IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received September 20, 2019, accepted October 10, 2019, date of publication October 16, 2019, date of current version October 31, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2947759

Non-Contact Infrared-Depth Camera-Based Method for Respiratory Rhythm Measurement While Driving

MARC MATEU-MATEUS, FEDERICO GUEDE-FERNÁNDEZ, MIGUEL A. GARCÍA-GONZÁLEZ, JUAN RAMOS-CASTRO, (Member, IEEE), AND MIREYA FERNANDEZ-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 MINECO under Project DEP2015-68538-C2-2-R, and in part by the Universitat Politècnica de Catalunya.

ABSTRACT This work proposes a new non-contact method based on Infrared and Depth cameras to measure respiratory rhythm in real-life situations. The proposed algorithm consists on using both video feeds to track the movements of the subject in real-time, compute the location of the face and calculate the most suitable ROI to extract the respiratory signal in adult population. 20 subjects were measured while driving in a car simulator with no constraints other than the simulator itself. The algorithm has been validated using a commercial thorax plethysmography system. An opportunistic approach has been used to obtain pieces from the tests for each subject, thus making a more realistic approach to the real-life situations where the signal is likely to contain errors. The breath-to-breath respiratory signal and the instantaneous frequency from both methods has been computed from each piece to characterise the error between the proposed method and the reference system. The results show a high correlation between the measured rhythm from the reference method and the proposed method, with relatively low error results and good sensitivity in cycle detection and low errors for the instantaneous frequency between methods. The error results have also been compared with the ones obtained in previous studies showing a good agreement between the obtained results and the ones presented in the previous studies. No relationship between the length of the pieces and the error has been found either for the respiratory cycle signal or the instantaneous frequency signal. The proposed algorithm can be used to measure respiratory rhythm in unconstrained conditions and with opportunistic measurements, thus making it suitable to perform in real-life situations while driving. Further studies taking into account vibrations or light changing conditions are needed to confirm that the proposed method performs with the same accuracy with these constraints.

INDEX TERMS Non-contact, respiration, IR, camera-based, point-cloud, computer vision.

I. INTRODUCTION

In 2015 the European Sleep Research Society made a study across 19 european countries [1] in which they estimated that the prevalence of falling asleep while driving was 17 % of the population across Europe. Moreover, they estimated that as a direct consequence 7 % of the sleepy drivers would be involved in an accident.

Nowadays detecting drowsiness while driving has become a major topic. Examples of these algorithms and systems comprise algorithms to detect attention while driving [2], electrocardiogram (ECG) or electroencephalogram (EEG) acquisition and analysis systems [3], [4], or even respiratory rhythm detection and analysis methods [5], [6]. One relevant example in this last category can be found in [7], which is based on the respiratory rate variability analysis in order to detect the fight against to fall asleep.

Acquiring physiological variables through unobtrusive methods is also a major topic nowadays. Examples of these methods comprise from Doppler radar [8] to infrared thermography [9] for respiratory signal extraction. Recently, methods that rely on computer vision algorithms to extract

The associate editor coordinating the review of this manuscript and approving it for publication was Vishal Srivastava.

physiological signals can also be found, from consumer-grade cameras [10], [11] to more advanced depth point-cloud based acquisition [12].

Recent studies of breathing detection using depth point-cloud or Time-of-Flight sensors can be found in the literature: the study [13] uses a Microsoft Kinect V2 camera to acquire the breath signal in subjects lying down in a static position by manually selecting specific pixels. The study [14] also uses a Kinect camera but in this case the breathing is obtained by a predefined region of interest (ROI) to assess expiratory volume. More advanced algorithms can be found in [15] as they compute the ROI based on the pose of the subject, although these algorithms have better performance in real-life situations than the two aforementioned articles, [15] does not compare the obtained breathing with a reference system and they only use fixed breathing frequencies to verify the breathing signal. The study [16] presents an hybrid approach to obtain the ROI based on inferring the position of the bed and the subject using the Infrared feed of a Kinect camera, and then applying the ROI to the Depth stream. This last study uses the power spectral density (PSD) to compute the breathing rate, so no information about the breath to breath respiratory periods is obtained. All the cited studies use subjects either in a seating position or lying down on a bed in a static position, hence no movement errors are taken into account.

Moreover, most of the aforementioned unobtrusive camera-based methods and the previous mentioned studies are not suitable to measure physiological signals in conditions such as in a car cockpit, due to environmental variables comprising vibrations, sudden lighting changes and movements of the subject. Another limiting factor is that most of these algorithms do not work in real-time, which is a key factor to extract real-time information from the extracted physiological variables.

The aim of this work is to present a non-contact camera-based method [10] to measure respiratory rhythm (instantaneous frequency changes in the respiratory signal) with an hybrid architecture, using a synchronized Infrared (IR) and Depth point-cloud video feeds provided by the same camera. The proposed algorithm is able to track the movements of the subject using the IR feed, making it robust to sudden posture changes and more robust than an RGB feed to illumination changes. Using the same IR feed the face of the subject is located in order to compute the best region of interest (ROI) in the thoracic area to extract the respiratory signal from depth point-cloud feed. By tracking the changes in the thorax displacement of the subject the respiratory signal can be extracted. The algorithm is designed to perform in real-time as the extracted signal is intended to be used for drowsiness detection while driving. In order to validate the proposed method, the extracted respiratory signal has been compared with a commercial thorax plethysmography system based on an inductive band (RespiBand system from BioSignalsPlux ^{TM}) that has been used as a reference system. A car simulator has been used to perform all the tests in static conditions, where the only moving part was the subject being measured. Both reference system and the two video feeds from the camera were recorded simultaneously with the same computer and synchronized among them.

II. MATERIALS AND METHODS

The proposed method is based on the simultaneous acquisition of the infrared frame (IR) that the camera will use to extract the depth information and the depth point-cloud (Depth) itself. The camera uses two infrared feeds and a IR laser projector to compute the disparity map between the two IR feeds and then converts it to a depth point-cloud. The proposed algorithm was built based on the OpenCV (Version: 3.4) and Dlib [17] (Version: 19.16) libraries. The whole code for the proposed algorithm has been written in Python (Version 3.7 with CPP bindings like cython). The feeds from the camera were obtained using the librealsense library (version 1.12.1).

A. PROPOSED ALGORITHM

Fig. 1 summarizes the operational flow of the algorithm including the IR and Depth frame and the processing steps taken to compute the respiratory signal.

1) INFRARED FRAME

The infrared frame (IR) is used to perform two tasks, the first is to detect the face of the subject, whose location and width will be used to compute the ROI that will be used by the depth point-cloud (Depth), and the second is to detect if the subject is moving. In order for this algorithm to work, two frames are needed: the previous frame and the actual frame. The movement of the subject is computed in the following way by the means of optical flow algorithms.

- First, the IR frame is decimated by a factor of 4 by scaling the image to a 1/4 of its original size. This step is necessary to compute the optical flow as the whole frame would decrease the performance of the algorithm and increase the computational cost, and further scaling would decrease the accuracy of the optical flow.
- Then, the optical flow is computed by the means of the pyramidal implementation of the Kanade-Lucas-Tomasi (KLT) [18] algorithm by comparing the last frame from the camera and the actual frame.
- The mask obtained from the background removal is applied to the last step in order to remove possible false positives due to background movements.
- The mean modulus of the optical flow is computed, if the modulus does not exceed a certain threshold the last frame is updated, if otherwise, the algorithm proceeds to the face detection stage.

The face detection stage is only performed when the algorithm detects that the subject is moving. To perform this stage the original IR frame is used, without any decimation, to increase the accuracy of the face detection algorithm. To detect the face of the subject the Dlib library [19] has

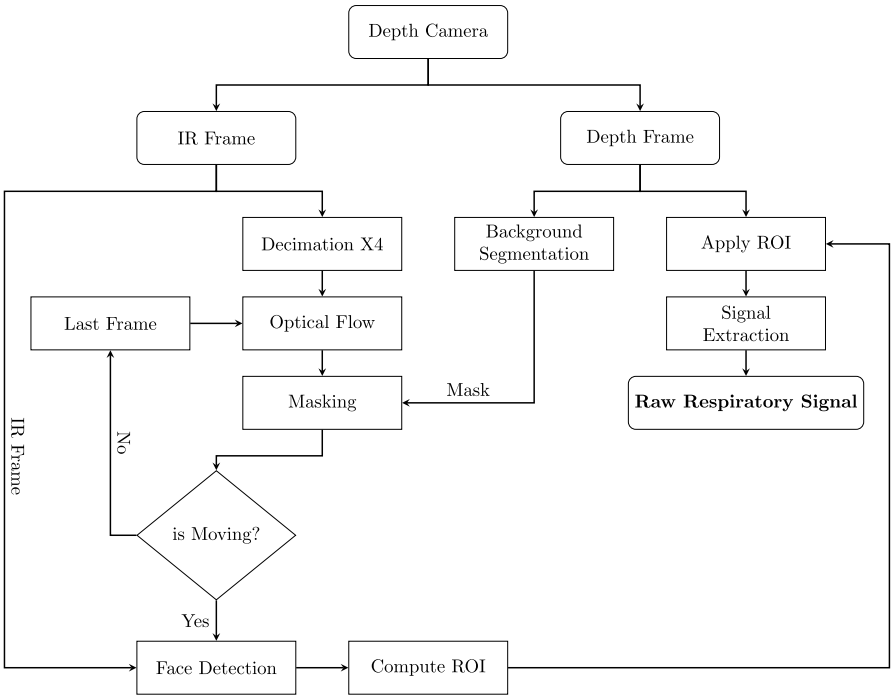


FIGURE 1. Flowchart of the proposed algorithm.

been used, this library uses the Histogram Oriented Gradients (HOG) feature combined with linear classifiers to detect and compute the position of human faces inside an image. Once a face is detected the ROI is computed as it can be seen in Equation 1 and depicted in Fig. 2a.

$$X_{roi} = X_i + w/2$$
$$Y_{roi} = Y_i + h * 2 \quad (1)$$

where (X_i, Y_i) are the left-corner coordinates of the rectangle that defines the position within the image of the face, w and h are respectively the width and the height of that given rectangle and finally, (X_{roi}, Y_{roi}) are the centre coordinates of the computed region. All the coordinates and the reference coordinates of the image can be appreciated in Fig. 2a.

The centre coordinates of the ROI are computed by taking into account the length of the rectangle that defines the face. The centre is located two times the height below the left-corner in the Y axis and half the width of the rectangle at the X axis as it can be seen in Fig. 2a. From that centre, the ROI is defined as a centred square of 32 by 32 pixels. This amount has been obtained by previous experiments as the one that yields better results, as a wider ROI produced slightly better results but required higher computational cost as well, and narrower ROIs produced poor quality results.

2) DEPTH POINT-CLOUD

The depth point-cloud (Depth) is used to perform two differentiated tasks: the first is to compute a background segmentation mask, and the second is to extract the respiratory signal given the ROI computed from the previous steps.

The background segmentation is performed by thresholding the point-cloud within the range of 0.5 m to 1 m. This distances have been chosen as they are the most probable range where the subject would be respectively from the camera, according to the camera position in the measurement setup. Once these thresholds have been applied to the point-cloud, a mask is generated by replacing the resultant points to the maximum value and the rest of the points to 0. This mask is the one used in the IR stage to filter the modulus of the optical flow.

Once the ROI is computed in the IR stage, the respiratory signal is extracted from the depth point-cloud by performing an average of all the points contained within the previous ROI. The computed ROI can be appreciated as the blue square depicted in Fig. 2b. The averaged distances are modulated by the movement of the thorax hence related to the respiration of the subject [20]. This final part of the algorithm is based on the patent Respiratory Signal Extraction [10] published in 2018 with number: WO/2018/121861 which has been validated in [21] versus other two video sources and a reference system for respiratory rhythm measurement.

B. MEASUREMENT SETUP

To acquire both the Infrared and Depth video feeds the RealSense ZR300 from Intel [22] was used. The camera was configured to acquire both feeds at 30 fps with a resolution of 492 x 372 pixels. This resolution was chosen as it is the highest resolution that supports distances as short as 0.5 m. The camera was placed at an approximate distance between

(a) (b)

FIGURE 2. Capture examples of the IR feed and Depth Point-Cloud from the preliminary tests (not the actual setup), with the metrics used to compute the ROI depicted in the IR feed and the final ROI represented as a blue square in the Point-Cloud. The ROI used to extract the respiratory signal (blue square) from the Depth Point-Cloud can be appreciated in (b).

0.6 - 0.8 m from the subject (this distance range was obtained after the analysis of the point-cloud for each subject).

The reference system used was the RespiBand plethysmographic system from BioSignalsPlux. This system is comprised of a Bluetooth transmitter and a thoracic band. The respiratory signal is acquired by sensing the volumetric changes of the thorax by the means of an inductive band. The chest wall strain is measured with a 12 bit ADC and sampled at 40 Hz, the signal is filtered with a 1st order analogue band-pass filter with frequencies 0.058 Hz and 0.9 Hz. The acquired signal is then sent to the computer by the means of a Bluetooth classic (2.0) serial port.

To record both video feeds and the reference system simultaneously, a custom program based on the ROS (Robotic Operative System) was used. Both signals where stored in a *.bag binary file with their respective timestamps so a posterior analysis could be performed. The ROS *.bag format allows to emulate the time when the events were recorded, using this emulation the algorithm has been tested for loss of information due to processing speed, hence testing it to perform in real-time environments. The version of ROS used was kinetic Kame on an Ubuntu 16.04 LTS.

The laptop used to acquire and process the measurements was an ASUS ROG gaming laptop with the following specifications: Intel i7- 4710HQ, Nvidia GeForce GTX 850M and 8 GB of RAM.

C. MEASUREMENT PROTOCOL

Twenty healthy subjects, 10 male and 10 female, with ages comprised between 23 and 49 years old (mean: 37.55 years, sd: 6.66 years), with height comprised between 148.5 *cm* and 195 *cm* (mean: 171.48 *cm*, sd: 10.38 *cm*), weight comprised between 49.5 *kg* and 114.5 *kg* (mean: 76.99 *kg*, sd: 18.02 *kg*) and body mass index (BMI) comprised between 17.9 kg/m^2 and 34 kg/m^2 (mean: 26.04 kg/m^2, sd: 4.87 kg/m^2) volunteered for the study.

The measurements were performed in a car simulator that consisted on: a full size car chassis (with steering wheel, pedals and car chair) and a highway simulator projected on a wall screen in front of the vehicle. The setup is depicted in Fig. 3.

(a) (b)

FIGURE 3. Photos of the car simulator in the facilities of the Biomechanics Institute of Valencia. (photos courtesy from the Biomechanics Institute of Valencia)

Each subject was asked to wear the RespiBand BioSignalsPlux system (reference system) and was asked to drive normally following the road indications, to behave as normal as they could and to act freely within the capabilities of the simulator. Before the beginning of the test, the subject was able to adjust the distance between the seat and the steering wheel. All the measurements, and the protocol to perform them, were dictated and performed by the Biomechanics Institute of Valencia. The authors did not partake in the realization of the tests, only the data obtained from the tests was processed by the authors.

D. SIGNAL PROCESSING

After the extraction of the respiratory signal by the means of the proposed algorithm, the normalisation, the piece selection and the performance analysis has been computed using Matlab version 2018a and for the statistical computation R version 3.6.0 has been used.

1) SIGNAL NORMALISATION

Prior to the cycle extraction from the respiratory signals obtained with the proposed algorithm and the reference system, and in order to perform its posterior comparison,

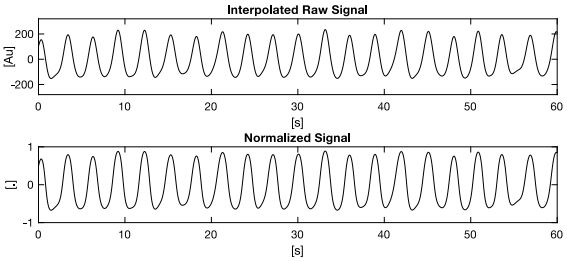


FIGURE 4. Comparison between raw signal and normalized signal from the Bioplux sensor.

the signals had to be normalized. The signal processing steps taken in both respiratory signals were the following:

- First, the respiratory signal extracted from the depth camera was interpolated at 40 Hz using a cubic spline. This step was necessary to normalize the sampling frequencies from both methods.
- A zero-phase 2nd order Butterworth digital bandpass filter was applied with cut-off frequencies between 0.05 Hz and 1 Hz. This filtering stage eliminates undesired components and removes possible drifts in the signal.
- A moving median filter [23] was applied to the signals in order to remove the peaks induced by the previous step, produced by the transitory periods that were a direct consequence of the change in the ROI due to the subject movements. The resulting signal is obtained by subtracting the median filter signal from the original signal. The window of the filter was set to three seconds, this length is enough to produce good results and shorter than a complete normal respiratory cycle.
- Finally, in order to compress the signal between -1 and 1, a non linear function was applied as defined in Equation 2 [24]

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

where $S[k]$ is the original respiratory signal, $\bar{S}$ is the mean of the respiratory signal and $S_n[k]$ is the normalized signal.

An example of the raw respiratory signal and the normalized signal for the reference system can be seen in Fig. 4.

2) OPPORTUNISTIC PIECE SELECTION

As the purpose of the proposed algorithm is to measure respiratory rhythm in real-life conditions, each test has been divided into variable-length pieces with a minimum duration of 60 s, as it will be the most probable scenario in real-life situations. Multiple pieces in an opportunistic manner have been taken for each user, each piece has been selected by visual inspection based on the amplitude of the signal from the proposed method. A piece is deemed valid if the signal amplitude is at least 0.5 and the length of this signal exceeds 60 seconds.

E. PERFORMANCE CHARACTERISATION

1) STATISTICAL PERFORMANCE

In order to characterise the relationship between the signals obtained with the proposed method and from the reference system, the respiratory cycle (RC) signal and the Instantaneous Frequency (IF) signal [21] has been computed for each one of the opportunistic pieces.

To avoid errors when comparing the reference system and the proposed method, the signals were aligned by the means of the intra-class Fisher correlation (ICC) [25], iterating around one period. As the maximum displacement between the reference system and the one from the proposed method is one period, using this length to slide the signals between each other ensures the ICC correlation is maximized, and that both signals are perfectly aligned between each other.

The steps needed to compute the RC series were the following:

- First, the percentile 65 was computed in both signals to obtain a threshold. This percentile was the one that yielded the best results.
- The previous threshold was used to detect the intersection with positive slopes in the respiratory signal.
- Finally, the time between consecutive slopes was computed to form the RC series.

The Instantaneous Frequency for both the proposed method and the reference system, was obtained following the same procedure that in [21] and defined in [26] as the first derivative of the instantaneous phase of the respiration signal:

- First, the Hilbert transform was obtained from both signals.
- The unwrapped phase of the Hilbert transform was obtained and corrected with increments of 2π to ensure continuity.
- The difference between adjacent angle samples was obtained conforming the instantaneous frequency in radians.
- A conversion between radians and Hz was performed.
- Finally, a Hodrick-Prescott [27] filter was applied to the resulting signal with a smoothing factor of $2 * 10^6$.

The Hodrick-Prescott filter is applied to the signal in order to remove the effects of impulsive noise produced by jumps of 2π radians in the argument of the Hilbert Transform.

2) DETECTION ASSESSMENT

In order to compute the cycle respiratory signal and to evaluate the performance of the cycle detector in the RC time series, a confusion matrix has been computed using the following parameters:

- True Positive (TP): number of respiratory cycles that have been detected in both the proposed method and the reference system.
- False Positive (FP): number of respiratory cycles that have been detected on the proposed method but not in the reference system.

- False Negative (FN): number of respiratory cycles that have been detected in the reference system but not in the proposed method.
- Sensitivity (SEN): ratio between TP and TP + FN.
- Positive predictive value (PPV): ratio between TP and TP + FP.

For further analysis, the FP and FN for both RC series have been discarded in order to avoid further errors, only the TP are taken into account to conform the RC series for both the proposed method and the reference system.

In order to asses the accuracy of the respiratory cycle (RC) detection on the proposed method, a cycle to cycle comparison between the proposed algorithm and the reference system has been performed and evaluated using the following methods: mean absolute error (MAE) (3) and mean absolute percentage of error (MAPE) (4). The standard deviation of the error (SDE) (5) has also been used for both RC and IF signals.

$$e_k[i] = S_k[i] - G_k[i]$$

$$MAE_k = \frac{1}{N}\sum_{i=1}^{N}|e_k[i]| \tag{3}$$

$$MAPE_k = \frac{\frac{1}{N}\sum_{i=1}^{N}|e_k[i]|}{\frac{1}{N}\sum_{i=1}^{N}G_k[i]} * 100 \tag{4}$$

$$\bar{e}_k = \frac{1}{N}\sum_{i=1}^{N}e_k[i]$$

$$SDE_k = \sqrt{\frac{\sum_{i=1}^{N}(e_k[i] - \bar{e}_k)^2}{N-1}} \tag{5}$$

In the previous expressions, $S[i]$ represents the RC series obtained from the proposed method, and $G[i]$ represents the RC series obtained from the reference system, N represents the total number of respiratory cycles for each piece and k represents the piece being analysed. No distinction between subjects has been made.

3) ERROR ASSESSMENT

Bland-Altman (BA) [28] plots comparing all the breath to breath cycles obtained from each method and the mean IF and the SD IF for each piece have been computed. No distinction between subjects or pieces has been made in the cycle BA, for the instantaneous frequency BA the samples have been grouped by pieces with no distinction between subjects. The mean IF is defined as the mean instantaneous frequency for a given piece, the SD IF is defined as the standard deviation of the instantaneous frequency for a given piece.

In both IF BA the median and the 95% reference interval (percentiles 2.5 and 97.5) has been used instead of the limits of agreement as the samples did not present a normal distribution as indicated by the Anderson-Darling Test (ADT) [29].

The ADT was also applied to the standard deviation of the error (for both cycles and instantaneous frequency) and the lengths of the pieces to verify if this data presented a Gaussian statistic or not, the tests from cycle SDE and the length of the pieces gave a $p > 0.05$ which indicate that they adjust to a normal distribution. On the contrary the test for the SDE of the instantaneous frequency gave a $p < 0.05$ which discards the null hypothesis that this data has a normal distribution.

In order to verify the hypothesis that there is no dependence between the length of the pieces and the SDE of the proposed method versus the reference system, an ANOVA test [30] has been performed taking into account the length of each piece and the mean SDE of the cycles for each piece. For the instantaneous frequency a Kruskal-Wallis test [31] has been performed taking into account the length and the mean IF SDE for each piece.

An ANOVA test between cycle SDE and PPV has also been performed to verify the relationship between the increase in False Positive detection and the increase in the cycle SDE.

III. RESULTS

A. SIGNALS

The number of pieces obtained from each subject was 6.53 ± 3.22 (mean ± sd) pieces, the length of each piece had a median of 167.5 s (percentile 25 %: 108 s, percentile 75 %: 271.38 s), the number of respiratory cycles per piece had a median of 47 cycles (percentile 25 %: 27 cycles, percentile 75 %: 77 cycles). The cycle length for piece had a median of 3.45 s (percentile 25 %: 2.98 s, percentile 75 %: 4.55 s) for the reference system and a median of 3.45 s (percentile 25 %: 2.98 s, percentile 75 %: 4.63 s) for the proposed method.

Fig. 5 has been obtained from one of the pieces from a given subject. An example of the comparison between the signal obtained from the proposed method and the reference system is depicted, along with the computed Respiratory Cycle series and the Instantaneous Frequency series for both methods.

In Fig. 5 it can also be observed that both respiratory signals from both methods are on top of each other, indicating a high concordance between them. On the respiratory cycle plot it can be seen that both signals follow the same temporal evolution with practically the same changes with small differences in the number of computed respiratory cycles. For the instantaneous frequency plot, it can be observed that both signals behave in the same way with the same temporal and amplitude evolution.

If the interval between 40 s and 60 s is taken as an example for both the respiratory cycle and the instantaneous frequency, when the RC signal increases (the respiratory cycle increase its length) an equivalent decrease can be seen in the IF signal (the instantaneous frequency decreases), the same behaviour can be appreciated if the IF signal increases.

B. PERFORMANCE

Table 1 summarizes the results for the confusion matrix, sensitivity (SEN) and the positive predictive value (PPV). The TP, FP and FN presented in Table 1 refer to the aggregated amount of TP, FP and FN without any distinction between

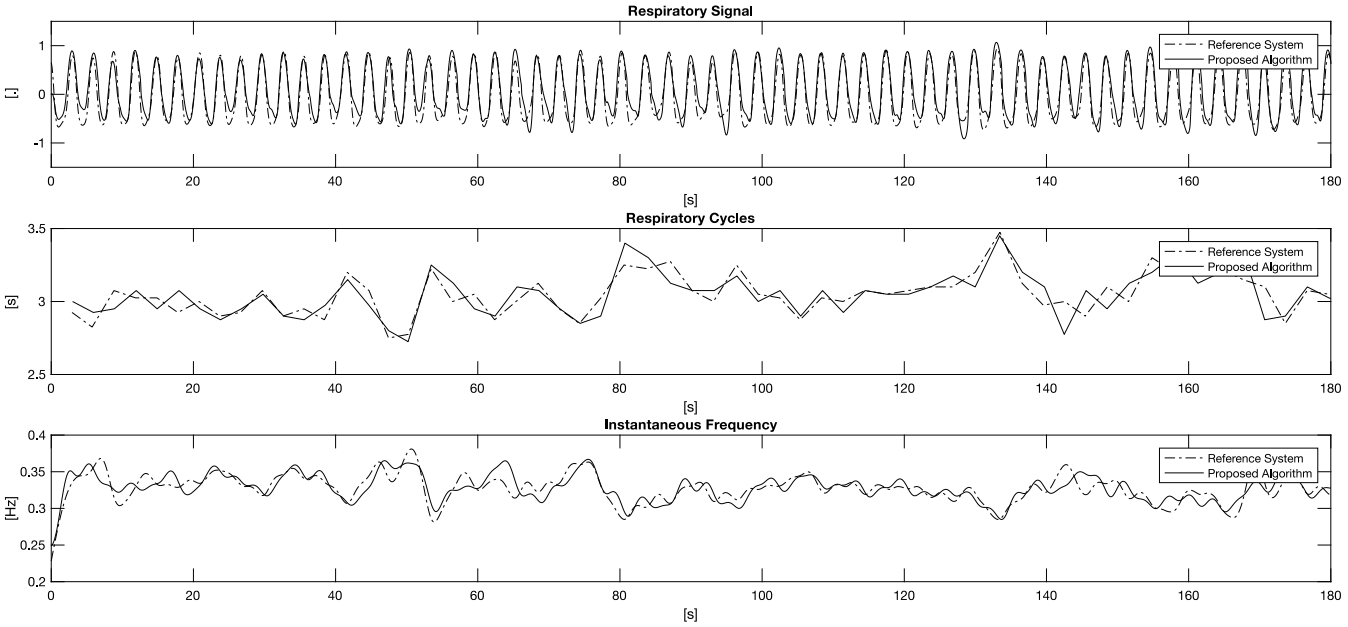


FIGURE 5. Comparison between the signals obtained from the reference system (Dashed) and the Proposed Method: raw respiratory signal, respiratory cycles and instantaneous frequency.

TABLE 1. Confusion elements, sensitivity and PPV.

TP	FP	FN	G. SEN (%)	G. PPV (%)	SEN (%)	PPV (%)
5796	1387	1711	77.21	80.69	80.12 ± 14.23	80.74 ± 13.40

Where G. SEN and G. PPV represent the Global values for these indicators.

TABLE 2. Breath to breath ICC, MAE and MAPE expressed as mean ± standard deviation.

Cycle ICC	MAE [s]	MAPE (%)
0.96	0.46 ± 0.25	8.49 ± 2.99

subjects or pieces. SEN (Global) and PPV (Global) have been computed using the aggregated number of TP, FP and FN making no distinction between subjects or pieces. SEN and PPV have been computed for each piece with the TP, FP and FN of each piece and the mean and standard deviation have been presented without making distinction between subjects.

Results for the mean correlation between the respiratory cycles for both methods and all pieces, the standard deviation of the error, the mean absolute error and the mean percentage error can be found in Table 2. The intra-class Fisher correlation (ICC) between the respiratory cycle signals has been obtained from the true positive (TP) RC series from both the proposed method and the reference system.

Both the MAE and MAPE results were obtained from the true positive respiratory cycle series. For the MAPE results in Table 2, as the MAPE is influenced by the length of the sample, the weighted mean and standard deviation has been used to avoid errors due to the difference in length between pieces.

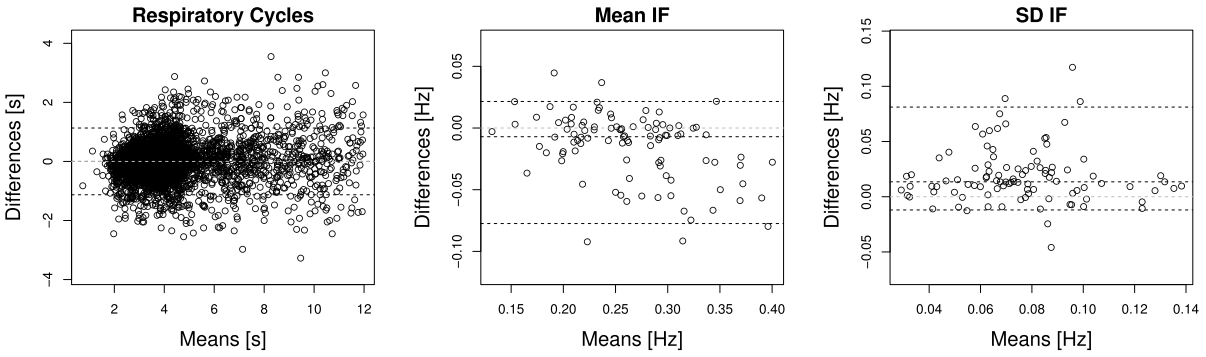
1) ERROR CHARACTERISATION

For the Bland-Altman representation, some outliers have been removed based on the following criteria: for the respiratory cycles, the cycles that exceeded more than 2 times the standard deviation of the group were eliminated. For the instantaneous frequency, if the difference for a piece was below the 2.5 % or above the 97.5 % percentile of the error, these pieces were discarded. A total of 100 cycles were removed and a total of 4 pieces were removed from both IF BA.

Fig. 6 contains three Bland-Altman plots, the first plot (Fig. 6a) compares the respiratory cycles for the reference system versus the respiratory cycles for the proposed method, where no distinction between subjects or pieces has been made. The second and third Bland-Altman plots (Figures 6b, 6c) compare respectively: the mean instantaneous frequency of the reference system versus the mean IF of the proposed method for each piece, and the standard deviation of the IF for the reference system versus the standard deviation of the IF for the proposed method for each piece. In these last BA plots no distinction between subjects has been made.

Table 3 contains the mean and standard deviation for the BA differences of the respiratory cycles and the median and 95% reference interval for the BA differences of both mean IF and the SD IF.

Table 4 summarizes the results for the standard deviation of the error for both respiratory cycles and the instantaneous frequency. In Table 4 the results for the statistical tests



(a) (b) (c)

FIGURE 6. **Bland-Altman of the computed periods, mean instantaneous frequency and standard deviation of the instantaneous frequency.**

TABLE 3. **Mean ± SD of the differences of Fig. 6a. Median and percentile [2.5% ; 97.5%] of the differences of Fig. 6b and Fig. 6c.**

Cycles [ms]	Mean IF [mHz]	SD IF [mHz]
1.013 ± 576.559	-7.05 [-77.43 ; 21.52]	13.5 [-11.98 ; 81.2]

TABLE 4. **Mean ± SD of cycle SDE, Median and interquartile range [2.5% ; 97.5%] of the IF SDE.**

SDE Cycles [s]	SDE IF[mHz]	Significance
0.62 ± 0.32	90.6 [28.27 ; 187.8]	NS_* ; $NS_\#$; †

The significance results for the ANOVA [30] and Kruskal-Wallis tests [31] are represented as follows: "$*$" for the results of the length vs the cycle SDE, "$\#$" for the length vs the IF SDE and "†" for the cycle SDE vs PPV. Regarding the significance: "NS_*" for $p > 0.05$, "$*$" for $p < 0.05$ and "$**$" for $p < 0.001$.

comparing the mean SDE and the length for each piece for both methods can also be seen, the results for these tests present non-significant differences between length and SDE. On the other hand, the results comparing cycle SDE with PPV do present significant differences.

IV. DISCUSSIONS

Regarding the accuracy of the cycle detection for both methods shown in Table 1 a global sensitivity of 77.21 % and a global PPV of 80.69 % can be appreciated. The results for the mean and standard deviation for both SEN and PPV taking into account each piece individually are slightly better than the global ones. SEN and PPV did not yield high results, but it must be taken in consideration that the driver did not have any restriction regarding the movements inside the cockpit of the simulator, thus producing abrupt shifts in the baseline of the optical method making each piece slightly different even within the same subject. After the normalisation, although most of the errors attributed to movements are corrected, as the cycle detection for each piece has been performed using a percentile 65 of the piece, the errors that remain in the signal have a very negative impact on the cycle detection which decreases the whole accuracy for the global values.

Another important factor to take into account to discuss the SEN and PPV results is the high variability between pieces for the same subject, which have a clear impact in the error. This behaviour can be clearly appreciated when taking into account the significance results between PPV and cycle SDE for each piece, which shows a $p < 0.05$. This result can be explained as: when the PPV decreases the cycle SDE increases and vice-versa due to detection of true positives (TP) and false positives (FP) in each piece. If a piece contains less TP and more FP, the piece will have less PPV and an increased cycle SDE than a piece with the same length that has more TP and less FP. This type of behaviour is potentiated by the freedom of movement given to the subjects and the errors on the detection due to these movements.

Taking into account the results in Table 2, a high correlation between the RC series from the reference system and the one from the proposed method can be appreciated, although this result is high, the correlation alone does not give information of how close the agreement is between both RC series. To get further into the agreement between the respiratory cycles obtained with the proposed method and the ones obtained with the reference system, the SDE, MAE and MAPE results should be taken into consideration.

As for the MAE results, these values indicate a relative low error between respiratory cycles obtained with the proposed method and the reference system. Moreover, if the error is compared to the mean duration of a period; the MAE result is roughly a 10 % of the median length of a cycle, thus relative low mean and SD values can be found. The MAPE results on the other hand are consistent with the MAE results as a weighted mean of 8.49 % with a weighted standard deviation of 2.99 % can be observed. To interpret these results, the constraints of the tests must be taken into account as the subject had no movement restrictions whatsoever during the test, hence these results can be interpreted as good results if the contour variables like the movement of the subject are taken into account.

In order to compare both methods taking into account the movements of the subject, and the reliability of the cycle

detection in these constraints, the Instantaneous Frequency of both signals has been computed. Using the IF instead of the cycle detection allows for the comparison of both methods without the need to use piece-related information, for example the percentile 60, to extract the IF.

Given the Bland-Altman results in Fig. 6, the plot from Fig. 6a presents a high agreement between methods with mean and standard deviation for the differences of: 10.13 ms $\pm$ 576.56 ms with no apparent bias error. For Figures 6b and 6c, although by visual inspection no bias can be appreciated and although the 0 value is contained inside the 95 % reference interval, it can not be assured that there is no bias present in the BA plot; as the samples do not present a normal distribution. Fig. 6b presents a low median value for the differences which can be interpreted as a good agreement between the proposed method and the reference system. For the results in Fig. 6c a low median can also be appreciated.

Taking into account the Reading results for the Depth method on [21], the Reading test is the one that presents more errors due to the nature of the movements while reading hence comparable to the involuntary movements that are produced while driving. If the results for this two last BA plots are compared with the Reading results on [21], it can be appreciated an increased error (differences) in both mean and sd than the results in the original study. This can be explained as in the original study the only variable was the subject reading out loud a text in an static posture, while in this study the subject had total freedom of movement within the limits and capabilities of the simulator.

Relative to the cycle SDE, relative low values can be appreciated in the results: mean of 0.62 s, and with an SD of 0.32 s. These results indicate a high agreement between the computed cycles of the reference system and the ones obtained from the proposed method. As for the SDE of the IF signal low values can also be appreciated, if the 95 % reference interval for the IF SDE is compared with the inter quantile range (IQR) obtained for the SDE Reading test on [21], it can be appreciated that the results are similar as the IQR from the Reading test is contained in the 95 % reference interval. These results can be compared as the Reading test from [21] is the test with less constraints of the study, moreover, as for the nature of the breathing involved, this test contains errors due to involuntary movements and due to the unique variations of the respiratory signal. These movements and unique respiratory patterns are also reflected in the current study as the subject was not asked to perform any specific action while performing the test, hence the error from the current test can be compared to the one from the Reading test in [21].

Regarding the statistical tests, both the ANOVA and the Kruskal-Wallis tests for the length vs the cycle SDE and length vs the IF SDE respectively, show non-significant differences. This can be interpreted as there is no relationship between the length of the piece and the error that is produced in either the respiratory cycle series or the instantaneous frequency series. These results have special relevance as it confirms that the use of variable-length pieces has no impact on the error between the proposed method and the reference system. Moreover, these results confirm that it is viable to use variable-length pieces in real-life systems to compare two alternative methods that measure respiratory rhythm. For the ANOVA between the PPV and the cycle SDE, significant differences can be found, which shows a relationship between the PPV and the cycle SDE. This can be explained as if the error between the reference system and the proposed method for the RC series increases, less cycle TP will be found hence a decrease in the PPV. If the SDE decreases more cycle TP can be found in both the reference system and the proposed method, hence the PPV will increase.

There were several limitations to this study, being the first the number of subjects that participated on the study. Only 17 of the 20 subjects could be used. On the three subjects that were discarded, the point-cloud feed was corrupted beyond recovery due to the distance between the camera and the subject being less than the minimum distance required by the camera.

The second limitation of this study was that although all the tests were performed in a simulator with no constraints to the subjects, the tests were not performed under real vibration or real light changing conditions. For these reasons it cannot be assured that the algorithm would perform identically in vibration or light changing conditions, as they will certainly have a negative impact in the respiratory signal extraction.

The last limitation of this study was the lack of an algorithm to automatically assess the quality of the signal. Having this algorithm would have provided an objective tool to either automatically select the length of each piece, or to provide a better guidance to select the pieces by visual inspection. Because in this study the pieces were obtained by visual inspection, some errors were introduced and some outliers had to be removed from the results. Once the aforementioned algorithm is in place there will be no need to remove any outlier. As a future work, this signal quality algorithm will be developed and added to the current method.

V. CONCLUSION

A new hybrid non-contact IR-Depth method to acquire respiratory signals has been presented. The proposed algorithm consists on detecting the movements of the subject by the means of optical flow algorithms on an IR image, and computing the most suitable ROI that it can later be used by the Depth camera to acquire the respiratory signal. The algorithm has been validated using a thorax plethysmography system as a reference system. The results showed a high correlation between the acquired respiratory cycles from the proposed method and the reference system, with relatively low error results and good sensitivity in the cycle detection. Regarding the IF results, no bias can be appreciated with low IF SDE indicating a good agreement between the proposed method and the reference system. The statistical results for the tests involving the cycle SDE and the IF SDE versus the length of the pieces show no correlation between each other, validating

the hypothesis that the respiratory signal can be chunked in an opportunistic manner based on the quality of the signal, without having any impact on its analysis and characterisation. In general, the proposed algorithm was able to detect the respiratory signal for the purpose of measuring respiratory rhythm (for both the cycles and IF), in real-time and real-life situations with high performance if compared to the reference system.

ACKNOWLEDGMENT

The authors would like to specially thank: FICOSA SA and the Instituto de Biomecánica de Valencia (IBV).

REFERENCES

[1] M. Gonçalves, R. Amici, R. Lucas, T. Åkerstedt, F. Cirignotta, J. Horne, D. Léger, W. T. Mcnicholas, M. Partinen, J. Téran-Santos, P. Peigneux, and L. Grote, "Sleepiness at the wheel across Europe: A survey of 19 countries," *J. Sleep Res.*, vol. 24, no. 3, pp. 242–253, 2015.

[2] B. Mandal, L. Li, G. S. Wang, and J. Lin, "Towards detection of bus driver fatigue based on robust visual analysis of eye state," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 3, pp. 545–557, Mar. 2017.

[3] C.-S. Wei, Y.-T. Wang, C.-T. Lin, and T.-P. Jung, "Toward drowsiness detection using non-hair-bearing EEG-based brain-computer interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 400–406, Feb. 2018. [Online]. Available: http://ieeexplore.ieee.org/document/8247228/

[4] K. T. Chui, K. F. Tsang, H. R. Chi, B. W. K. Ling, and C. K. Wu, "An accurate ECG-based transportation safety drowsiness detection scheme," *IEEE Trans. Ind. Informat.*, vol. 12, no. 4, pp. 1438–1452, Aug. 2016.

[5] N. Rodríguez-Ibáñez, M. A. García-González, M. Fernández-Chimeno, and J. Ramos-Castro, "Drowsiness detection by thoracic effort signal analysis in real driving environments," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug./Sep. 2011, pp. 6055–6058. [Online]. Available: http://ieeexplore.ieee.org/document/6091496/

[6] N. R. Ibanez, M. F. Chimeno, J. J. R. Castro, M. A. G. Gonzalez, E. M. Masip, and D. B. Matinez, "Method and system for determining an individual's state of attention," U.S. Patent 8 734 359 B2, May 27, 2014.

[7] F. Guede-Fernández, M. Fernández-Chimeno, J. Ramos-Castro, and M. A. García-González, "Driver drowsiness detection based on respiratory signal analysis," *IEEE Access*, vol. 7, pp. 81826–81838, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8744224/

[8] Y. S. Lee, P. N. Pathirana, R. J. Evans, and C. L. Steinfort, "Noncontact detection and analysis of respiratory function using microwave Doppler radar," *J. Sensors*, vol. 2015, Feb. 2015, Art. no. 548136. [Online]. Available: http://www.hindawi.com/journals/js/2015/548136/

[9] C. B. Pereira, X. Yu, M. Czaplik, R. Rossaint, V. Blazek, and S. Leonhardt, "Remote monitoring of breathing dynamics using infrared thermography," *Biomed. Opt. Express*, vol. 6, no. 11, p. 4378, Nov. 2015. [Online]. Available: https://www.osapublishing.org/abstract.cfm?URI=boe-6-11-4378

[10] M. F. Chimeno, J. R. Castro, M. A. G. Gonzalez, F. G. Fernandez, M. M. Mateus, N. R. Ibañez, B. B. Pujols, and J. M. A. Gomez, "Respiratory signal extraction," WO Patent 2018 121 861, Jul. 5, 2018. [Online]. Available: https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2018121861

[11] C. Massaroni, D. L. Presti, D. Formica, S. Silvestri, and E. Schena, "Non-contact monitoring of breathing pattern and respiratory rate via RGB signal measurement," *Sensors*, vol. 19, no. 12, p. 2758, 2019.

[12] A. Procházka, M. Schätz, O. Vyšata, and M. Vališ, "Microsoft Kinect visual and depth sensors for breathing and heart rate analysis," *Sensors*, vol. 16, no. 12, p. 996, Jun. 2016. [Online]. Available: http://www.mdpi.com/1424-8220/16/7/996

[13] E. Silverstein and M. Snyder, "Comparative analysis of respiratory motion tracking using Microsoft Kinect v2 sensor," *J. Appl. Clin. Med. Phys.*, vol. 19, no. 3, pp. 193–204, 2018.

[14] C. Sharp, V. Soleimani, S. Hannuna, M. Camplani, D. Damen, J. Viner, M. Mirmehdi, and J. W. Dodd, "Toward respiratory assessment using depth measurements from a time-of-flight sensor," *Frontiers Physiol.*, vol. 8, p. 65, Feb. 2017.

[15] J. Kempfle and K. Van Laerhoven, "Respiration rate estimation with depth cameras: An evaluation of parameters," in *Proc. 5th Int. Workshop Sensor-Based Activity Recognit. Interact.*, 2018, pp. 1–10.

[16] M. Martinez and R. Stiefelhagen, "Breathing rate monitoring during sleep from a depth camera under real-life conditions," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2017, pp. 1168–1176.

[17] D. E. King, "Dlib-ml: A machine learning toolkit," *J. Mach. Learn. Res.*, vol. 10, pp. 1755–1758, Jan. 2009.

[18] J.-Y. Bouguet. (2001). *Pyramidal Implementation of the Affine Lucas Kanade Feature Tracker—Description of the Algorithm*. [Online]. Available: http://robots.stanford.edu/cs223b04/algo_tracking.pdf

[19] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 2010, pp. 886–893. [Online]. Available: http://ieeexplore.ieee.org/document/1467360/

[20] S. Kumagai, R. Uemura, T. Ishibashi, S. Nakabayashi, N. Arai, T. Kobayashi, and J. Kotoku, "Markerless respiratory motion tracking using single depth camera," *Open J. Med. Imag.*, vol. 6, no. 1, pp. 20–31, 2016. [Online]. Available: http://www.scirp.org/journal/doi.aspx?DOI=10.4236/ojmi.2016.61003

[21] M. Mateu-Mateus, F. Guede-Fernández, V. Ferrer-Mileo, M. A. García-González, J. Ramos-Castro, and M. Fernández-Chimeno, "Comparison of video-based methods for respiration rhythm measurement," *Biomed. Signal Process. Control*, vol. 51, pp. 138–147, May 2019. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S1746809419300400

[22] Intel. (2017). *Intel RealSense 3D Camera*. [Online]. Available: https://click.intel.com/media/ZR300-Product-Datasheet-Public-002.pdf

[23] G. R. Arce, *Nonlinear Signal Processing: A Statistical Approach*. Hoboken, NJ, USA: Wiley, 2004. [Online]. Available: https://www.wiley.com/en-es/Nonlinear+Signal+Processing:+A+Statistical+Approach-p-9780471676249

[24] A. Martínez, R. Alcaraz, and J. J. Rieta, "Application of the phasor transform for automatic delineation of single-lead ECG fiducial points," *Physiol. Meas.*, vol. 31, no. 11, pp. 1467–1485, Nov. 2010.

[25] R. A. Fisher, *Statistical Methods for Research Workers*, vol. 5, F. E. Crew and D. W. R. Culter, Eds., 6th ed. London, U.K.: Oliver & Boyd, 1936.

[26] B. Boashash, "Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals," *Proc. IEEE*, vol. 80, no. 4, pp. 520–538, Apr. 1992. [Online]. Available: http://ieeexplore.ieee.org/document/135376/

[27] R. J. Hodrick and E. C. Prescott, "Postwar U.S. Business cycles: An empirical investigation," *J. Money, Credit Banking*, vol. 29, no. 1, pp. 1–16, Feb. 1997. [Online]. Available: http://www.jstor.org/stable/2953682?origin=crossref

[28] D. Giavarina, "Understanding bland Altman analysis," *Biochem. Med.*, vol. 25, pp. 141–151, Sep. 2015.

[29] T. W. Anderson and D. A. Darling, "Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes," *Ann. Math. Statist.*, vol. 23, no. 2, pp. 193–212, 1952.

[30] R. V. Hogg and J. Ledolter, *Engineering Statistics*. New York, NY, USA: Macmillan, 1987.

[31] W. H. Kruskal and W. A. Wallis, "Use of ranks in one-criterion variance analysis," *J. Amer. Stat. Assoc.*, vol. 47, no. 260, pp. 583–621, 1952.

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 current research interests include unobtrusive monitoring of physiological variables using video-based method, 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 current research interests include driver monitoring, mobile health, and machine learning.

MIGUEL A. GARCÍA-GONZÁLEZ received the Ingeniero de Telecomunicación degree and the Doctor Ingeniero Electrónico degree from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1993 and 1998, respectively. He is currently an Assistant Professor of electronic engineering, UPC. He teaches courses in several areas of medical and electronic instrumentation. He is engaged in the 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, and nonlinear dynamic techniques, and noninvasive measurement of physiological signals.

JUAN RAMOS-CASTRO 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 Electronic Engineering Department as a Lecturer and he has been an Associate Professor, since 1997, 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 FERNANDEZ-CHIMENO (M'90) received the Ingeniero de Telecomunicación and Doctor Ingeniero de Telecomunicación degrees from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 1990 and 1996, respectively. She has been a Vice-Dean of the Telecomunication Engineering School (ETSETB), from 1996 to 2000. She is currently an Associate Professor in electronic engineering with the UPC. She is also a Quality Manager of the Electromagnetic Compatibility Group (GCEM)—one of the centers of the Technological Innovation Network of Generalitat de Catalunya (autonomical govern of Catalonia), Technical University 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.