula (Eq. 4). The atan2() function has been used instead of the classical implementation, as it takes into account the real quadrant of the phase, thus having a continuous range between ($-\pi, \pi$]. Once the phase of the flow field is computed for all the vectors, an averaging of all the values is performed (Eq. 5) to obtain the sample of the raw respiratory signal for that given frame.

Fig. 1 depicts the flow field vectors during inhalation (Fig. 1a) and exhalation (Fig. 1b). In both figures, the direction of the flow field can be appreciated (the cross represents the current position while the line depicts the previous positions): during inhalation, as the thorax rises, the flow field is oriented to the first quadrant, thus the angle of the flow field at the maximum inhalation will present values comprised between $[0, \frac{\pi}{2}]$, whereas during exhalation, when the thorax returns to its initial state, the flow field is re-oriented to the third quadrant and will present values at the peak exhalation comprised between ($-\pi, -\frac{\pi}{2}$]. This phenomena is the result of tracking the movements involved in the thoracic cage and abdomen while breathing, and by computing the changes in the phase of the flow field for each frame, and by concatenating the resultant averaged angles, the raw respiratory signal can be conformed.

\[ b: \text{QUALITY INDEX} \]

The quality index is based on the hypothesis that the modulus of the flow field yields higher amplitudes when the subject is moving or talking than when the subject is only breathing. The rationale behind this hypothesis is: as breathing is a slow frequency signal, the frame by frame movement will be small thus producing a small optical flow modulus. On the other hand, abrupt or sudden movements, like talking or performing an action, will report higher movement between frames hence yielding a higher modulus. To obtain the proposed quality index two different operations must be performed: the first
corresponds to a window where the subject has been reading for 30 seconds.

In Fig. 2a, as there is no movement because the subject is breathing regularly, the statistic is practically symmetrical thus the computed skewness will produce low values. As depicted in Fig. 2b, whenever a movement is registered and the increased modulus enters the sliding window, the statistic of the window becomes more asymmetrical and will produce an increase in the skewness. While the movement is taking place, as the window only comprises a few respiratory cycles, the increased modulus will maintain the asymmetry in the statistic (Fig. 2c). Finally, when the movement stops, the statistic will slowly become more symmetric and the skewness will decay to low or near zero values. This example is still valid even if the source of the movement is an occlusion or a sudden movement of the subject.

The third step consists on a recursive low-pass filter (Eq. 9), its function is to remove undesired components of the signal and to smooth its transitions. Multiplying factors of 0.9 for the previous sample and 0.1 for the current sample, have been chosen empirically based on preliminary results.

C. MEASUREMENT SETUP
To acquire the lateral image of the thorax, a Logitech™ C920 camera has been used. The Logitech C920 camera is a Universal Video Class (UVC) device, which enables a full control of the camera parameters such as exposition and focus, among others. Provided the configuration capabilities, the camera was configured at its native resolution of 640×480 px at 24 fps with YUY2 color-space, with fixed focus and exposure. The remaining parameters of the camera were set with their default values. As the camera was placed very close to the subject, it is important to establish the field of view of the camera, which is: 78°, 70.42°, 43.3° of diagonal, horizontal and vertical respectively. This field of view applies to all of the aspect ratios, but as the used resolution uses a 4:3 aspect ratio instead of a 16:9, the horizontal field of view was cropped to the matching ratio.

The reference system used to validate the algorithm was the RespiBand™ wearable inductive respiratory plethysmography (RIP) system from BioSignalsPlux [15]. The system is comprised of an inductive strap (chest band) and a Bluetooth transmitter. The respiration of the subject is detected by the means of the RIP signal and transmitted to the computer via a Bluetooth using a Serial Port Profile (SPP). The signal is sampled at 40 Hz and quantized with a resolution of 12 bit, the system also has a passband filter embedded into the analogue stage with cut-off frequencies of 0.058 Hz and 0.9 Hz.

Fig. 3a shows the position of the camera on the setup. The camera was mounted on a tripod and placed approximately at 7-9 cm from the thorax of the subject, between the 7th and the 10th rib. Given the FOV of the camera, the aspect-ratio and the approximate distances from the subject, the effective area of vision of the thorax is comprised between 41.17 cm² and 68.07 cm², which gives a ratio of 86.37 px/cm and 67.18 px/cm as it can be inferred from Fig. 3b. A table was placed in front of the setup to ease the conduct of the tests. Finally, the RespiBand system was placed on the thorax of the subject beneath the pectoral muscle.

To record all the data from the camera and the reference system simultaneously, a custom program based on ROS™ (Robotic Operative System) was used [21]. To save all the data from the cameras and the RIP sensor, the format *.bag has been used, as it allows to emulate the time when the frame was captured, thus enabling the synchronisation between systems without additional efforts and to check for data loss. The version of ROS used was the Melodic Morenia under Ubuntu 18.04 LTS. Finally, the laptop used to capture and process all the data had an Intel i7-4710HQ processor, with an Nvidia GeForce GTX 850M graphics card and 8 GB of RAM.

D. MEASUREMENT PROTOCOL
Fifteen healthy subjects, eight males and seven females volunteered for the study. Table 1 contains all the anthropometric data of the subjects including: age, height, width and body mass index (BMI).
The measurements were performed in a controlled environment, and each subject gave her/his written consent to participate in the study. This study was performed in accordance with the principles of the Declaration of Helsinki [22], and all the measurements performed complied with the regulations of the Universitat Politècnica de Catalunya (UPC).

Before any measurement, the subjects were instructed to put on the chest strap for the RIP system, to seat comfortably on a chair and to rest the arms comfortably in the table in front of the setup. The camera was then calibrated so the distance between the subject and the camera was approximately the same among measurements and subjects. Each subject was asked to perform four different tests with the aim to validate the proposed method against the reference method, using known respiratory frequencies, under unconstrained breathing, and finally, to validate the quality index obtained from the modulus of the optical flow.

To perform the respiratory signal validation in the frequency domain, two tests involving a controlled breathing at 0.1 Hz and 0.3 Hz were performed with a duration of 3 minutes each. To help the subjects breathe at the designated frequency, a custom visual aid was developed which consisted on a moving bar with 2/3 of the given period for inhaling and 1/3 for exhaling.

In order to validate the performance of the proposed method in free breathing conditions, another 10 minute test with no respiratory constraints (the subject could breath freely) was performed. To guarantee a natural respiratory frequency (without external interference), each subject was asked to breath normally while watching a documentary.

Finally, to validate the quality index, another 10 minute test was performed. The test consisted on three different tasks: free breathing, reading out loud a text and solving a sudoku puzzle. In this last test the subject was allowed to move freely while performing a sudoku, but had to remain silent and breath normally. To validate if reading out loud a text reported different results than solving a sudoku, a two minute period of free breathing between reading and solving the sudoku was established. The complete timetable of this test is depicted in Tab. 2.

At the beginning of each test, each subject was asked to perform an apnoea in order to align the signals from reference method and the proposed algorithm. Moreover, the aforementioned tests were performed sequentially, being the 0.1 Hz test the first and the unconstrained test the last, with a 3 minute interval between tests.

---

E. SIGNAL PROCESSING

Once the raw respiratory signals were obtained from the video by the means of the proposed algorithm, the normalisation, quality index and performance characterization were computed using Matlab (Version 2019b) and R (Version 3.6.0) software.

1) RESPIRATORY SIGNAL NORMALISATION

Prior to any comparison between the respiratory signal obtained with the proposed method and the reference method, the extraction of the respiratory cycles or the instantaneous frequency, a normalisation of both signals has been performed. The signal processing steps on both respiratory signals were the following:

- First, both respiratory signals were interpolated at 80 Hz using a cubic spline in order to normalize the sampling frequencies from both methods.
- A zero-phase 2nd order bandpass digital Butterworth filter with cut-off frequencies between 0.05 Hz and 0.6 Hz, was applied to both signals. In the case of the signal from the proposed method, this step is necessary to eliminate high frequency noise produced by the uncertainty in the optical flow computation. In the case of the reference method, it is used to remove the artefacts from the thoracic inductive band.
- Finally, a compression between −1 and 1 arbitrary units has been performed by applying to both signals a non linear function based on the arctangent function. Eq. 10 describes the formal definition of the applied transformation.

\[
S_n[k] = \arctan \left( \frac{S[k]}{\sqrt{\sum_{i=1}^{N}(S[i]-\bar{S})^2}} \right) \sqrt{2} \tag{10}
\]

where \( S[k] \) is the raw respiratory signal, \( \bar{S} \) is the mean of the raw respiratory signal, and finally \( S_n[k] \) is the normalized respiratory signal.

Fig. 4 depicts an example of the raw respiratory signal obtained by the proposed algorithm and its normalized counterpart.

---

### TABLE 1. Mean ± SD of the anthropometric data of the subjects.

| Age [years] | Height [cm] | Weight [kg] | BMI [kg/m²] |
|-------------|-------------|-------------|-------------|
| 26.6 ± 3.54 | 170.47 ± 9.26 | 64.6 ± 11.97 | 22.09 ± 2.72 |

### TABLE 2. Timetable of the unconstrained test.

| Duration | Activity                        |
|----------|---------------------------------|
| 3 min    | Free Respiration                 |
| 1 min    | Read a Text                      |
| 2 min    | Free Respiration                 |
| 4 min    | Solve a Sudoku                   |

---

VOLUME 8, 2020

---

FIGURE 4. Example between filtered and normalised respiratory signal obtained with the proposed algorithm. Signal extracted from the 0.3 Hz test.
2) RESPIRATORY INSTANTANEOUS FREQUENCY ESTIMATION

The IF of the respiratory signal from the reference method and the proposed algorithm has been obtained by means of the synchrosqueezing transform [23]. This method has been extensively used to extract the IF from ECG and respiration [24]–[26], or even to separate components from non-stationary signals [27].

To compute the IF from a signal, first a continuous wavelet transformation (CWT) (Eq. 11) has to be applied (a convolution between a wavelet and the signal). This step has to be repeated with an arbitrary number of wavelets of different period length (scales) in order to obtain the different harmonics within the signal.

To obtain the synchrosqueezing transformation from the previous CWT, the following mathematical definitions from Daubechies et al. [23] have to be taken into account. Taking as an example a source signal of the type: \( s(t) = A \cos(\omega t) \) the CWT can be rewritten as follows: given a wavelet \( \psi \), and following the mathematical approaches and definitions described in [23], if the wavelet \( \psi \) is concentrated only in the positive frequency axis, the following equations can be obtained.

\[
W_s(a, b) = \int s(t) \psi^\dagger \left( \frac{t-b}{a} \right) dt
\]

\[
W_s(a, b) = A \frac{1}{4\pi a^2} \frac{1}{b^w} \psi^\dagger(aw)e^{ibw}
\]

where \( a \) represents a certain scale (wavelet transformation), \( b \) represents a certain translation, \( W_s(a, b) \) is a CWT component for a given scale and shift, \( A \) represents the amplitude of the input signal, \( \psi \) represents the given wavelet, and finally \( \psi^\dagger \) represents the Fourier transform of the wavelet \( \psi \).

Taking into account the previous definition (Eq. 12), a candidate instantaneous frequency [23] can be computed with the following equation (Eq. 13) given that \( W_s(a, b) \neq 0 \).

\[
\omega(a, b) = -f(W_s(a, b))^{-1} \frac{\partial}{\partial b} W_s(a, b)
\]

where \( \omega(a, b) \) is the candidate IF, and \( W_s(a, b) \) represents the CWT for a given sample and scale.

Finally, to compute the synchrosqueezing transform, [23] the spectral power density of the different frequency components have to be concentrated into narrower bands, and taking into account that only a discrete number of scales “bins” have been computed \((a_k)\), this transformation will also be binned in \( a_k \) intervals (discrete frequency bins). The following equation (Eq. 14) defines this last step of the synchrosqueezing transformation.

\[
T_s(\omega_1, b) = (\Delta \omega^{-1}) \sum_{a_k|\omega(a_k, b) - \omega_1| \leq \Delta \omega/2} W_s(a_k, b) a_k^{-\frac{3}{2}} (\Delta a)_k
\]

where \( T_s(\omega_1, b) \) is the obtained synchrosqueezing transform, \( \Delta \omega \) is the distance between adjacent \( \omega \) that determine the discrete frequency bins, \( \Delta a \) represents the distance between adjacent scales, and finally, \( a_0 \) and \( a_k \) are respectively discrete frequency and scale components.

In order to obtain the desired frequency component from the whole synchrosqueezing transform (ridge), the component which presents the maximum power has been selected. To ensure that no “hops” between adjacent frequencies with similar power are produced, a penalty to reduce this frequency jumps has been put in place.

All the previous calculations have been performed using the tools provided on the Matlab environment. The functions used were: \( \text{wsst} \) and \( \text{wsstridge} \) with a penalty of 20 in the ridge computation.

3) RESPIRATORY CYCLE EXTRACTION

The respiratory cycle time series (RC) for both methods have been obtained following the same procedure that in [9]. The steps needed to compute the RC signal are the following:

- First, the percentile 65 is computed from the respiratory signal in order to obtain a threshold.
- This threshold is used to detect intersection with positive slopes in the respiratory signal.
- Finally, the time between consecutive slopes is computed to form the RC signal.

To conform the respiratory cycle time series, only the cycles that have been correctly detected in both methods have been taken into account.

F. PERFORMANCE CHARACTERISATION

To evaluate the performance of the proposed algorithm, three different methodologies have been proposed. For the 0.1 Hz and 0.3 Hz respiratory frequency test, two Bland-Altman (BA) plots of the mean IF and STD IF for each of the tests have been computed and evaluated. For the free breathing test, the respiratory cycles (RC) series and IF signal have been evaluated in terms of accuracy of detection of the cycles, error between series and BA analysis. Finally for the unconstrained test, different statistical methods have been used to assess the performance of the quality index. The following subsections describe how the different indicators have been obtained.

1) STATISTICAL PERFORMANCE

To determine the accuracy of the respiratory cycle detection on the proposed method for the free breathing test, a cycle to cycle comparison has been performed between the cycles obtained with the proposed method and the reference method using the same methodologies than in [9], which have been evaluated using the following indicators:

- True Positive (TP), number of respiratory cycles detected in both methods.
- False Positive (FP), number of respiratory cycles detected in the proposed method but not in the reference method.
- False Negative (FN), number of cycles detected in the reference method but not in the proposed method.
- Sensitivity (SEN), is the ratio between the TP and the total sum of TP and FN.
Positive predictive value (PPV), is the ratio between TP and the sum of TP and FP. The FP and FN are only used to obtain the SEN and PPV indicators, and only defined to clarify the computation of these two indices.

To quantify the error between RC series, the error has been defined as the TP from the proposed method minus the TP from the reference method. From the previous error definition, the following indicators have been evaluated: mean absolute error (MAE) and mean absolute percentage of the error (MAPE). A more detailed explanation of the cycle detection, the obtaining of the accuracy indicators and the error quantification can be found in [9].

The Spearman correlation has been obtained for the RC series and IF signal of the free breathing test, between the proposed method and the reference method for all the subjects on a sample by sample basis.

Finally, the Standard Deviation of the Error (SDE) has been obtained on a sample by sample basis for the constant frequency and free breathing test. On the constant frequency test the SDE has been computed for the IF signal, while for the free breathing test the SDE has been obtained either for the RC series and the IF signal.

2) ERROR ASSESSMENT

Four Bland-Altman (BA) plots [28] for the 0.1 Hz and 0.3 Hz test and three BA plots for the free breathing test have been performed. For the free test, in the cycle BA no distinction between subjects has been made as all the individual cycles from all the subjects have been used. For the instantaneous frequency BA plots, the Mean IF is defined as the mean instantaneous frequency for a given subject, and the STD IF is defined as the standard deviation of the instantaneous frequency for a given subject. In all the BA involving the IF, the median and the 95% interval (percentiles 2.5 and 97.5) have been used instead of the limits of agreement, as indicated by the Anderson-Darling Test (ADT) [29], because the samples did not present a normal distribution. The ADT was also applied to the standard deviation of the error (SDE) for both the IF for all the tests and the RC signal. The SDE of the IF for all the tests showed a $p < 0.05$ which discards the null hypothesis that the samples present a normal distribution. On the contrary the SDE of the cycles show a $p > 0.05$ which indicate that they adjust to a normal distribution.

In order to evaluate and validate the quality index, a series of statistical tests have been performed. Prior to any test, the quality index signal from each one of the subjects has been manually labelled and cropped into three different pieces: Free, Reading and Activity. The Free section includes the signal from the minute 0.5 to the minute 2.5, the Reading section is comprised between the minutes 2.5 to 4.5, and finally the Activity section is comprised between the minutes 5.5 until 10. Both the free breathing and reading have a length of 2 minutes while the activity piece has 4.5. The difference in lengths is justified as the unpredictable nature of the movements during activity requires a greater length to ensure that all the samples contain at least 2 to 3 movements. Once each piece has been correctly labelled, the mean amplitude of each piece has been obtained to perform the statistical tests. Prior to any test, an ADT test was performed to verify if the samples had normal distribution or not. The results of the ADT test showed a non-normal distribution for the free piece, for this reason non-parametric statistical tests have been used.

The tests used to perform the analysis of the quality index were the Kruskal-Wallis test [30] and the Nemenyi post-hoc test with a Tukey distribution for the post-hoc analysis, based on the PMCMR R package [31]. Where the pieces of the test containing free breathing, reading and activity have been compared between each other. Finally to ensure that the proposed index can determine correctly if the subjects are breathing normally or on the contrary are talking/reading or performing an activity, an Area Under the Curve (AUC) has been computed by the means of the perfcurve Matlab function, which internally uses a cross-validation algorithm to assess the true positive rate vs the false positive rate. To perform this analysis the free pieces have been labelled as “1” indicating a good quality of the signal, and the reading and activity pieces have been labelled as “0” indicating poor signal quality.

III. RESULTS

A. CONSTANT BREATHING TESTS

The results presented below refer to the first two breathing tests (constant rate), where the performance of the proposed algorithm regarding the instantaneous frequency estimation of the respiratory signal is evaluated.

1) SIGNALS

Fig. 5a depicts an example of the obtained respiratory signal and Fig. 5b the instantaneous frequency (IF) for the 0.1 Hz and 0.3 Hz tests. As it can be seen in Fig. 5b the IF of both methods is centred at 0.1 Hz and 0.3 Hz respectively, as the subject was instructed to breathe at those frequencies. In Fig. 5a a high concordance between the respiratory signal and the obtained IF from the reference method and the IF from the proposed method can be appreciated. For the 0.1 Hz test in Fig. 5a, the respiratory signal from the proposed method presents small ripples if compared to the signal obtained from the reference method. this can be caused by the subject not able to continue inhaling air due to the extremely slow respiratory frequency.

Table 3 contains the mean and standard deviation of the IF of the reference method and the proposed method for all subjects, expressed as median and interquartile range (IQR) [25; 75] as the samples do not present a normal distribution. This table serves as a reference to compare, for both methods, the base frequency of both the mean and standard deviation of the IF. It can also be noted that both the reference method and the proposed method present practically the same median for either the 0.1 Hz and 0.3 Hz test respectively.
M. Mateu-Mateus et al.: Camera-Based Method for Respiratory Rhythm Extraction From a Lateral Perspective

FIGURE 5. a) Respiratory signals obtained from the reference method and the proposed method for the 0.1 Hz and 0.3 Hz test respectively. b) IF obtained from the reference method and the proposed method for the 0.1 Hz and 0.3 Hz test respectively.

TABLE 3. Median and interquartile range [25; 75] of the Mean IF and STD If of the constant breathing tests, where RM stands for Reference Method, and PM stands for Proposed method.

| Test  | Mean IF RM [mHz] | STD IF RM [mHz] | Mean IF PM [mHz] | STD IF PM [mHz] |
|-------|------------------|------------------|------------------|------------------|
| 0.1 Hz| 99.79 [98.64; 99.93] | 0.92 [0.79; 1.02] | 99.71 [99.58; 99.92] | 1.16 [0.79; 1.54] |
| 0.3 Hz| 299.89 [299.76; 300.1] | 4.51 [3.86; 4.93] | 299.81 [299.69; 300.08] | 4.89 [4.03; 5.24] |

FIGURE 6. Bland-Altman plots of the instantaneous frequency for constant respiration tests, with rows and columns representing the 0.1 Hz, 0.3 Hz test, Mean and STD respectively.

2) PERFORMANCE

Table 4 contains the median and IQR [25; 75] of the SDE and Median and Percentile [2.5; 97.5] of the differences of Fig. 6 BA plots for the 0.1 Hz and 0.3 Hz tests.

| Test  | SDE [mHz] | BA Mean [mHz] | BA STD [mHz] |
|-------|-----------|---------------|--------------|
| 0.1 Hz| 1.02 [0.83; 1.12] | -0.08 [-0.37; 0.17] | 0.14 [-0.31; 0.73] |
| 0.3 Hz| 2.78 [2.71; 3.55] | -0.04 [-0.26; 0.2] | 0.37 [-1.49; 1.39] |

column contains the BA comparison for the standard deviation of the IF from the reference method versus the proposed method. Each point in the BA plots represents a subject, and no segregation between subjects has been made. In the Bland-Altman representation of the IF for the 0.1 Hz and 0.3 Hz test (Fig. 6), one outlier has been removed from each test. To remove these outliers the following criteria was used: if the difference between the reference method and the proposed method exceeded 1.5 times, below or above, the 95% reference interval, the sample was removed from the BA plot. This criterion was applied to remove the errors produced by the Synchrosqueezing transform when excessive jumps in the base frequency produced erroneous instantaneous frequency estimations.

Table 4 contains the median and IQR [25; 75] of the standard deviation of the error (SDE) IF between the reference method and the proposed method on a sample by sample basis, and the median and 95% reference interval of the BA plots in Fig. 6 for the 0.1 Hz and 0.3 Hz tests.

B. FREE BREATHING TEST

The following results are the ones regarding the free breathing test, where the performance of the respiratory cycle
extraction, as well as the estimation of the instantaneous frequency of the respiratory signal, are evaluated.

1) SIGNALS

Fig. 7 shows an example of the obtained respiratory signal, respiratory cycle time series and instantaneous frequency signal for both the proposed method (solid line) and the reference method (dashed line). On Fig. 7 it can be seen that the depicted respiratory signals in the first plot are on top of each other, indicating a high concordance between them. For the cycle time series in the second plot, both signals have the same temporal evolution with practically overlapping each other with only small differences in the number of periods. Finally, for the instantaneous frequency plot, it can also be appreciated that both signals follow the same amplitude changes and temporal evolution, indicating a high concordance between methods.

Table 5 presents the mean and standard deviation IF of the free breathing test for the reference method and the proposed method for all subjects, expressed as median and IQR [25; 75] as the samples do not present a normal distribution. It can be noted that the IQR of the mean IF increases if compared to the 0.1 Hz and 0.3 Hz tests as the respiratory signal was obtained in free breathing conditions so each individual breaths at his/her own pace. Moreover, the standard deviation of the IF is higher than in the previous tests because the subjects are not tasked with following a constant pace. The median for both methods is approximately 0.24 Hz which is inside the 14 to 20 breaths per minute range that corresponds to the typical respiratory frequency for adult population at rest [32].

2) PERFORMANCE

Table 6 summarizes the results of the cycle detection performance, the standard deviation of the error (SDE) and the correlation results either for the respiratory cycle series and the instantaneous frequency. The SEN and PPV indicators have been computed for each subject, and represent the accuracy of the system to detect respiratory cycles from the proposed method. For the SEN the median and interquartile range [25; 75] is presented instead of the mean ± std as the SEN for each subject did not present a normal distribution as a result of a very few false negatives. The mean ± std is presented for the PPV.
C. QUALITY INDEX

The next results refer to the unconstrained test, where the quality index is evaluated along with the relationship between the length of the sliding window and its statistical performance.

1) SIGNALS

Fig. 9 shows an example of the respiratory signals from the reference method and the proposed method as well as the quality index for the unconstrained test. The quality index is shown in the last plot of Fig. 9, where four different sections can be easily identified. The section from 0 to 3 minutes corresponds to a free breathing period, in the section from the minute 3 to minute 4 the subject was asked to read out loud a text, the section from the minute 4 to minute 6 corresponds to a second free breathing period, and finally from the minute 6 to the end of the signal the subject was asked to solve a sudoku puzzle (perform an activity with free movement).

In the first two plots, in the respiratory signal, when the subject is asked to read a text or perform an activity, both signals from the reference and the proposed method become less predictable and do not present any periodicity whatsoever. This is reflected in the quality index as an increase of the baseline. Another characteristic of Fig. 9, is in the transitions where the subject begins to read or begins to perform an activity, where it can be clearly seen that the quality index presents an abrupt baseline shift, indicating an instant degradation of the respiratory signal. As for the last plot of Fig. 9, increased values for the quality index report a degradation of the signal (while reading or performing an activity), while lower values indicate good signal quality (the periods where the subject is breathing normally).

2) PERFORMANCE

Table 8 shows the post-hoc results between the sections of the unconstrained breathing test where different sliding window lengths have been used. To perform this analysis the mean of the quality index signal of the whole piece (free, reading and activity) has been computed for each subject and aggregated by type. The result of the Kruskal-Wallis test with window lengths between 5 s to 30 s returned p-values inferior to $7.748 \cdot 10^{-5}$, being the maximum returned p-value for all these windows, which corresponds to the 30 second window. The 1 s window length returned a p-value of $3.899 \cdot 10^{-4}$.

### TABLE 7. Mean ± STD of cycle BA, median and 95% reference interval of the Mean IF and STD IF BA plots.

| Parameter     | Mean ± STD | Median | 95% Reference Interval |
|---------------|------------|--------|------------------------|
| Cycle [s]     | 0 ± 0.35   | 1.94 [0 ; 6.13] | 0.17 [-3.14 ; 5.11]  |

The MAE and MAPE are also presented in Table 6. Both indicators have been computed only taking into account the respiratory cycles of the given subject, hence an aggregated result in the form of mean ± std is given for both indicators. The SDE of the respiratory cycle series and the instantaneous frequency has been computed in a sample by sample basis for each subject. Finally, the Spearman correlation has been computed for both the RC series and IF signal between the reference method and the proposed method. The aggregated results for all the subjects are presented as a mean ± std for the SDE of the cycle time series and the correlation results, and as a median and interquartile range [25; 75] of the instantaneous frequency SDE, as this last one does not present a normal distribution.

Three Bland-Altman plots are presented in Fig. 8, the first BA plot compares the cycles obtained with the reference method and cycles obtained with the proposed method, where no distinction between subjects has been made and all the individual cycles have been used. A total of 2007 cycles from all the subjects have been used in this analysis. The second and third plots compares the mean and standard deviation of the instantaneous frequency of the reference method versus the proposed method for each subject. For the IF BA plots, each dot represents a subject and no segregation between subjects has been made. Two outliers have been removed from the IF BA plots following the same criteria as in the constant frequency tests. No outliers have been removed from the cycle BA plot.

Table 7 contains the differences of the BA plots in Fig. 8, where the cycle differences are expressed as mean ± std. The IF BA plots results, as they do not present a normal distribution, are expressed as median and 95% reference interval.
The post-hoc tests evaluate the statistical interaction between pieces of different kind: free versus reading, free versus activity and reading versus activity, with very significant differences \( p < 0.001 \) between free vs reading for all the sliding windows, significant differences \( p < 0.05 \) for the free vs activity, and no significant differences \( p > 0.05 \) between reading and activity. Regarding the AUC (where free breathing was labelled as “1” and reading/activity as “0”), all the window lengths present a value superior to 0.8.

Fig. 10 depicts a box plot of the samples present in each piece for all the subjects with a sliding window of 10 seconds. This particular length has been chosen as the default window length as it is the one that presents the lowest \( p \)-value between free versus activity, a good \( p \)-value in the free versus reading case and a very good AUC. As it can be seen, the disposition of each box, given a certain piece, is in concordance with the statistical results shown in Table 8.

**TABLE 8. Unconstrained test statistical results.**

| Window length [s] | Free vs Reading | Free vs Activity | Reading vs Activity | AUC     |
|-------------------|-----------------|------------------|---------------------|---------|
| 1                 | 0.0002          | 0.0607           | 0.2121              | 0.8     |
| 5                 | 1.6e-07         | 0.0046           | 0.0046              | 0.9578  |
| 10                | 2.1e-06         | 0.0011           | 0.3316              | 0.9533  |
| 20                | 2.1e-05         | 0.0015           | 0.5675              | 0.9244  |
| 30                | 0.0001          | 0.0025           | 0.7332              | 0.8956  |

Results for different sliding window lengths, \( p \)-values for the post-hoc statistical test between sections and area under the curve (AUC) of the true positive rate versus the false positive rate.

**FIGURE 10. Boxplot of the quality index for the free, reading and activity pieces using a 10 second window.**

![Boxplot of the quality index for the free, reading and activity pieces using a 10 second window.](image)

The constant frequency breathing tests, as it can be appreciated in Table 3, both tests for both methods present very similar median and IQR range for the mean IF, being the 0.3 Hz the test that presents a higher IQR range. For both methods and tests, the STD is comparable either in median and IQR range.

**IV. DISCUSSION**

**A. CONSTANT BREATHING TESTS**

As it can be appreciated in Table 3, both tests for both methods present very similar median and IQR range for the mean IF, being the 0.3 Hz the test that presents a higher IQR range. For both methods and tests, the STD is comparable either in median and IQR range.

Regarding the constant frequency breathing tests, in Fig. 5a the 0.1 Hz respiratory signal for the proposed method presents distortions in the signal due to forced breathing at slow frequencies, when the subject can no longer inhale air due to the slow respiratory frequency, and stops breathing to hold the air. While the subject is waiting for the visual aid to start its descend, the abdominal muscles bounce producing notches in the respiratory signal. This notches are also present in the...
signal from the reference method, but as the proposed method measures the changes at the intercostal muscles and abdomen this effect is more prominent.

This notches present in the 0.1 Hz test for the proposed method (Fig. 5a), produce harmonics in the instantaneous frequency. By using the synchrosqueezing transform, the fundamental harmonic at 0.1 Hz can be differentiated from the harmonics produced by the notches, due to the property of the synchrosqueezing to concentrate the spectral power density in narrow frequency bands. Using other methods to obtain the instantaneous frequency, for example the Hilbert transform, could lead to errors due to the aforementioned harmonics.

For the BA plots depicted in Fig. 6, no apparent bias can be appreciated and although the 0 value is inside the 95% reference interval, as the samples do not present a normal distribution, it can not be assured that the BA plots do not contain any bias. Either for the Mean IF and the STD IF plots, for both tests, a low median can be appreciated in Table 4, which can be interpreted as a good agreement between the reference method and the proposed method. Regarding the SDE for both tests in Table 4, values within the range of the STD results in Table 3 are reported indicating a low error between methods. The median of the SDE for both methods show very low values with a narrow IQR range, thus indicating a good precision of the proposed method to obtain the respiratory signal for the constant frequency respiratory tests.

**B. FREE BREATHING TEST**

Regarding the free breathing test, the results in Table 6 referring to the accuracy and performance of the cycle detection, report a median sensitivity of 100% with a very narrow IQR, indicating that most of the respiratory cycles have been detected correctly. A mean PPV of 94.7% can also be observed with low std values, indicating that from the detected cycles very few are false positives. This false positives can be due to small movements of the subject or differences between the thoracic and abdominal respiration [14], [18], as the reference method measuring point is beneath the pectoral muscle and the camera is measuring the respiration at the side of the thorax between the 7th and 10th rib. Taking into account both SEN and PPV in Table 6, these numbers indicate a high successful detection rate of the respiratory cycles.

For the MAE and MAPE results in Table 6, the MAE results indicate a low error between the cycles obtained with the proposed method and the reference method. The MAPE results, on the other hand, are consistent with the MAE results with a mean of 4.87% and std of 1.57% yielding low error results. Relative to the Cycle SDE low values can be appreciated in the results with mean of 0.34 s and std of 0.16 s. As for the correlation results for the RC series a correlation of 0.94 with a small std can be appreciated. All these results indicate a high agreement between the cycles obtained with the proposed method and the ones obtained with the reference method.

| MAE [bpm]   | MAPE [%]   | SDE Cycle [bpm] | Cycle BA [bpm] |
|-------------|------------|-----------------|----------------|
| 0.758 ± 0.278 | 5.353 ± 1.975 | 1.331 ± 0.64   | 0.082 ± 1.476  |

The instantaneous frequency results for the free test (Table 5) present similar median and IQR range either in the Mean IF and the STD IF between both methods. The IF SDE reports a median of 12.95 mHz which indicates a high agreement between both methods. Regarding the MAPE results of the IF signals, a mean of 0.9 with an std of 0.09 can be appreciated, being the increased std in this case due to the small discrepancies in the IF computation. In general, it can be seen that the correlation results are in agreement with the SDE results for the instantaneous frequency analysis.

Given the BA plots in Fig. 8, the Cycles BA presents a high agreement between methods with mean and std for the differences of 0 s and 0.35 s, with no apparent bias or trend. For the BA of the Mean IF, a bias of 2 mHz can be appreciated, this bias if compared to the mean IF for both methods yields less than 1% of the IF.

This bias can be due to small differences between the real sampling frequency and the theoretical sampling frequency of the camera. This small differences can produce misalignments between the signal obtained from the reference method and the signal obtained from the proposed method, thus producing a bias on the instantaneous frequency.

Regarding the STD IF BA plot, no apparent bias can be appreciated, although the 0 is included inside the 95% reference interval, as the samples do not present a Gaussian distribution, it cannot be assured that no bias is present inside the BA plot. The STD IF BA also yields low median near 0 with a narrow 95% reference interval.

To compare the results for the cycle respiratory signal with other methods in the literature, the results in this study have been adapted to the metric of breaths/min (bpm), which is commonly used in other studies. These results are shown in Table 9.

The obtained results of this study can be compared to the methods proposed by Massaroni et al., and the results obtained in [12] and [33]. These two works are based on acquiring the respiration at the pit of the neck by the means of an RGB camera. Although the methods in [12], [33] do not use the same measuring point as our method, and given that no previous methods could be found in the literature that use the same measuring region in adult population, they pose as a good reference of video-based methods for respiratory acquisition.

Comparing the obtained results with the ones shown in [33]: the mean MAE reported on [33] is 0.55 breaths/min with a max MAE of 1.23 breaths/min, our method although has a mean MAE of 0.758 it has less std than the one reported in [33]. Regarding the BA analysis [33] reports −0.03 ± 1.78 breaths/min, our method reports 0.082 ± 1.476 which presents more bias if compared

---

**TABLE 9. Mean ± STD of the MAE, MAPE, SDE and BA differences of the RC series expressed in breaths per minute (bpm).**
with [33] but with less STD. When comparing the results in [12], the overall MAE for the free breathing in [12] shows 0.39 breaths/min with BA results of near 0 mean and 1.02 breaths/min for the limits of agreement. Our method if compared with the latest study presents an increased MAE and increased BA results. These differences, although small, can be justified if it is taken into account the measurement point and the amount of cycles that have been used to perform the analysis. Overall the proposed method yields similar results than the ones reported in the literature, with small differences in MAE and the BA analysis.

**C. QUALITY INDEX**

Given the quality index in Fig. 9, in the last plot, a direct effect of the window length can be appreciated. Depending on the window length the quality index will have longer or shorter response time, where the signal will still predict a “bad quality” even though the signal is good. This is due to the window still having “bad samples” in it. Until all the samples from a “bad region” are removed from the window the quality index will not decrease. This effect can be clearly seen in Fig. 2 where three pieces of the signal with its histograms are depicted. Moreover, the statistical results of the analysis of the sliding window length versus the quality index (Table 8), indicate a clear dependence between the length of the window and the estimation of the quality of the signal. If the window length decreases, the statistical difference between pieces increases. This is true until the window has become sufficiently small (1 s window) that only a few periods of respiratory signal are included, then there is not enough information inside the window to accurately compute the quality index. For all these reasons, when choosing a suitable window length, the trade-off between specificity and response time has to be taken into account. A window of 10 s of length has been chosen as it is the one that reports a good trade-off between specificity, with an area under the curve of 0.9533 and a good statistical differences between the free breathing piece and the reading and activity pieces. Moreover, the response time of the used window is sufficiently small to not alter the results of the respiratory signal analysis.

Regarding Fig. 10 depicting the results of the 10 s window, it can be seen that the median and interquartile range (IQR) of the free distribution is below the reading and activity distributions, validating the statistical results in Table 8. As for the reading and activity, the values are very similar between each other with the difference that the reading piece has less IQR, which can be explained if it is taken into account that the reading was an action that every subject performed in the same way, while in solving a sudoku each subject had freedom of movement and proceeded differently.

**D. LIMITATIONS OF THE STUDY**

There were some limitations to this study. The first limitation was that the measurement of the reference method was performed at the chest while the camera measures the intercostal and abdominal region, which can produce different results depending on how the subject breaths. Although the obtained results do not show significant error neither in the IF nor the cycle detection, regarding this issue, a new validation with two reference methods between chest and abdomen has to be performed to completely discard errors due to abdominal movements. The second limitation to this study was that the narrow age gap between the subjects. To make the study extensible to all adult population, more specially to elder people where the respiratory mechanics can change, measurements with a wider age interval should be performed. The last limitation of the study was that the tests were not performed in changing lighting conditions or external vibrations. While the proposed algorithm will still be valid if the camera is substituted by an IR camera, which will mitigate the possible errors due to light changing conditions, a test with external vibrations should be performed. As external vibrations will have an impact in the obtained respiratory signal and quality index.

**V. CONCLUSION**

A new video-based method that tracks the thoracic-abdominal changes through the means of optical flow has been presented. The novelty of the method resides in obtaining the respiratory signal from the phase of the optical flow. Moreover, a quality index based on the modulus is presented. The algorithm has been validated using a thoracic inductive plethysmograph as a reference system.

The results of the constant frequency breathing test yield a high agreement between the proposed method and the reference method, demonstrating the viability of the method in retrieving the instantaneous frequency of the respiratory signal.

The free breathing test results show a high sensitivity in the detection of the respiratory cycles, with low error results in the cycle by cycle comparison. For the IF, the results show a good agreement between methods with small differences between the IF of the proposed method and the reference method.

The hypothesis that the modulus of the optical flow could be used as a quality indicator has been validated. The results for the quality index indicate that the algorithm can discern between normal breathing and non-normal breathing (reading/talking or movements). Moreover, the trade-off between specificity and the sliding window length for the computation of the quality index has been assessed.

Regarding the privacy of the subject, if compared with other video-based methods that use a frontal image (including the face of the subject), the proposed method does not pose any privacy issue as the camera is located at the side of the subject, and no recognizable part of the subject is shown. In addition, another advantage of the proposed location is that the possible occlusions or background errors are naturally minimized, and as the camera is located close to the subject there is no need for high pixel density to acquire the respiratory signal. As for the applicability of the method in real-life, one direct field of application would be on
automotive environments. The camera could be easily placed on the side of the driver, where the possible occlusions produced by steering on methods based on frontal cameras would be greatly mitigated. Moreover, by using the obtained quality index, any movement, occlusions or even talking could be immediately detected and assessed. Even though the method yields good results on healthy subjects, if the method were to be applied on subjects with respiratory diseases or respiratory alterations, further studies would be required to assess the different contributions of the respiratory mechanics to the proposed method.

In general, the proposed method can be used to extract the respiratory signal from healthy adult subjects at the side of the thorax, presenting a high agreement with the reference method either in cycle detection and instantaneous frequency. Thanks to its simplicity, the proposed algorithm performs with real-time constraints, extracting the respiratory signal from the phase of the optical flow while asserting the quality of the signal from the modulus.

REFERENCES

[1] A. Fleisch, “Der pneumotachograph; ein apparat zur geschwindigkeitsregi-
sterierung der atemluft,” Pflügers Arch., für die Gesamte Physiologie des
Menschen und der Tiere, vol. 209, no. 1, pp. 713–722, Dec. 1925.

[2] I. Yoshiha, Y. Shimada, and K. Tanaka, “Evaluation for clinical of a
hot-wire respiratory flowmeter.” J. Appl. Physiol., vol. 47, no. 5,
pp. 1131–1135, 1979.

[3] M.-N. Fiamma, Z. Samara, P. Bonnacier, T. Similowski, and
C. Straus, “Respiratory inductive plethysmography to assess respiratory
variability and complexity in humans.” Respiratory Physiol. Neurobiol.,
vol. 156, no. 2, pp. 234–239, May 2007. [Online]. Available:
https://linkinghub.elsevier.com/retrieve/pii/S156990480600334X

[4] Z. Zhang, J. Zheng, H. Wu, W. Wang, B. Wang, and H. Liu,
“Development of a respiratory inductive plethysmography module
supporting multiple sensors for wearable systems.” Sensors, vol.
12, no. 10, pp. 13167–13184, Sep. 2012. [Online]. Available:
http://www.mdpi.com/1424-8220/12/10/13167

[5] M. Mercuri, Y.-H. Liu, I. Lorato, T. Torfs, F. Wieringa, A. Bourdoux,
and C. Van Hooft, “A direct phase-tracking Doppler radar using wavelet
independent component analysis for non-contact respiratory and heart
rate monitoring.” IEEE Trans. Biomed. Circuits Syst., vol. 12, no. 3,
pp. 632–643, Jun. 2018.

[6] Y. He, C. Gu, H. Ma, J. Zhu, and G. V. Eleftheriades, “Miniaturized
circularly polarized Doppler radar for human vital sign detection.” IEEE
Trans. Antennas Propag., vol. 67, no. 11, pp. 7022–7030, Nov. 2019.

[7] H. Rehouna, R. Nourine, W. Bouachir, P. Jouvet, and S. Essouri,
“3D imaging system for respiratory monitoring in pediatric intensive care
environment.” Comput. Med. Imag. Graph., vol. 70, pp. 17–28, Dec. 2018,
doi: 10.1016/j.compmedimag.2018.09.006.

[8] J. Kemptke and K. Van Laerhoven, “Respiration rate estimation with depth
cameras,” in Proc. 5th Int. Workshop Sensor-Based Activity Recognit.
Interact., New York, NY, USA, Sep. 2018, pp. 1–10. [Online]. Available:
https://dl.acm.org/doi/10.1145/3266157.3266208

[9] M. Mateu-Mateus, F. Guede-Fernandez, M. A. Garcia-Gonzalez,
J. Ramos-Castro, and M. Fernandez-Chimeno, “Non-contact infrared-
depth camera-based method for respiratory rhythm measurement while
driving.” IEEE Access, vol. 7, pp. 152522–152532, 2019. [Online]. Available:
https://ieeexplore.ieee.org/document/8871160/

[10] F. Schrumpf, C. Monch, G. Bausch, and M. Fuchs, “Exploiting weak head
movements for camera-based respiration detection,” in Proc. 41st Ann.
Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Jul. 2019, pp. 6059–6062.

[11] K. Zhu, M. Li, S. Akbarian, M. Hafezi, A. Yadollahi, and B. Taati,
“Vision-based heart and respiratory rate monitoring during sleep—A valida-
tion study for the population at risk of sleep apnea.” IEEE J. Transl. Eng.
Health Med., vol. 7, Oct. 2019, Art. no. 1900708.
MARC MATEU-MATEUS received the degree in telecommunication engineering from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2015, where he is currently pursuing the Ph.D. degree in electronics. His research interest includes unobtrusive monitoring of physiological variables using video-based method. He is interested in deep learning and computer vision algorithms.

FEDERICO GUEDE-FERNÁNDEZ received the degree in telecommunication engineering and the Ph.D. degree from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2012 and 2018, respectively. His research interests include driver monitoring, mobile health, and machine learning.

MARC MATEU-MATEUS

MIGUEL ÁNGEL GARCÍA-GONZÁLEZ received the Ingeniero de Telecomunicación and Doctor Ingeniero Electrónico degrees from the Universitat Politècnica de Catalunya, Barcelona, Spain, in 1993 and 1998, respectively. He is currently an Assistant Professor of electronic engineering with the Universitat Politècnica de Catalunya. He teaches courses in several areas of medical and electronic instrumentation. He is involved in research on instrumentation methods and ECG, arterial blood pressure, and EMG measurements. His current research interests include time series signal processing by time-domain, frequency-domain, time-frequency spectra, nonlinear dynamic techniques, and noninvasive measurement of physiological signals.

JUAN JOSÉ RAMOS-CASTRO (Member, IEEE) received the degree in telecommunication engineering and the Ph.D. degree from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1992 and 1997, respectively. In 1992, he joined the Department of Electronic Engineering as a Lecturer. Since 1997, he has been an Associate Professor, teaching courses in several areas of electronic instrumentation. He is currently a Member of the Biomedical Research Center, UPC. His current research interests include biomedical and electronic instrumentation.

MIREYA FERNÁNDEZ-CHIMENO (Member, IEEE) received the Ingeniero de Telecomunicación and Doctor Ingeniero de Telecomunicación degrees from the Universitat Politècnica de Catalunya, Barcelona, Spain, in 1990 and 1996, respectively. She has been a Vice-Dean of the Telecommunication Engineering School (ETSETB) from 1996 to 2000. She is currently an Associate Professor of electronic engineering with the Universitat Politècnica de Catalunya. She is also a Quality Manager of the Electromagnetic Compatibility Group (GCEM), Technical University of Catalonia. GCEM is one of the centers of the Technological Innovation Network of Generalitat de Catalunya (Autonomous Govern of Catalonia). She teaches courses of electronic instrumentation, acquisition systems, and electrical safety. She is the coauthor of Electronic Circuits and Devices (Edicions UPC, 1999), and Automatic Test Systems (Edicions UPC, 1999) both published in Spanish or Catalan. Her current research interests include biopotential measurements (high-resolution ECG, beat-to-beat ECG monitoring, and heart rate variability) and electromagnetic compatibility, mainly oriented to medical devices and hospital environments.

MIREYA FERNÁNDEZ-CHIMENO