UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament d'Enginyeria Electrònica

A contribution to unobtrusive video-based measurement of respiratory signals.

Thesis submitted in partial fulfilment of the requirement for the PhD Degree issued by the Universitat Politècnica de Catalunya, in its Electronic Engineering Program.

Doctoral thesis by:
Marc Mateu-Mateus

Thesis advisors:
Juan José Ramos-Castro
Mireya Fernández-Chimeno

Barcelona, October 2020

Acknowledgements

En primer lloc m'agradaria agrair als meus directors Juan Ramos i Mireya Fernández per la seva dedicació, paciència i confiança, sense les quals no hagués estat possible realitzar aquesta tesi. Sobretot per donar-me llibertat per explorar i provar noves idees, molts cops a priori poc versemblants, les quals han contribuït a la qualitat d'aquesta tesi. També agrair al Miguel Ángel Gárcia tot el suport i ajuda al llarg d'aquests anys, i al Alfonso Méndez per la infinitat de converses i la inestimable ajuda rebuda. I tants altres professors/res del departament que m'han ajudat tots aquests anys.

A tots els meus companys del laboratori: Sergio, Antonio, Albert... però en especial al Federico Guede i al Victor Ferrer. Sense la paciència del Federico i la seva incansable ajuda, aquesta tesi no hagués estat la mateixa. Victor, crec que sense tu continuaria sense entendre l'estadística, i en especial que és un ANOVA. A la resta de companys, dir que els nostres debats i jocs de taula, han estat una part essencial d'aquesta etapa de la meva vida.

Agrair al Xavi Marti tot el suport i l'amistat que m'ha brindat aquests anys i per l'oportunitat de sortir de la meva zona de confort, fet que ha contribuït a millorar la qualitat d'aquesta tesi.

Al Miquel Llull, per la seva inestimable amistat i les nostres converses culinari-tecnològiques, les quals m'han brindat l'oportunitat de plantejar i debatre idees (algunes més viables que d'altres), què han contribuït notablement al progrés personal durant la realització d'aquesta tesi.

Vull fer una menció al CMU Penyafort-Montserrat i en especial a la seva directora a la Montse Lavado, per donar-me l'oportunitat de poder viure i gaudir del col·legi major aquests anys.

Agrair a tots els amics/gues i companys/es que han estat al meu costat en aquesta etapa de la vida, i agrair especialment a la Eli el seu suport.

Al Ramon Granés per tots aquests anys en els que m'ha ajudat i ensenyat.

Finalment, agrair als meus pares Isabel i Josep, al meu germà Josep, als meus tiets Mª Rosa, Josep i Mercè i a tota la resta de família pel seu continu suport en aquesta etapa de la meva vida.

Abstract

Due to the growing popularity of video-based methods for physiological signal measurement, and taking into account the technological advancements of these type of devices, this work proposes a series of new novel methods to obtain the respiratory signal from a distance, based on video analysis. This thesis aims to improve the state of the art video methods for respiratory measurement, more specifically, by presenting methods that can be used to obtain respiratory variability or perform respiratory rhythm measurements. Moreover, this thesis also aims to present a new implementation of a time-frequency signal processing technique, to improve its computational efficiency when applied to the respiratory signals.

In this document a first approach to video-based methods for respiratory signal measurement is performed, to assert the feasibility of using a consumer-grade camera, not only to measure the mean respiratory rate or frequency, but to assert if this hardware could be used to acquire the raw respiratory signal and the respiratory rhythm as well. In this regard a new video-based method was introduced that measures the respiratory signal of a subject at a distance, with the aid of a custom pattern placed on the thorax of the subject.

Given the results from the first video-based method, a more broad approach was taken by comparing three different types of video hardware, with the aim to characterise if they could be used for respiratory signal acquisition and respiratory variability measurements. The comparative analysis was performed in terms of instantaneous frequency, as it allowed to characterise the methods in terms of respiratory variability and to compare them in the same terms with the reference method.

Subsequently, and due to the previous obtained results, a new method was proposed using a stereo depth camera with the aim to tackle the limitations of the previous study. The proposed method uses an hybrid architecture were the synchronized infrared frame and depth point-cloud from the same camera are acquired. The infrared frame is used to detect the movements of the subject inside the scene, and to recompute on demand a region of interest to obtain the respiratory signal from the depth point-cloud. Furthermore, in this study an opportunistic approach is taken in order to process all the obtained data, as it is also the aim of this study to verify if using a more realistic approach to respiratory signal analysis in real-life conditions, would influence the respiratory rhythm measurement.

Even though the depth camera method proved reliable in terms of respiratory rhythm measurement, the opportunistic approach relied on visual inspection of the obtained respiratory signal to properly define each piece. For this reason, a quality indicator had to be proposed that could objectively identify whenever a respiratory signal contained errors. Furthermore, from the idea to characterise the movements of a subject, and by changing the measuring point from a frontal to a lateral perspective to avoid most of the occlusions, a new method based on obtaining the movement of the thoraco-abdominal region using dense optical flow was proposed. This method makes us of the phase of the optical flow to obtain the respiratory signal of the subject, while using the modulus to compute a quality index.

Finally, regarding the different signal processing methods used in this thesis to obtain the instantaneous frequency, there were none that could perform in real-time, making the analysis of the respiratory variability not possible in real-life systems where the signals have to be processed in a sample by sample basis. For this reason, as a final chapter a new implementation of the synchrosqueezing transform for time-frequency analysis in real-time is proposed, with the aim to provide a new tool for non-contact methods to obtain the variability of the respiratory signal in real-time.

Resum

A causa de la creixent popularitat en la mesura de senyals fisiològics amb mètodes de vídeo, i tenint en compte els avenços tecnològics d'aquests dispositius, aquesta tesi proposa una sèrie de nous mètodes per tal d'obtenir la respiració a distància mitjançant l'anàlisi de vídeo. Aquesta tesi té com a objectiu millorar l'estat de l'art referent a la mesura de senyal respiratòria mitjançant els mètodes que en ella es descriuen, així com presentar mètodes que puguin ser usats per obtenir la variabilitat o el ritme respiratori. A més, aquesta tesi té com a objectiu presentar una nova implementació d'un mètode de processat de senyal temps-freqüencial, per tal de millorar-ne l'eficiència computacional quant s'aplica a senyals respiratoris.

En aquest document, es realitza una primera aproximació a la mesura de senyal respiratòria mitjançant mètodes de vídeo per tal de verificar si és factible utilitzar una càmera de consum, no només per mesurar el senyal respiratori, sinó verificar si aquest tipus de hardware també pot ser emprat per obtenir el ritme respiratori. En aquest sentit, es presenta en aquest document un nou mètode d'adquisició de senyal respiratòria a distància basat en vídeo, el qual fa ús d'un patró ubicat al tòrax del subjecte per tal d'obtenir-ne la respiració.

Un cop obtinguts els resultats del primers resultats, s'han analitzat tres tipus diferents de càmeres, amb la finalitat de caracteritzar-ne la viabilitat d'obtenir el senyal respiratori i la seva variabilitat. L'estudi comparatiu s'ha realitzat en termes de freqüència instantània, donat que permet caracteritzar els mètodes en termes de variabilitat respiratòria i comparar-los, en les mateixes condicions, amb el mètode de referencia.

A continuació, s'ha presentat un nou mètode basat en una càmera de profunditat estèreo amb la finalitat de millorar i corregir les limitacions anteriors. El nou mètode proposat es basa en una arquitectura hibrida la qual utilitza els canals de vídeo infraroig i de profunditat de forma sincronitzada. El canal infraroig s'utilitza per detectar els moviments del subjecte dins l'escena i calcular, sota demanda, una regió d'interès que s'utilitza posteriorment en el canal de profunditat per extreure el senyal respiratori. A més a més, en aquest estudi s'ha utilitzat una aproximació oportunista en el processat del senyal respiratori, donat que també és un dels objectius d'aquest estudi, verificar si el fet d'utilitzar una aproximació més realista en l'adquisició de senyal, pot influir en la mesura del ritme respiratori.

Tot i que el mètode anterior es mostra fiable en termes de mesura del ritme respiratori, la selecció oportunista del senyal necessita d'inspecció visual per tal de definir correctament cada fragment. Per aquest motiu, era necessari definir un índex de qualitat el qual permetés identificar

de forma objectiva cada tram de senyal, així com detectar si el senyal conté errors. Partint de la idea de caracteritzar el moviment del subjecte de l'estudi anterior, i modificant el punt de mesura frontal cap a un de lateral per tal d'evitar oclusions, es proposa un nou mètode basat en l'obtenció del moviment toràcic-abdominal a partir del flux òptic del senyal de vídeo. Aquest mètode recupera el senyal respiratori del subjecte a partir de la fase del flux òptic, tot calculant un índex de qualitat a partir del mòdul.

Finalment, i tenint en compte els diferents mètodes de processat utilitzats en aquesta tesi per tal de obtenir la freqüència instantània, es pot apreciar que cap d'ells és capaç de funcionar en temps real, fent inviable l'anàlisi de la variabilitat respiratòria en sistemes reals amb processat mostra a mostra. Per aquest motiu, en el capítol final d'aquesta tesi, s'ha proposat una nova implementació de la transformació "synchrosqueezing" per tal de realitzar l'anàlisi temporal-freqüencial en temps real, i proveir d'una nova eina per tal d'obtenir la variabilitat respiratòria en temps real, amb mètodes sense contacte.

Contents

List of Figures

List of Tables

Acronyms

ADT	Anderson-Darling test
AM	Amplitude modulation
AUC	Area under the curve
BA	Bland-Altman
BMI	Body mass index
BPM	Breaths per minute
BW	Baseline wandering
CPU	Central processing unit
CUDA	Compute Unified Device Architecture
CW	Continuous wave
CWT	Continuous wavelet transform
DIS	Dense Inverse Search
DRL	Dosimetric reference limit
ECG	Electrocardiogram
EEG	Electroencephalography
ERL	Exposure reference level
ERV	Expiration reserve volume
FFT	Fast fourier transform
FIFO	First In First Out
FLANN	Fast Library for Approximate Nearest Neighbours
FM	Frequency modulation
FN	False negative
FOV	Field of view
FP	False positive
GPU	Graphics processing unit

HOG	Histogram Oriented Gradients
IBV	Biomechanics Institute of Valencia
ICC	Intra-class correlation
ICNIRP	International comission on Non-Ionizing Radiation Protection
IF	Instantaneous frequency
IFFT	Inverse fast fourier transform
IQR	Interquartile range
IR	Infrared
IRV	Inspiratory reserve volume
KLT	Kanadae-Lucas-Tomasi
LED	Light emitting diode
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MEMS	Microelectromechanical systems
mHealth	Mobile health
MINECO	Ministerio de Economia, Industria y Competitividad
NCD	Non-communicable diseases
NIR	Near-infrared
ORB	Oriented FAST and Rotated Brief
PD	Photodiode
PPG	Photoplethysmography
PPV	Predictability of positive value
PSD	Power spectral density
RANSAC	Random Sample Consensus
RC	Respiratory cycle
RF	Radio frequency
RGB	Red-green-blue
RIIV	Respiratory-induced intensity variation
RIP	Respiratory inductive plethysmograph
ROI	Region of interest
ROS	Robotic Operative System

RR	Respiration to respiration
RSA	Respiratory sinus arrhythmia
RV	Residual volume
RX	Receive
SD	Standard deviation
SDE	Standard deviation of the error
SEN	Sensitivity
SIFT	Scale-Invariant Feature Transform
SST	Synchrosqueezing transform
TLC	Total lung capacity
TP	True positive
TV	Tidal volume
TX	Transmit
VC	Vital capacity

Chapter 1

Introduction

The ability to measure physiological signals from the human body can be considered one of the most significant achievements of humanity in the last century [1]. Since the first electrocardiogram (ECG) invented by Dr Willem Einthoven in 1901 [2], nowadays measuring physiological signals has evolved in a way that has become an standard procedure in medicine. Moreover, the information that can be obtained from the different signals of the body has improved the diagnosis and treatment of multiple diseases, as well as allowed the measurement of other health related parameters that could not have been obtained otherwise [3]. As an example, by obtaining the heart rate variability, which is derived from the ECG, multiple heart diseases can be assessed [4], but also important, the stress levels of a subject can be objectively assessed as well [5].

Measuring these variables had been traditionally performed on ambulatory setups, or in a very controlled environments with dedicated or ad-hoc expensive hardware. Nowadays, due to the technical and technological improvements, an everyday increasing number of these physiological variables can be measured with mobile devices, making these measurements feasible without the need of specialised hardware, even in uncontrolled environments or during daily life activities [6, 7, 8].

As vital sign measurement is becoming more ubiquitous to society due to the proliferation of these technical and technological improvements, a new concept based on personal monitoring has emerged; mobile health (mHealth). mHealth encompasses a series of services, from health tele-monitoring to self-quantification, being this last one defined as monitoring and acquiring physiological variables by the means of mobile devices. In 2014 the European Commission reported [9] that mHealth solutions favouring self-measurement could potentially help in the detection and diagnosis of multiple non-communicable diseases (NCD) that otherwise would not have been detected, in a way, that this practices could play a crucial role in the health sector of the member states. Another survey performed in 2015 from the World Health Organization [10] regarding the adoption and implementation of mHealth on the European Union, reported that more than 46 % of the state members had already mHealth policies in place. By promoting mHealth among citizens, and the use of mHealth applications for self-monitoring, could potentially provide real-time and portable access to health care while promoting more healthy lifestyles and self-quantification.

Furthermore, the use of mobile devices and ad-hoc sensors to perform physiological measurements instead of expensive medical equipment, has experimented an exponential increment in the last decade [11, 7]. Moreover, with the growth of the aforementioned mHealth policies on the European Union, the need to develop and present new novel methods to obtain physiological variables with these type of devices will certainly present a challenge regarding the limitations of these devices, as most of them were not designed to measure vital signs. Furthermore, as the measurements will certainly not be performed in controlled environments, the acquisition systems based on mobile devices present new opportunities to researchers to tackle how the measurements have to be performed, and more importantly, how to assess the quality of the obtained signals. Nevertheless, not all the physiological variables are suitable to be obtained throughout these methods, for instance the signals that require specific interfaces, for example, encephalography which requires a certain type of electrodes distributed on a certain position around the head [12], as well as its signal conditioning circuits, may present a challenge if this signal was to be obtained via a mobile device.

Another limitation of these type of methods is the interaction with the user, for instance, on ambulatory setups the subject has clear constraints on how the measurement has to be performed while no interaction from the subject with the measuring device is needed. mHealth setups, on the other hand, and more specifically for those applications of mHealth involving self-quantification, the subject is the one responsible to perform the measurement with no necessary previous knowledge of how the measurement has to be performed, moreover, the subject has to interact with the device while the physiological signals are being acquired, which could lead to errors in the measurement or in worst case scenarios to a lack of adherence to self-quantification [13, 14]. Regarding this last limitation, other studies in the literature remark the use of unobtrusive methods for this same purpose, being those based on wearable devices that can communicate with the smartphone or computer to transmit the physiological data [15, 7]. Wearable devices are not exempt of limitations regarding usability, being the extended period of use the main issue of abandonment of these devices, as they can become uncomfortable to wear or even produce skin irritation if used during a prolonged period.

Recently, to solve the issues produced by the wearable sensors and mobile measurements, and due to the miniaturization and performance increase of the existing hardware, new techniques based on non-contact measurements have gained more popularity. Those techniques are based on sensing vital signs at a distance, relying on sensors and methods that allow for a measurement without the need of being in direct contact with the subject. Those methods are based on multiple technologies, being the most relevant the ones based on video and the ones based on Doppler radar [13, 8]. With these two methods the limitations regarding the subject interacting with the measurement devices, as well as some of the aforementioned limitations, are mitigated as the measurements can be performed at a distance without the need of the subject to be in direct contact with the sensor or the measuring device. Even though these methods have an advantage in this regard they also present some drawbacks, as an example, video-based methods could raise concerns about the privacy of the subject, while radar-based methods could rise concerns about electromagnetic emissions and its impact on human health. Furthermore, there are only a few video-based methods on the literature that do not require any interaction with the subject [16], for example: selecting a region of interest or standing/seating in a concrete position.

One of the main limitations of these methods is the quality of the extracted signals, as the measurements are performed without contact, the physiological signals that can be obtained are limited. Moreover, the errors in the obtained signals have a strong dependence on the movements of the subject and other environmental factors such as lighting conditions or even vibrations [17, 18, 19, 20]. However, and even though the aforementioned limitations regarding the measurement quality and the user interaction, mHealth technologies and the use of non-medical hardware to perform physiological measurements has the potential to produce a huge benefit to society.

Regarding the different physiological variables that can be obtained from the human body, as an example, two types of signals can be drawn [12, 21]: the ones that can be obtained by measuring bio-potentials, and the ones that can be measured by the mechanical displacement of certain regions of the body. The first ones are produced by nervous and muscular tissue, and are transmitted to the skin via ionic exchanges between cells. The latter ones are produced by the mechanical movement of different muscles and organs inside the body. These two categories are not completely independent between each other, as an example, the heart rate can be either measured by sensing the electrical activity of the heart by the means of an electrocardiograph, and can also be obtained by sensing the movement of the chest produced at each heart beat by the means of a ballistocardiograph.

There are other signals that can only be measured either electrically or mechanically, being one example the encephalography signal which can only be obtained by measuring the biopotential changes at different regions of the head, which are related to the activity of the brain. Nevertheless, as the human body is a complex system, there are other physiological variables that can be directly obtained by measuring the mechanical changes, and be indirectly obtained by the modulation it produces onto other vital signs, as an example, the respiratory signal can be directly measured either by the displacement of the chest or by the volume of air that is being displaced, but as the lungs change the position of the heart whenever they inflate or deflate, the respiratory signal can also be obtained as a modulation of the ECG and pulse wave. For this reason, the respiratory signal has a special relevance as it is a physiological variable that can be easily obtained, even through the naked eye.

The respiratory signal has been traditionally used in medicine to assess different pathologies related to the respiratory system or different sleep disorders, not much attention has been paid to this signal from the area of biomedical research, as an example, there are hardly any specific algorithms for analysing this type of wave if compared with other biomedical signals such as the electrocardiogram. Nowadays, it has been demonstrated in the literature that the respiratory signal contains as much information about the physiology of the body [22] as other signals such as ECG or the pulse wave, making the study of this signal a current research topic. As stated by Tipton *et al.* [23] the frequency and amplitude of the respiratory signal can experiment changes whenever a subject is presented with fear, pain or other physiological stressors such as hypoxia or hyperoxia. Another study by Nicoló *et al.* [24] showed the correlation between changes in frequency and amplitude of the respiratory signal whenever a person is performing a physical activity and how this changes are modulated by the type of exercise that is being performed. Finally, In the research group where this doctoral thesis has been developed, for years there has been a line of work for the detection and assessment of drowsiness at the wheel by the means of the analysis of the respiratory signal [25, 26]. One of the results of this line of research, was the proposition of a new method based on the variability of the respiratory signal in order to obtain an indicator of the drowsy state of the subject.

As of this day, the main problems of measuring respiration throughout non-contact methods is that the obtained signals cannot be used to detect instantaneous changes in neither the amplitude and/or frequency, with enough precision and accuracy so the aforementioned indicators and parameters, such as stress or drowsiness, could be correctly assessed. This poses as an opportunity to develop methods that are capable of having all the advantages of non-contact methods, while at the same time be able to measure the respiratory signal of the subject with enough accuracy and precision so this indicators could be extracted. On the other hand, the presented methods should also be able to overcome the limitations regarding the user interaction and usability. For example, by restricting the user interaction with the measurement setup and the measurement itself, favouring the repeatability of the measurements as the human error factor is greatly attenuated.

1.1 Objectives

Inspired by the prior statements about self-monitoring and physiological measurement, the aim and main objective of this thesis is develop new methods for non-contact unobtrusive monitoring of the respiration of the subject, ensuring that all the developed methods are capable to obtain signals with enough quality, to be used in the future to extract relevant physiological information from the respiratory signal.

Once the main challenge is achieved, the second main objective of this thesis is to adequate the proposed algorithms to comply with a series of constraints that truly allow unobtrusive monitoring, which includes real-time performance and ubiquitous measurements. Real-time performance is crucial as it provides the subject with near-instantaneous information regarding the obtained respiratory signal. On the other hand, ubiquitous measurements imply the ability to perform a measurement without the need of the subject to perform any specific action, as an example, select a region of interest or adopt a specific posture.

In order to comply with the two main objectives, a series of sub-objectives have been outlied:

1. Validate whereas a method based on a consumer video camera is suitable to extract the respiratory signal, and successfully develop a video-based method to obtain the respiratory signal of a subject.

2. Evaluate the respiratory signal obtained with different video-based methods, based on dif-

ferent hardware, to assess the accuracy and precision of the methods in terms of respiratory rhythm analysis.

3. Improve the proposed methods based on the findings of the previous objectives, regarding the differences in the used hardware and the limitations found.
4. Asses the viability of using an opportunistic approach in the measurement of the respiration based on the quality of the signal.
5. Improve the usability and ubiquity of the proposed methods, to allow a null user interaction while automatically assessing the quality of the obtained signals.

As a final sub-objective of this thesis, an optimization of the algorithms for respiratory signal analysis to perform in real-time has been proposed, so the signals obtained with the proposed methods can be analysed on-line while are being obtained.

1.2 Thesis Framework

This PhD thesis has been developed in the Electronic and Biomedical Instrumentation group, within the Department of Electronic Engineering from the Universitat Politécnica de Catalunya. The Electronic and Biomedical Instrumentation Group has been actively researching for more than 30 years in the field of biomedical engineering, and during the last year, some of the research has been focused on unobtrusive methods [27] and the assessment of healthy lifestyle habits [28].

This study was performed as a part of the MINECO project ”DESARROLLO DE METODOS DE MEDIDA DEL NIVEL DE ESFUERZO/RECUPERACION EN LA PRACTICA DE EJERCICIO FISICO, BASADOS EN LA ACTIVIDAD CARDIOVASCULAR, TEMPERATURA Y RESPIRACION” with reference DEP2015-68538-C2-2-R, and as part of the non-competitive project ”DETECCIÓN DE SOMNOLENCIA EN CONDUCTORES” with reference C-10888. Both projects share as goal the obtention of physiological variables with unobtrusive methods, the first project with the aim of monitoring healthy lifestyle habits, while the second project is aimed to drowsiness detection in drivers. Nevertheless, both projects have benefited from the work performed in this thesis, as the methods presented can be used in both scenarios. Moreover part of the work described in this thesis has been patented under the title *Respiratory Signal Extraction* with patent number WO2018/121861.

Part of this thesis has been published and already submitted in recognized journals in the field of electronic engineering, biomedical engineering and instrumentation.

1.2.1 Structure

The present thesis is structured in 8 chapters including the introduction, which encompasses a thorough state of the art revision, the five fundamental chapters which conform this thesis, and finally, the conclusions and future work.

An state of the art overview is presented in **Chapter 2**, where different methods that can be used to obtain the respiratory signal are reviewed, as well as an introduction to the respiratory signal, its mechanics and the signal processing techniques that can be used to obtain relevant information from the raw respiratory signal.

Chapter 3 contains the first method proposed on this thesis to obtain the respiratory signal from a subject using an RGB camera. The method is based on capturing the thoraco-abdominal movements by the means of a custom pattern placed on the chest of the subject. The proposed method is designed to perform in real-time, with no interaction with the subject whatsoever. This method was crucial to the development of this thesis as it was the starting point and the first approach to non-contact video-based methods. This chapter fulfils the first objective of validating a consumer-grade camera method for respiratory extraction as well as the obtention of the respiratory signal of a subject.

Chapter 4 provides a comparison between non-contact video methods for respiratory signal measurement, fulfilling the second objective. This chapter focuses on how the respiratory signal obtained with three different methods under a controlled environment compare with the signal obtained with a reference method. Three different cameras are used to perform the analysis with three different respiratory signal extraction methods, being the one based on an RGB camera the one studied in chapter 3. A time-frequency analysis is also performed to assess the differences between the methods and the reference method. This chapter provides a better understanding of the limitations and advantages of each one of the used methods, as well as a first approach on respiratory rhythm measurement using time-frequency analysis. The results obtained in this chapter provided the basis for the next chapters of this thesis

In **Chapter 5** the development of an hybrid Depth-IR method to obtain the respiratory signal from the subject while seating, is presented. On the light of the findings in chapter 4 regarding the results of the Depth method, and as this type of camera presents some advantages in respect of the RGB camera, the method presented in chapter 5 has been designed to overcome the limitations present in the method proposed in chapter 3, thus fulfilling the third objective of this thesis. As the chapter 3 method relied on a pattern detection to properly work, this posed as a limitation on environments with light changing conditions or vibrations. On the other hand the method presented in chapter 4 as it relies on Depth point-cloud to obtain the respiratory signal, the lighting limitations were not as severe. For this reasons, and also in fulfilment of the first two objectives, a method based on depth point-cloud and IR video was presented. The method presented in this chapter was designed to function in low-light environments and in light changing conditions. In fulfilment of the fourth objective of this thesis, the hypothesis of obtaining the respiratory signal in an opportunistic approach has been assessed and validated in this study.

Chapter 6 presents the culmination of the learnings and findings from the three previous chapters as well as achieving the fifth objective of this thesis, where the limitations of the previous methods regarding the position of the subject and the need to detect either a pattern or the face of the subject are corrected. The proposed method in chapter 5 was conceived based on the depth-IR method findings, as it makes use of the same type of algorithm that is used in chapter 5 to detect the subjects movement. Even though in the presented chapter an RGB camera is used, an IR camera could be used as well as the method does not require any specific camera hardware. The novelty of the proposed method lies on obtaining the respiratory signal of the subject while at the same time obtaining a quality indicator in real-time. Moreover, as the measuring point is located at the side of the subject and no region of interest is needed for the method to work, the null user interaction is achieved.

Chapter 7 presents a novel implementation of the Synchrosqueezing transform method. This chapter is relevant to this thesis as it provides with the necessary tools to extract relevant information from the respiratory signal in real-time, while on the other hand providing a tool to evaluate all the signals extracted from the aforementioned methods. This final chapter attain the final objective of this thesis, by providing a real-time evaluation tool to process the respiratory signal.

Finally, **Chapter 8** presents the general discussions and conclusions of this thesis.

1.3 List of Journal Articles

List of journal articles derived from this thesis:

The first journal article is entitled ”Comparison of video-based methods for respiration rhythm measurement.” [29], and it is derived from chapter 4.

The second journal article is entitled ”Non-contact Infrared-Depth Camera-based Method for Respiratory Rhythm Measurement While Driving.” [30], and it is derived from chapter 5.

The third journal article is entitled ”Camera-Based Method for Respiratory Rhythm Extraction From a Lateral Perspective.” [31], and it is derived from chapter 6

Chapter 3 has been submitted as an article to the Biomedical Signal and Processing Control Journal.

Chapter 2

State of the Art

This chapter presents a brief introduction, description and the nowadays knowledge of the fundamental concepts, methods and methodologies that have been used by the author to elaborate this thesis.

2.1 Breathing

Breathing is the mechanical process by which the air moves in and out of the lungs. The purpose of breathing is to allow intake of fresh air, where oxygen (O_2) and carbon dioxide (CO_2) are exchanged, and allows its posterior release to the atmosphere. The act of inspiration (inhaling air) is produced due to the contraction of the muscles in the chest, where the fibres of the diaphragm muscle shorten, which translates into a piston-like movement that pulls the lungs down, while at the same time the external intercostal muscles contract, expanding the thoracic cage which enlarges the thoracic cavity, hence producing a negative pressure (relative to the atmospheric pressure) that favours a flow of air (Boyle's Law) into the lungs/alveoli. On the other hand, expiration (exhaling air), is produced by the thoracic muscles (diaphragm and external intercostal muscles) relaxation, which will return to its initial state thus reducing the volume of the chest cavity, producing a positive pressure which favours the release to the atmosphere of the air inhaled during the inspiratory process [12].

Even though the muscles in the thoracic cage (intercostal muscles and diaphragm) are the main responsible of the expansion and contraction of the upper torso, there are studies in the literature [32, 33] that show that there are other factors that contribute to the chest wall mechanics, being one of them the abdominal movement. In [32] the movements of the thoracic cage and abdomen are measured separately while breathing demonstrating that the respiratory movement has at least two degrees of freedom, where the volume change due to the abdominal muscle contraction and expansion play a key role in the act of breathing. The study presented in [34], also demonstrated that the abdominal muscles have an important role to prevent rib cage distortion while breathing during exercise.

The combined effort of the thoracic and abdominal muscles [32, 33], along with multiple external factors, influence the amount of air that enters the lungs. To characterize and define the volume of air that is being displaced in and out of the lungs in each breath, the following definitions have been taken into account [12, 35, 36]:

- *Total lung capacity (TLC)*: is the maximum air volume that can be inhaled into the lungs.
- *Vital capacity (VC)*: is the maximum volume of air that can be exhaled after a deep breath (when the maximum air volume has been inhaled). This value strongly depends on the subject and it is strictly related with age and gender.
- *Residual volume (RV)*: is the minimum volume of air that remains in the lungs after a long expiration. The residual volume also depends on the subject and it is also related with age and gender.
- *Tidal volume (TV)*: is the average volume of air expired and inspired at each breath. This value is related to normal breathing function and it occurs spontaneously. Tidal volume variability is related to different physiological processes, as an example: the tidal volume will not be the same if the subject is at rest versus performing physical activity.
- *Inspiratory reserve volume (IRV)*: is the extra volume of inspired air over the Tidal Volume level when performing a deep breath (forced inspiration).
- *Expiration reserve volume (ERV)*: is the maximum volume of expired air when exhaling, normally by performing a forced expiration.

Figure 2.1 contains a representation of the different parameters related to breath volume.

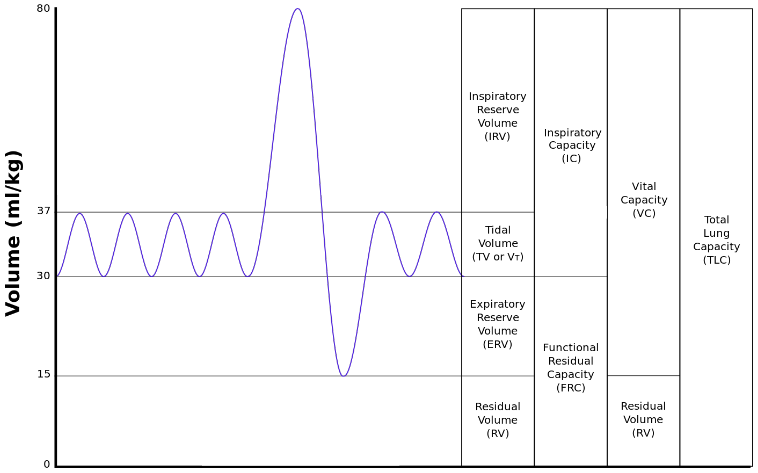


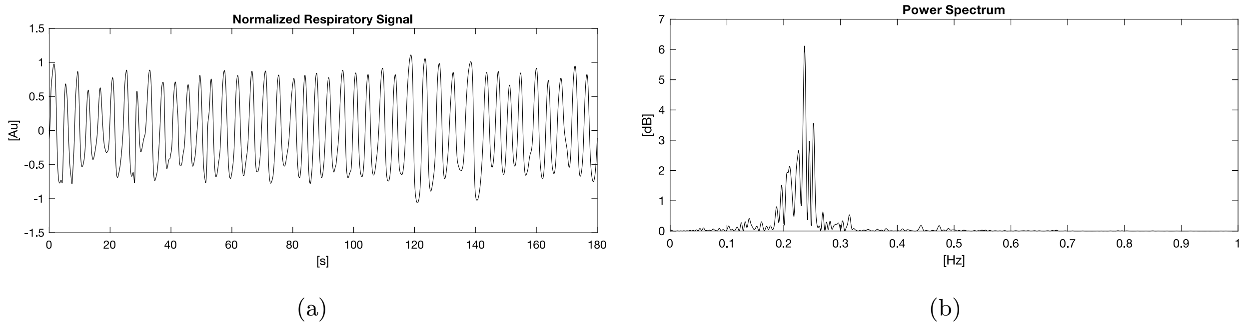
Figure 2.1: Lung volume versus time and its parameters, representation by Kapwatt under CC BY-SA 3.0 (https://commons.wikimedia.org/w/index.php?curid=74891988).

Tidal volume (TV) has a special relevance to this thesis, as it is the magnitude that will be measured, with different methods, throughout the presented work. The importance of the TV is that its temporal evolution and variability contains information about the subject and its activity, for example: if the subject is performing a physical activity or even if the subject is talking [37, 24].

Breathing not only consists on inhaling and exhaling air, the rate at which the air is breathed also contributes to the normal behaviour of the respiratory system. The normal respiratory rate of a healthy adult could vary between 14 to 20 breaths per minute at rest [38], being the rate at which inspiration and expiration occurs an involuntary reflex regulated at the medulla in order to maintain the homeostasis between O_2 and CO_2 at the alveoli. This respiratory rate is not completely regular as it can change due to external factors such as: speech, exercise, walking, stress or even environmental changes. Respiratory rate can also be voluntarily forced by the subject to maintain (hyperapnoea) or to exhale (apnoea) all the air from the lungs [39, 23, 40].

2.1.1 Respiratory Signal

The respiratory signal is the time series obtained by continuously measuring the breath of a subject (tidal volume). The respiratory signal is conformed by breathing cycles, and it can be conceptualized as the temporal evolution of the tidal volume. Each breathing cycle comprises a combination of inspiratory and expiratory movements, whereas the proportion between inspiration and expiration is not constant. As the breath of a subject is not constant [41], neither will be the respiratory signal.



(a) (b)

Figure 2.2: Normalized Respiratory Signal from a reference method and its PSD of unconstrained breathing (free breathing).

Figure 2.2 depicts an example of a normalized respiratory signal, obtained with a thoracic inductive plethysmograph, and its power spectral density (PSD). As it is seen in Figure 2.2b the average density in the power spectrum is comprised between 0.18 Hz and 0.27 Hz, with a peak frequency of 0.24 Hz, which can be converted to 14.4 breaths per minute. Note that the peak frequency is within the range of 14 to 20 breaths per minute at rest [38] as stated in the literature. The PSD bandwidth is due to the variability of the respiratory signal. This variability occurs naturally as it is directly related to the tidal volume, hence the frequency of each respiratory cycle can increase or decrease depending on the situation or environment [39, 42] thus producing a certain bandwidth on the PSD.

In order to characterise the respiratory signal in terms of rate and frequency, the time period of each respiratory cycle is obtained. A respiration cycle, as defined previously, is the time that takes to perform a whole breath (inspiration and expiration). In order to obtain the duration of each cycle, and taking into account the tidal volume and its morphology, the time between positive slopes of two consecutive breaths is computed. Figure 2.3 depicts the process by which the breathing cycles are extracted. The respiratory cycle series, which is composed by the concatenation of the time between consecutive breaths, is also shown in Figure 2.3.

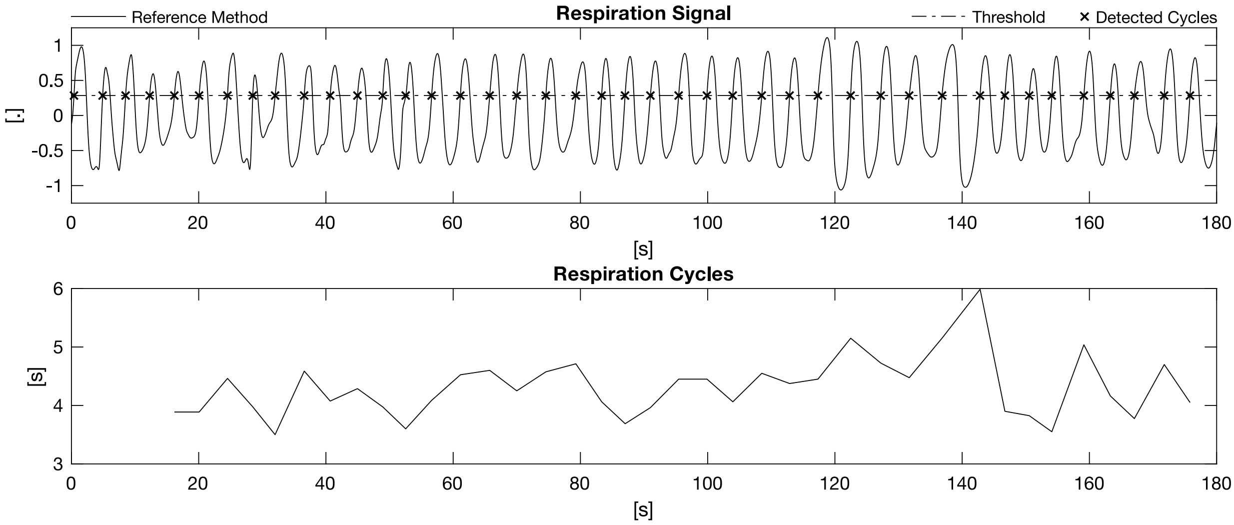


Figure 2.3: Respiratory signal (Figure 2.2a) with the detected Breathing Cycles and threshold, and the representation of the duration of the respiratory cycle signal.

Studies involving the respiratory signal have been traditionally performed to assess sleeping disorders such as sleep apnoea or related to cardiovascular diseases [43, 22, 44, 45]. Nowadays the analysis of this signal has exceeded the medical practice [40] to reach new fields of study such as: drowsiness detection [26] or even to asses the contribution of the respiratory rate to concentration on meditation practice [46].

2.2 Overview on Contact Methods

In this section a subset of technologies to acquire the respiratory signal based on contact methods are reviewed. These kind of methods consist on transducers or sensors that are placed in direct contact with the body of the subject, an example of those comprise: electrodes, thermistors, inductive bands or even turbines and optical fibres. The aim of these sensors is to obtain through different physical properties of the body (electrocardiogram modulation, change in temperature at the nostril...) the respiratory signal of the subject. The methods that make use of these type of transducers can be categorized in two groups: obtrusive and unobtrusive, regarding the type and location of the sensor within the body of the subject, and the level of perturbation caused by the sensor on the daily life of the subject.

2.2.1 Obtrusive Methods

Obtrusive methods are defined as those that impede the normal behaviour of the subject while performing a measurement, whereas the location of the sensor or the sensor itself are the cause.

Airflow-derived Respiration

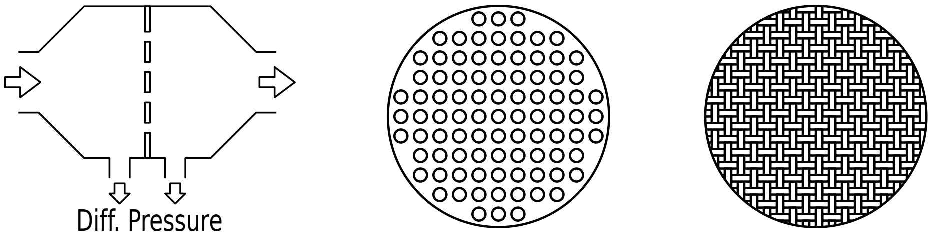
The first methods that were used to obtain the respiratory signal were based on quantifying the amount of air that goes in and out of the lungs, an example of these methods comprise: neumotachographs, body plethysmographs, hot wire anemometers or turbine-based flowmeters among others.

As there are plenty of methods to acquire airflow-derived respiration, only the neumotachograph is going to be described in detail. Historically this is one of the most relevant methods, while nowadays it is typically used as a reference to validate new methodologies in ambulatory setups. Other methods like hot wire anemometers or turbine-based flowmeter will also be briefly described as they present the current state of the art on these type of devices.

Pneumotachograph

Pneumotachograph derives from the word pneumo and tacho, which mean lungs and speed respectively in Greek. This method is based on quantifying the amount of displaced air (volume) that goes in and out of the lungs [35]. To perform this measurement, a Pneumotachograph is composed by a cavity in which the air flows and a membrane (or a fine mesh) dividing the cavity to produce a resistance to the flow of air, inducing a difference of pressure between the two halves proportional to the flow. This membrane can be of two types: Fleisch were the resistance in the airflow is produced by capillary tubes [47] and Lilly which consists on a fine wire mesh [48]. The width of the orifices within the mesh or the capillary, determine the sensitivity (proportional to the airflow resistance) of the Pneumotachograph. Figure 2.4 shows an schematic drawing of a Pneumotachograph and the two types of membrane Fleisch: 2.4b and Lilly: 2.4c.

The difference in pressure between the two halves of the cavity can then be measured by two pressure transducers (one for each halve), or a differential pressure sensor, to obtain the amount of air that flows within the cavity. The difference in pressure inside the two cavities is due to the Venturi effect, being the difference in pressure proportional to the square of the flow of air. Pneumotachographs can differentiate between inhaled and exhaled air since the flow of air is bidirectional, but this also poses an inconvenient as the exhaled air does not have the same chemical composition, temperature and humidity as the inhaled air. The difference between inhaled and exhaled air can produce a change in the baseline over time and also produce sensitivity changes if the sensor is not compensated for these effects [35].



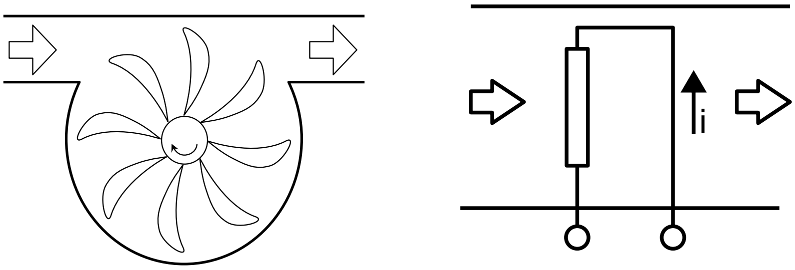
(a) (b) (c)

Figure 2.4: a) Schematic drawing of a Pneumotachograph, with two types of membranes: b) Fleisch and c) Lilly

This method is normally used in medical practice to perform ambulatory measurements of the respiratory flow in both adults and infants. Another use of this type of device is to monitor artificial breathers or embedded into medical devices. For research purposes it is also used as a reference method to asses new techniques to obtain the respiratory signal [49].

Other Methods

Turbine flowmeters (Figure 2.5a) are devices that consist on a turbine inside a cannula that is usually placed at the end of a mouth piece. The turbine can freely rotate in both directions with the flow of air [50], with the rotation speed of the turbine inside the cannula being proportional to the flow of air that goes through. The direction of the rotation can also indicate if the air is being inhaled or exhaled. This type of devices have a linear response to the flow of air, if the flow increases the rotation of the turbine increases as well. As the flow of air is quantified by a mechanical movement, this devices are robust to temperature and humidity changes within the range of the breathed air. Turbine flowmeters can be found in certified spirometers used in medical practice, moreover, new designs have been presented with MEMS components [51] making this type of device ideal to integrate in portable spirometers.



(a) (b)

Figure 2.5: a) Schematic drawing of a Turbine flowmeter and b) Hot-wire anemometer

Hot-wire anemometers (Figure 2.5b) are made with one or more heated wire inside a tube that can exchange heat with the flow of air. The amount of flow inside the tube is measured by sensing the current that goes through the wires: as more flow of air passes through the tube the wires get colder, thus changing the resistivity proportional to the temperature change, hence the current that flows through the wires will also change [52, 53]. This type of sensors present a non-linear response to the flow of air (small changes in the air flow produce a significant variation on the current that flows through the wire), and a flat frequency response, making them ideal for acquiring the respiratory signal on infants [54]. One downside of this type of sensor is that humidity and moisture changes can affect the accuracy of the sensor if not compensated correctly [55]. Nevertheless, both the Turbine flowmeter and the Hot-wire Anemometer are used as a gold standard to validate other methods to obtain the respiratory signal, and are usually integrated, as a reference monitor, in artificial respirators and other medical devices.

Although the Pneumotachograph, Turbine flowmeter and the Hot-wire Anemometer are widely used in medical environments or for the validation of other systems, one of the main drawbacks of these type of devices is the presence of a collector in order to acquire all the inhaled and exhaled air. As this collector is placed on the face of the subject it can alter the way the subject breaths [56]. This collector can be a face mask that covers the whole nose and mouth region, or it can also be a mouth piece where a clip needs to be placed on the nose of the subject. Either way the collector alters the way of normal breathing and perturbs the normal behaviour of the subject during the measurements.

ECG-derived Respiration

The electrocardiogram (ECG) is the measurement on the surface of the body of the electrical activity of the heart induced by action potentials produced by the sinoatrial node (sinus node) [21] during each heart beat. These action potentials induce ionic currents that are measured through the skin of the subject by the means of transducers called electrodes. In order to obtain the ECG wave, two or more electrodes must be placed on the skin to obtain the difference in voltage over time between at least two of the selected locations [21], hence obtaining the ECG wave. In Figure 2.6 a classical ECG wave is depicted, composed by the P, QRS complex and T waves.

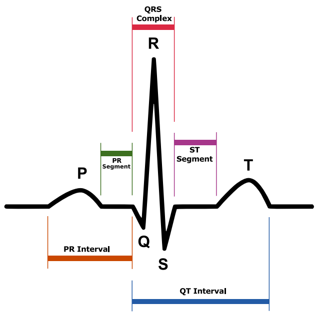
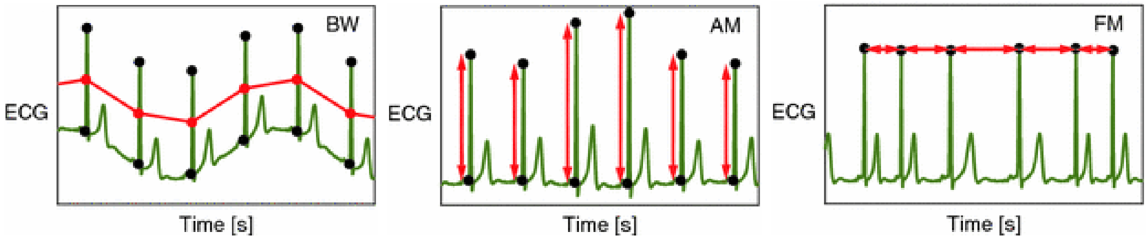


Figure 2.6: ECG with P, QRS and T waves. Public Domain image extracted from: https://commons.wikimedia.org/wiki/File:SinusRhythmLabels.svg

Breathing affects the ECG wave as a modulation of the signal, an example of these modulations are depicted in Figure 2.7. These modulations can appear as baseline wandering (BW), amplitude modulation (AM) and frequency modulation (FM). Each one of these effects is produced by a different physiological process: the BW and AM are produced by the changes in the electrical vector of the heart relative to the electrodes produced by the muscle activity involved in breathing [57], thus producing a baseline shift in the ECG wave and an AM modulation to the ECG. The FM modulation, on the other hand, is produced by the respiratory sinus arrhythmia (RSA), which is caused by the acceleration and deceleration of the heart beat due to the inhalation and exhalation movements [58, 59]. It has been extensively demonstrated in the bibliography that these modulations can appear with different strengths or magnitudes depending on the subject, as the physiological process that produce them may not be present or dampened. As an example, the age of the subject has a direct influence on the FM modulation, as the RSA and chest wall expansion both diminish with age [60]. Figure 2.7 shows the three aforementioned modulations.



(a) (b) (c)

Figure 2.7: Effects of the respiratory signals in an ECG wave. Figures modified by the author, adapted from Charlton *et. al.* [60] (CC BY-NC 4.0: http://creativecommons.org/licenses/by-nc/4.0/).

The respiratory signal is obtained from these modulations as it is depicted in Figure 2.7 using the following signal processing techniques: the BW is obtained by tracking the position (baseline) of the R peak of the QRS complex, the respiratory signal derived from the AM modulation is acquired by tracking the amplitude changes of the R peak in the QRS complex, and finally the respiratory information from the FM modulation is retrieved by computing the spectrum of the ECG signal or by computing the time difference between R peaks in the QRS complex [61, 62]. In this latter modulation, the mean respiratory frequency is contained in the maximum peak of the spectrum within the low frequency range (being the peak centred around 0.23 Hz to 0.33 Hz in average [38]).

Nowadays multiple novel algorithms to extract respiration from ECG have been presented [63, 64, 65, 66] with a wide variety of signal processing techniques. Although the novelty of the methods, one major issue still remain; the ECG acquisition itself. The ECG signal can only be obtained using electrodes in direct contact with the skin, which makes these methods not suitable for daily use, or even to perform respiratory signal measurements in environments different from a medical centre or a controlled setup. Moreover, most of these methods use subjects in a lying position or seated, hence not taking into account the errors produced by the movements of the subject into the respiratory signal.

Thermal-derived Respiration

When air is inhaled from the mouth or the nose, the air is heated while travelling through the trachea to the lungs, and after the O_2 and CO_2 exchange has been produced, this heated air is exhaled. At the same time at the mouth or nose region, while the air is being inhaled, the difference in pressure produces a drop in temperature, and while the air is exhaled, the temperature rises as more heated air from the lungs is pushed out. The temperature of the inhaled/exhaled air ranges from $30.2 \pm 1.7°C$ to $32.5 \pm 1.1°C$ being the room temperature $25°C$ [67]. The lowest temperature at the mouth/nostrils occurs at the end of the inspiration while the highest occurs at the end of the expiration.

These temperature changes are proportional to the amount of inhaled/exhaled air hence proportional to the breathing. For this reason, measuring the changes in temperature at the nostril has been proven a valid method to obtain the respiratory signal of a subject [68, 69, 70]. Figure 2.8 depicts how the temperature at the nostrils is measured using a thermal sensor.

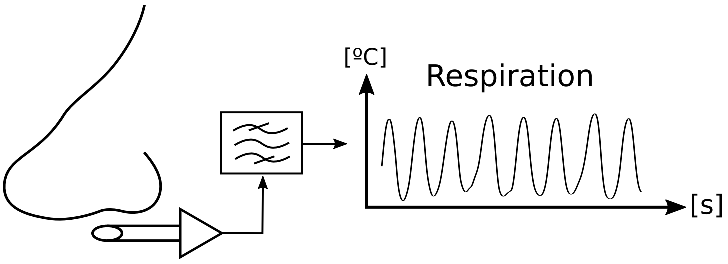


Figure 2.8: Schematic drawing of a thermal sensor under the nose.

Different type of sensors can be used to measure the temperature changes in the mouth/nostril area, two examples of these sensors are thermistors and thermocouples.

Thermistors, in one hand, are temperature dependant resistors that when heated its resistance will change in function of the applied temperature. Commercial thermistors are small, easy to integrate, do not require complex electronics to interface and present a high sensitivity within the range of the thermal changes of the breathing.

An inconvenient of this type of sensor is that presents a high response time, directly related with the sensor size, which may be a problem in order to detect rapid breathing changes or even to obtain the respiratory signal in real-life situations where the breathing is not constant (during exercise or in stressful situations). Nevertheless, this last limitation can be overcome if the sensor is sufficiently small.

On the other hand, thermocouples produce a voltage as a result of the Seebeck effect (thermoelectric effect). This type of sensor is made as the result of the union of two different metals, whereas a temperature change in this junction produces a voltage between the two conductors. This type of sensor is commercially available with different types of junction materials, and each material combination has different properties that confer them unique characteristics to be used in a certain setup. Moreover, small commercial thermocouples are available, where the accuracy is good enough to measure respiration, and furthermore, its response time is faster than thermistors. The disadvantages of this specific type of sensor are that the thermocouple will present different mechanical properties like corrosion or humidity depending on the materials of the junction, which will lower the sensitivity and also have an impact to the cost of the sensor. The most important disadvantage if compared to the thermistor, is the need of more complex electronics to interface with the sensor and with its encapsulation.

Although thermal-derived respiration (using thermistors or thermocouples) has been typically used as a gold standard to validate other respiratory acquisition techniques given their relative low-cost and ease-of-use [71, 72], no newer methods appear to be recently published based on this technique. On the contrary, new methods based on thermographic cameras have been gaining popularity in the recent years, where the thermal changes at the nostrils are measured using this type of sensors [73].

2.2.2 Unobtrusive Methods

Unobtrusive methods are defined as those used to perform measurements without affecting or impeding the normal behaviour of the subject. The unobtrusive nature of the sensor does not depend on the location or on the sensor itself, but on the ease of use and that they usually do not require any interaction/preparation from the subject whatsoever.

Photoplethysmography

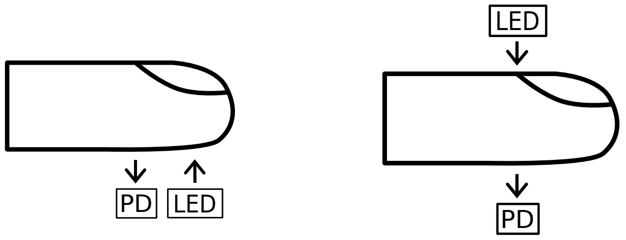
PPG is a method that measures volume variation of a tissue (plethysmogram) via an optical method. One example of signal that can be obtained with this method is the pulse wave of the heart. This signal can be obtained by measuring the variation in the light absorption of different tissues when perfused with blood, where an increase in blood volume produces an increased absorption of light due to a proportional increase in the optical path [74]. The Beer-Lambert law (Equation 2.1) describes how the emitted light is being absorbed in a non-scattering medium, in presence of a substance and proportional to the optical path [75].

$$log_{10}(\frac{I_0}{I}) = \sum_i \epsilon_i C_i d_i \tag{2.1}$$

Where I_0 represents the incident light on the sample, I is the exiting intensity and the summation represents the contribution of all the substances (tissue, blood...) present in the sample. ϵ_i

represents the extinction coefficient (absorption) of each substance, C_i represents the concentration of each substance in the sample and finally d_i represents the contribution to the path length for each substance.

In order to obtain the PPG wave an light emitting diode (LED), red (660 nm) or near infrared (880 nm to 940 nm), must be placed near the skin [74]. There are two combinations of LED and photodiode (PD) that are typically used to obtain the variation in absorbed light: the first one is an LED and a PD placed side by side (reflection) and in contact with the measured tissue, while in the second configuration the photodiode is placed at the opposite side of the LED with the measured tissue in between. Figure 2.9 shows both possible configurations of a PPG sensor. Typically this type of sensor is placed at the tip of the finger, although other body locations such as the ear lobe have been traditionally used [76].



(a) (b)

Figure 2.9: Schematic drawing of the possible combinations of LED and PD. a) Represents the reflection configuration and b) represents the transmission configuration.

The PPG wave is produced by the pressure wave that propagates through the cardiovascular system after each heart beat, where the blood flows through the principal arteries (aorta artery) to the peripheral arteries (distal artery) [76]. The morphology of the PPG wave can be affected by multiple factors such as the diameter and vasodilation of the arteries, blood pressure and changes due to the age of the subject [77]. Figure 2.10 contains a representation of the PPG signal, in this figure the rising flank produced by the incoming blood pressure wave is appreciated, as well as the dicrotic notch produced by the closure of the aortic valve.

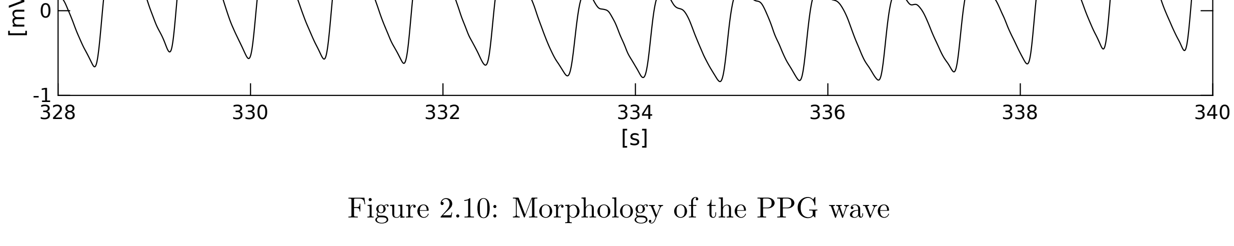
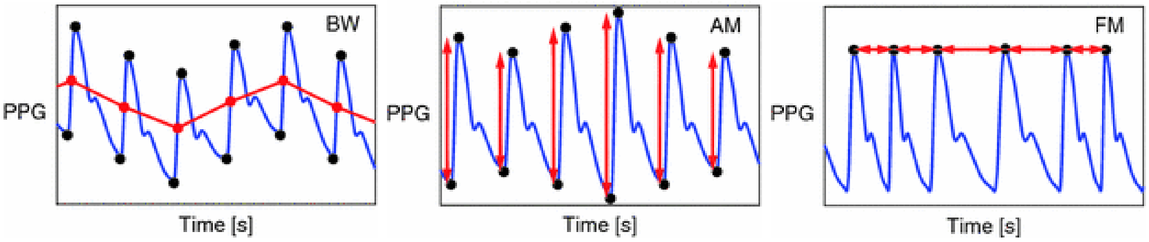


Figure 2.10: Morphology of the PPG wave

There are other physiological processes and properties that affect the PPG wave, these are translated into pulsatile and non-pulsatile variations on the signal. The non-pulsatile contributions are related to the light abortion of the tissue surrounding the blood vessels (bone, skin...), whereas the pulsatile components are related to the changes in the cardiovascular system (systole and diastole) and the respiratory system [78]. Another contributing factor to the pulsatile component, that is strongly related to the respiratory system, and more relevant for certain wavelengths, is the light absorption depending on the amount of haemoglobin in the blood. As more haemoglobin is present in the blood, certain wavelengths will be more absorbed and as a result the measured signal (for those wavelengths) will vary in amplitude [79, 80].

Breathing affects the PPG wave in the same way it affects the ECG wave, also presenting three distinctive types of modulation due to breathing: BW, AM and FM [61]. Although the modulations are the same, the physiological process associated to the modulations are fundamentally different in the case of BW and AM. BW is produced by the changes in the blood perfusion of the tissue as a result of the variation in the intrathoracic pressure transmitted through the arterial tree, this effect is due to the blood exchange between the pulmonary circulation and the systemic circulation. BW modulation can also be found in the literature referred as respiratory-induced intensity variation (RIIV) [80]. The AM modulation is also produced by the reduction of volume during inhalation due to changes in the intrathoracic pressure, hence reducing the amplitude of the signal. FM modulation is produced by the same physiological effect that produced the FM modulation in the ECG wave, the respiratory sinus arrhythmia [61, 80]. Figure 2.11 represent the different respiratory modulations on the PPG signal.



(a) (b) (c)

Figure 2.11: Different respiratory signal modulations on the PPG wave. Figures modified by the author, adapted from Charlton *et. al.* [60] (CC BY-NC 4.0: http://creativecommons.org/licenses/by-nc/4.0/).

The respiratory signal from the modulations present in the PPG wave (Figure 2.11) are obtained as follows: for the BW the respiratory signal is obtained by tracking the position of a given fiducial point (minimum, maximum...) at each heartbeat, for the AM modulation the respiratory signal is acquired by tracking the amplitude changes between the maxima and the minima of the PPG wave, finally the respiratory signal from the FM modulation is obtained by computing the time difference between consecutive peaks of the PPG wave.

Nowadays multiple algorithms using a wide variety of signal processing techniques can be found in the literature [61, 80, 79, 81, 82], which describe how to obtain the respiratory signal, with more novel and reliable approaches, from the PPG wave. Moreover, due to the last decade growth in the smartphone industry, nowadays it is possible to obtain the PPG signal using the camera of the smartphone instead of using specialised equipment, or even using wearable devices such as smartwatches, making the measurement of the PPG feasible for everyone.

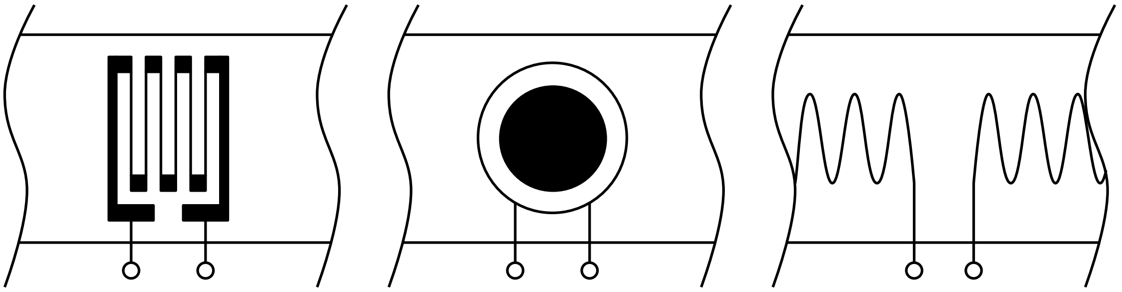
As a limitation of this method to obtain the respiratory signal, and because the PPG wave strongly depends on the subject [77], the acquisition of the respiratory signal through this method may contain errors depending on the age of the subject, as some of the modulations may not be present or with low variability. Although the latter limitation can introduce difficulties in the acquisition of the respiratory signal, it does not render the method unusable. The most important limitations of this method are lighting changes and body motions. As the method is affected by the surrounding light (mostly artificial light) [76], this poses as a serious issue as it can compromise the measurement. Motion artefacts on the other hand, and given that the most common measurement point is the finger, any motion of the hand can change the optical path hence changing the baseline of the PPG wave. This two limitations can render the measurement of the respiratory signal through PPG unusable, as the respiration is obtained from secondary parameters of the PPG signal (modulations), if the signal is contaminated by motion noise or with parasitic components due to external light, and although the PPG wave may still be valid, the respiratory signal may not be retrieved successfully.

Respiratory Effort

The act of breathing involves many physiological processes in order to let air in and out of the lungs, one of this processes consists on contracting and relaxing the inspiratory muscles (diaphragm, intercostal, abdominal muscles...). This contraction and relaxation produce a mechanical change (crucial to the act of breathing) in the thoracic cage [32] that can be measured by the means of a plethysmograph around the thorax [12, 83].

The advantage of using this type of sensors is that no previous conditioning or preparation of the subject is needed in order to acquire the respiratory signal. As these devices do not require direct contact with the skin of the subject, or any particular preparation whatsoever, they can be worn at any time, making these type of sensors specially useful in situations where the subject needs to perform different tasks, or behave normally, while being monitored.

Multiple sensors have been presented to measure the cross sectional area of the chest (volume changes) during respiration. The typical implementations found in commercial devices are: resistive (Figure 2.12a), piezoelectric (Figure 2.12b) and inductive (Figure 2.12c) plethysmographs. All three devices have in common that use a chest strap around the thorax of the subject, but the means to acquire the respiratory signal is fundamentally different for each case.



(a) (b) (c)

Figure 2.12: Schematic drawing of a resistive, piezoelectric and inductive plethysmograph respectively.

Resistive and Piezoelectric Plethysmograph

Resistive plethysmographs are usually build with a resistive strain gauge attached to a semi-elastic chest strap. A strain gauge is a type of transducer used to measure changes in the length of an object as a result of the force applied to it. These type of transducers change its electrical properties (resistivity) proportionally to the fractional change in its length (strain) [12].

The working principle of a resistive plethysmograph, is that the strain gauge within the sensor changes its resistivity in function of the change in volume of the thoracic cage, hence proportional to the tidal volume [32]. A simple case of how this type of sensor works is illustrated in the following example: as the thoracic cage expands (inhaling) the mechanical strain in the gauge increases which produces a proportional increase in its impedance, when the thoracic cage returns to its resting state (exhaling) the mechanical strain decreases as well as the gauge impedance. By measuring the electrical resistance of the gauge, by the means of simple electronics, the respiratory signal is obtained.

Other type of transducers have been presented that take advantage of the displacement of the thorax (strain) to obtain the respiratory signal. Piezoelectric devices are known to produce a voltage difference when strain is applied to them [84], using the same principle of strain increase and decrease proportional to the respiration, and with simple electronics to acquire the voltage difference at the piezoelectric transducer, the respiratory signal is extracted.

Although there are commercial products available to perform respiratory measurements using resistive plethysmographs, their use is limited to educational purposes and they are not commonly used in research studies: one possible explanation to this fact is that these type of transducers suffer from motion artifacts, which means that any involuntary movement of the subject will have a negative impact on the acquired signal. This makes these type of sensors not useful in environments where the subject has to perform different tasks or even when the subject is lying on a bed. Other type of issues related to these type of sensors include the linearity of the relationship between strain and impedance/voltage at the output, which is crucial to obtain an undistorted respiratory signal. One example of strain gauge transducer is the SS5LB [85] sensor from Biopac Systems$^{\mathrm{TM}}$.

Another example of commercial system is the 1370-Kit from SleepSense^{TM} which incorporates a piezoelectric transducer [86].

Inductive Plethysmograph

Inductive Plethysmograph uses a coil that encircles the thoracic cage woven in an "S" or zig-zag pattern along the strap, an example of this type of strap is depicted in Figure 2.12c. By measuring the changes in the inductance of the coil, due to the change in volume of the thorax (cross-sectional area) [32], the respiratory signal is obtained. A typical method to measure the changes in an inductance is to form an LC resonator circuit that oscillates at the frequency given by Equation 2.2. The respiratory signal is extracted by tracking the frequency changes (traditionally with a Phase Lock Loop) of the LC resonator circuit. Other implementations can also be found in the literature [87], but the same inductance change due to the displacement of the thorax still applies.

$$f_i = \frac{1}{2\pi}\sqrt{\frac{1}{L_i C}} \tag{2.2}$$

Where f is the resonant frequency, C is the total capacitance (cable and resonant capacitance) and L_i is the instantaneous inductance of the coil.

Respiratory inductive plethysmograph has been considered an alternative to pneumotachographs when measuring tidal volume (with a previous calibration [88, 89]). Also, this method has been widely used as an unobtrusive method to acquire the tidal volume of the subject without the need to use a face mask [90, 91, 92] or a mouth piece. Moreover, several studies have demonstrated that this method is suitable to be used as a gold standard (reference method) to acquire the respiratory signal in environments where the pneumotachograph, or other obtrusive methods, can not be used [37, 93].

One of the advantages of this method, is that it does not have the issues of the resistive or piezoelectric plethysmography, thus making it robust to postural changes and even, to a certain point, to motion artefacts [93]. This method also presents more accuracy when measuring the respiratory signal than the two aforementioned methods, as the measurement of the inductance (through its resonator circuit) can be performed with better accuracy, while being more robust against motion artefacts.

In this thesis the inductive plethysmograph (RespiBand from BioSignalsPlux^{TM} [94]) has been used as a reference method to validate all the presented methods.

2.3 Overview on Non-Contact Methods

In this section an overview of the recent developments on non-contact methods to acquire the respiratory signal will be performed. These kind of methods consist on devices that can measure at a distance (without contact), either with electromagnetic waves or by light, the physiological variables of the body. As they are more relevant to this thesis, only the methods to obtain the respiration of the subject will be reviewed.

The methods that will be outlined in this section fall in two categories, depending on the means used in order to measure the physiological variables: electromagnetic waves (radar) or light/image variations (video methods). The later can also be divided into three categories given the technology that they use: video, depth or thermographic cameras.

2.3.1 Doppler Radar

Doppler radar can be used to detect human vital signs such as heart beat or respiration, the way a Doppler radar works is by transmitting a continuous wave (CW) to a target (human body). Once it hits the target, part of the wave will be reflected and the other part will be scattered throughout the target. The reflected signal is then received and demodulated to extract the variations in distance between the emitter and the target [95]. The physiological principle which allows a radar to be used to extract physiological information from the body is the following: if the body of the subject is still (which means there is no movement relative to the radar), the movements of the thoracic cage will be proportional to two factors: the mechanical activity of the heart and the displacement produced by the respiration [32]. The movements produced by the mechanical changes in the body, due to the heart and respiration, are tracked by computing the distance between the target and the system by the means of the Doppler effect.

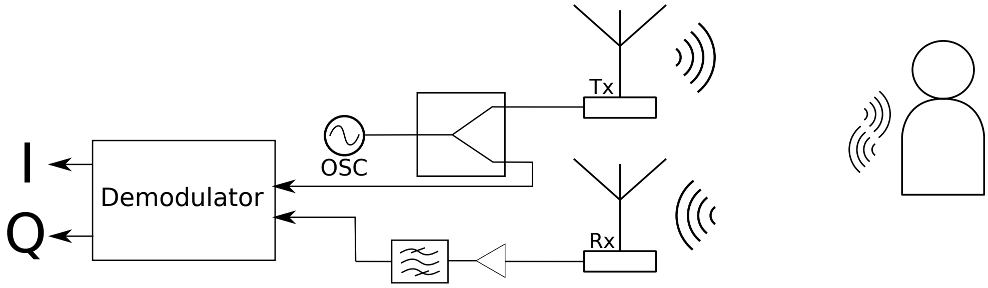


Figure 2.13: Schematic drawing of a radar configuration used to acquire physiological variables.

In Figure 2.13 a classical implementation of a Doppler radar is depicted. The typical architecture uses two antennas, one to transmit (TX) the signal to the target and one to receive (RX) the bounced signal from the target. There is an alternate configuration that uses a single antenna and a duplexer circuit to share the TX and RX channels [96, 19]. This last configuration is less used as it suffers from a problem called null detection, which cause the fading of certain harmonics (odd-order) and the strengthening of others (even-order), as this effect also affects the fundamental harmonics [97, 98] it becomes a problem when acquiring physiological variables. For this reason the typical implementation used to acquire physiological variables uses more than one antenna, so a quadrature transceiver can be implemented, which does not suffer from this kind of issues [99, 100, 97, 101, 102].

Doppler radar for physiological measurements can use a wide range of fundamental frequencies from 100 MHz up to 100 GHz, each frequency has its benefits and downsides. As an example, depending on the frequency, the penetration in the skin will vary as demonstrated in [103], this can become an issue or a benefit depending on the application. Nowadays, the presence of radars in the automotive industry has become a standard, and because the hardware has become incredibly cheap and small, new implementations that make use of the same hardware and the same frequency bands (77 GHz and 27 GHz), can be found in the literature to acquire heart rate and respiration [104, 100]. On the other hand, most of the implementations still use the traditional frequency bands such as: 900 MHz, 2.4 GHz and 5 GHz [105, 106, 101].

The signals transmitted from a Doppler radar can be described as a sinusoidal wave with a given frequency and with a certain phase [107], independently of the antenna or the transmitter topology:

$$T(t) = A\cos(2\pi ft + \phi(t)) \tag{2.3}$$

Where $T(t)$ is the transmitted signal to the target, A is the amplitude of the signal, f is the frequency of the transmitted signal, and $\phi(t)$ is the phase of the transmitted signal.

The received signal can be described with the following equations, given that the quadrature transceiver is the most used in the literature, the received signal has been decomposed to its in-phase and quadrature components.

$$\begin{aligned} R_m(t) &= B\cos(2\pi ft - \frac{4\pi(d_0 + x(t))}{\lambda} + \phi(t - \frac{2d_0}{c})) \\ R_I(t) &= A_I\cos(\frac{4\pi x(t)}{\lambda} + \theta_0 + \Delta\phi(t) + \frac{\pi}{4}) \\ R_Q(t) &= A_Q\cos(\frac{4\pi x(t)}{\lambda} + \theta_0 + \Delta\phi(t) - \frac{\pi}{4}) \end{aligned} \tag{2.4}$$

Where $R_m(t)$ is the reflected signal from the target, B is the reflected amplitude, d_0 is the distance between the acquisition system and the target, $x(t)$ is the displacement of the target relative to the system, λ is the wavelength of the radar carrier wave and ϕ is the phase of the reflected signal. $R_I(t)$ is the in-phase component (I) and $R_Q(t)$ is the quadrature component (Q) of the reflected signal. A_Q and A_I are respectively the amplitudes in I and Q components, θ_0 and $\Delta\phi(t)$ are the residual phase shift.

To extract the respiratory signal from the frequency changes due to the displacement of the thorax, present in both I and Q components as described in the Equations 2.4, multiple approaches have been presented in the literature for the past few years. The presented algorithms vary from the basic demodulation of the frequency shifts at its respective harmonics (heart and respiration) [100], to more complex approaches that use advanced signal processing techniques to obtain better quality signals or even to overcome issues intrinsic to the CW Doppler radar [108, 109].

There are several limitations on using Doppler radar as a method to obtain physiological variables, the first one being; if the attack angle is not perpendicular to the target the bounced signal can present errors, in order to avoid this issue, multiple radar architectures have been presented in the literature [107]. Another limitation is the random body movements, as these type of movements usually occur with higher amplitude than the displacement due to the physiological processes, making it impossible to recover any physiological signal. To compensate for this type of error multiple strategies have been presented [19, 20]. Finally, another limitation of this type of systems is the null transmission depending on the distance to the target, this type of error occurs when the reflected signal produces destructive interference on the receiver. Nevertheless, studies that solve this particular problem can be found in the literature [110]. Nowadays, architectures based on radar specifically designed to avoid the aforementioned issues have been presented, [101, 111] are an example of such implementations.

As Doppler radars emit electromagnetic radiation, it imposes a series of inherent risks or drawbacks that have to be taken into account prior to the use of these type of systems to acquire physiological variables. The first drawback is that Radar-based systems must ensure electromagnetic compatibility with other radio frequency (RF) devices, as the radar radiation can cause interference to other systems using the same frequency spectrum (i.e. Wi-fi at 2.4 GHz or 5 GHz), or even affect electronic devices that can potentially fail due to the interference caused by the radar signal. Guaranteeing that the acquisition system will not interfere with other equipments is a safety requirement that must be taken into account.

Another drawback and one of the main risks of using this type of technologies, is that the subject is illuminated with non-ionizing electromagnetic radiation. To ensure that no harm is done to the subject for long exposure times (long measurements), a series of requirements must be met. As an example, following the directives IEEE C95.1:2019 [112] and ICNIRP 2018 [113], the dosimetric reference limit (DRL) for the whole body between 100 kHz and 6 GHz must be below or equal to 0.08 W/kg and the DRL between 6 GHz to 300 GHz must be below or equal to 20 W/m^2. Another important parameter is the exposure reference level (ERL), which between 2 GHz and 300 GHz must be below 10 W/m^2 (measured over a 30 minute period). These examples only apply to stationary signals which implies a stable amount of radiation, if the source emits an intense pulse, then the following rules should be applied: incident energy density restrictions for intense pulses in the millimetre-wave frequency range (30 GHz to 300 GHz) has to be below $0.2\tau^{\frac{1}{2}}\ kJ/m^2$, where τ is the pulse width in seconds. The IEEE C95.1:2019 [112] directive also specifies that: if the maximum average power is below 20 mW (measured over 6-min period) the evaluation of the DRL between 100 kHz and 6 GHz is not necessary, as the DRL will not surpass the established limit. If the Radar based respiratory acquisition system complies with the above specifications of DRL and ERL, the system does not pose a risk to the health of the subject.

2.3.2 Video Analysis

In the past few years, the development in smartphone and consumer electronic technologies has pushed the video hardware (cameras) in a way that practically every piece of consumer technology includes one of this sensors. As the hardware has greatly improved during the past few years, in terms of resolution and framerate, and thanks to the ubiquity of these devices, the complexity of acquiring physiological variables through them has been reduced to practically a software-based problem. Another advantages of using video analysis to obtain the respiration, or other physiological variables, are: video acquisition is completely unobtrusive as the subject may not even be aware that it is being monitored, the measurement is performed at a distance as no direct contact with the subject is needed, and moreover, multiple subjects can also be measured simultaneously. These reasons make video-based methods a very promising field of study.

Thermal Camera

Thermographic sensors are a special type of cameras that are capable of sensing temperature at a distance. The operational principle behind this special hardware is the ability to measure long-wavelength infrared, which is the part of the spectrum that correspond to the wavelengths between 8 - 15 μm. This type of electromagnetic radiation is emitted by any object that is beyond the absolute zero Kelvin, hence, by sensing this type of radiation, the temperature of the object can be inferred [114]. Nowadays, the field of use for these type of sensors has broadened from surveillance to medical practice, thus making the technology more available to the general public [114].

The use of this type of cameras to obtain the respiratory signal of a subject, is based on the same principle than the thermistor under the nose. As an example, whenever the air is inhaled the inflow of air produces a change in the temperature at the nostrils, and when the air is exhaled, the temperature at the nostril region is increased by the flow of heated air from the lungs. As this type of cameras allow to measure temperature at a distance, by measuring the temperature changes at the nostrils or mouth, the respiratory signal of the subject can be obtained. In the recent years, and due to the technical improvements in this area, thermal cameras have experienced an increase in popularity in the measurement of physiological signals. Moreover, multiple examples can be found in the literature [73, 115, 116, 117] of methods that use thermographic images to obtain the respiratory signal of a subject.

Nevertheless, there are two major limitations of using this type of technology for measuring physiological variables, being the first issue the calibration of the sensor. For example, as the imager sensor is being heated while performing the measurements, if the sensor is not cooled properly or a previous pre-heat and calibration is performed, a drift in the measured temperature will appear over time [118, 119]. The second limitation of these devices is the availability of the hardware; as of nowadays, even though the hardware is more accessible, the required sensor resolution for this type of application is still expensive. Finally, a minor limitation of this method is the resolution and framerate of the used camera. Even though the current technology is capable of providing high resolution thermal imagers, the most common type of commercially available thermal cameras, only have low resolution and low sampling frequency. For this reason if a low resolution/framerate camera is used and placed a few meters away from the subject, or the subject is moving, the measurements cannot be performed.

Video Camera

Video cameras can be found in practically every consumer device, from smartphones to even smart fridges. The ubiquity of these devices has favoured an increase in number of pixels and its framerate, making video cameras a powerful tool to measure the real world. As a result, video-based methods to obtain physiological variables such as heart rate or respiration have become increasingly popular. As the aim of this thesis is to obtain the respiratory signal, only those methods that measure respiration will be overviewed.

Multiple techniques are found in the literature to obtain the respiratory signal from video analysis using red-green-blue (RGB) or infrared (IR) cameras [50, 120, 121, 122, 123]. The physiological principle which allows these video-based methods to acquire respiration is the displacement of the thoraco-abdominal wall. As stated before, this displacement is due to the volume increase and decrease of the thoracic cage due to the air inhalation and exhalation.

Among the multiple algorithms presented in the literature to extract the respiratory signal from video analysis, there is an image processing techniques that has been traditionally used [124, 121] to directly or indirectly obtain the respiration signal. This technique is called optical flow [125, 126, 127], and it also has a special relevance to this thesis as it is used throughout the presented work.

Optical flow was first presented by Kanade-Lucas-Tomasi [128], this algorithm tracks the changes between two consecutive images, allowing a continuous tracking of a feature throughout the time. A modern implementation of this algorithm is described in [125], which allows the tracking of the optical flow of an image in real-time. The next equation (Equation 2.5) describes the transformations and operations that are needed to compute the optical flow between two consecutive frames.

As shown by Bouguet *et. al* in [125], given two consecutive frames I and J, the optical flow is defined as: tracking the evolution of a feature (its new location) from frame I to the frame J. To find this new location, and considering the points defined as $\mathbf{u} = [u_x \ u_y]^T$ and its new location $\mathbf{v} = [v_x + d_x \ v_y + d_y]^T$, the goal of the optical flow is to ensure that $I(\mathbf{u})$ and $J(\mathbf{v})$ are "similar". To do so, the error function shown in Equation 2.5 must be minimized for every feature to track.

$$\epsilon(\mathbf{d}, \mathbf{A}) = \sum_{x=-w_x}^{w_x} \sum_{y=-w_y}^{w_y} (I(\mathbf{x} + \mathbf{u}) - J(\mathbf{A}\mathbf{x} + \mathbf{d} + \mathbf{u}))^2 \tag{2.5}$$

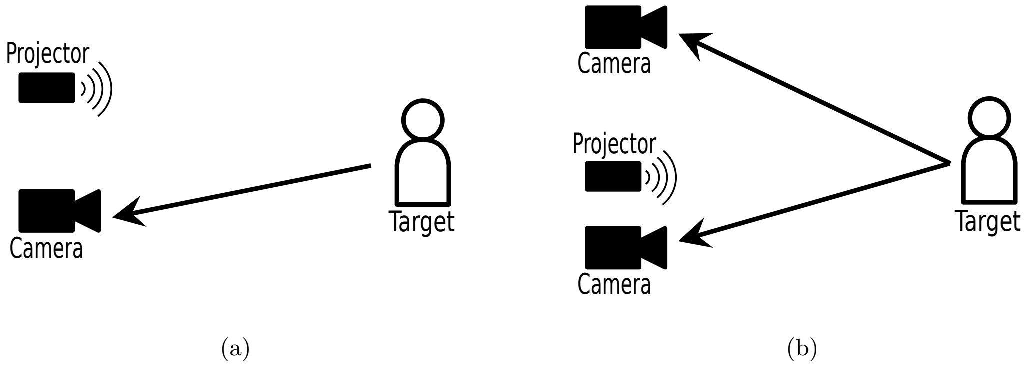
Where $\epsilon(\mathbf{d}, \mathbf{A})$ is the error function to minimize, $\mathbf{d}$ represents the image velocity (affine tracking) at $\mathbf{u}$, $\mathbf{A}$ is the affine transformation matrix at $\mathbf{u}$, and $I(\mathbf{x})$ and $J(\mathbf{x})$ represent two values (pixels) at the coordinates $\mathbf{x} = [x \ y]^T$. The integration region where the error function must be minimized (also known as the search region for a given feature) is defined by w_x and w_y.

To avoid the computational cost of performing optical flow tracking for the whole image (given a high resolution image), most of the recently presented algorithms make use of a region of interest (ROI) where the optical flow algorithm is applied to obtain the respiratory signal. An example of such implementation can be found in [121, 50], they use a user defined ROI to narrow the region to be processed and to increase the accuracy of the optical flow when acquiring respiration, by measuring the evolution of the thorax position through time. Other algorithms based on different signal and image processing techniques can also be found in the literature [129, 130, 131].

One of the main issues of the state of the art literature that uses video-based methods to obtain the respiratory signal, is that most of them are validated, or even designed to perform, in very controlled environments where the subject is lying on a bed [123] or with a very static posture [50]. This can pose as an issue if these algorithms want to be applied on real-life situations where the subject can freely move at any time. Other issues of the state of the art methods are the need to select a ROI to start acquiring the respiration, or even that most of them do not perform in real-time. Moreover, the main drawback of using video-based methods is that the illumination changes of the measured region can produce errors that completely invalidate the measurement. This type of issue is resolved by using IR cameras in external environments, or to perform the measurements in-door in order to have a constant source of light.

Depth Camera

Depth cameras have been slowly introduced into consumer electronics as well, an example is the Face-ID^{TM} from Apple, which uses a depth camera (embedded into different Apple products) to verify the identity of the subject. Another example of the evolution of these devices in the recent years is found in the Kinect^{TM} camera from Microsoft. If compared with the latest equivalent device from Intel, the RealSense^{TM} SR300, both cameras use an IR camera and a coherent near-infrared (NIR) projector (as depicted in Figure 2.14a) to compute the depth map. Another type of Depth camera is the RealSense^{TM} ZR300 (or its updated version D435i), that uses two IR cameras separated by a fixed distance and an IR projector (as depicted in Figure 2.14b).



(a) (b)

Figure 2.14: Schematic drawing of the two types of depth camera architecture: a) depicts a mono depth camera and b) depicts an stereo camera.

In the camera depicted in Figure 2.14a the depth map is computed as follows: the IR laser projector illuminates the target with a predefined pattern (vertical lines being the most common) and as the pattern illuminates the scene, and as the scene can contain multiple objects, the pattern gets distorted when hitting objects at different depths. The IR camera then captures the scene (with the distorted pattern) and the difference between the original and the distorted pattern is computed to obtain the depth map [132].

On the contrary, in stereo cameras (Figure 2.14b) the depth map is obtained by following a series of steps, first a comparison of the images from the left and the right camera is performed, thus obtaining a series of key features in both images. Then the algorithm computes the disparity between the same feature from both images, with this disparity and knowing the distance between the two cameras, the depth map can be computed [133].

Nowadays both the mono and stereo cameras have resolutions of at least 720 pixels and a framerate of at least 60 fps [134], thus making them a good choice to obtain physiological variables through the analysis of the depth map.

The physiological principle by which the respiratory signal is obtained from the depth map is the same as the one explained for the video camera, the difference in this case is, while the video camera sensed the X/Y displacement of the chest inside the image (2D), using a depth map (3D) allows to track the Z axis displacement, hence obtaining better resolution with less errors as the actual movement of the chest is in the Z axis. As an example, a very basic algorithm to obtain the respiratory signal from the depth map is to perform an average of the z vectors in the thoracic region, hence obtaining the mean displacement over time of the thorax.

Multiple methods have been presented over the years that use this kind of technology to obtain the respiratory signal [135, 136, 137, 138, 139, 140]. The common thing about the aforementioned methods is that make use of a Microsoft Kinect$^{\text{TM}}$ hardware, and although the results are good, the use of this hardware poses a limiting factor on the sample frequency and the resolution of the method, and needless to say the accuracy of the generated Depth map. Recent studies such as [16] make use of more advanced hardware (Intel RealSense$^{\text{TM}}$ SR300) which outperforms the Microsoft Kinect$^{\text{TM}}$.

The principal limitation of these studies is that they use a ROI that has to be predefined in order to obtain the movements in that region, if the subject moves, a new ROI would be needed. Other limitation is that all the studies where performed in very controlled environments where no movements (from the subject or external) or light changing conditions where taken into account, the later being of special relevance as all the aforementioned methods make use of mono cameras (either Kinect or SR300). This type of camera needs a perfectly illuminated scene in order to obtain the Depth map, in presence of an external IR source the depth map could contain errors making it not usable to detect any physiological variable.

2.4 Respiratory Signal Processing

All the methods and algorithms reviewed so far are designed to obtain the raw respiratory signal. In its raw stage, the respiratory signal does not provide much information about the subject apart from its tidal volume (amplitude of the signal). To extract relevant information from the respiratory signal, a series of pre-processing and processing techniques must be applied. One of the main drawbacks of these types of signals, is that almost no literature exists that provides specific signal processing methods to extract relevant information from the respiratory signals. One workaround to this issue, and as the respiratory signal is in fact a biological signal, is to apply the same methodologies described in the literature for other biopotential signals (ECG, electroencephalography (EEG)...).

A very common problem with non-contact methods is that the acquired signals are not as "clean", or noise free, as the ones obtained with contact methods. These errors can be attributed to involuntary movements or even due to the very nature of the acquisition method, as an example: video methods can have errors due to external light, where radar methods on the other hand, can have issues when demodulating due to null transmission or random body movements. All these issues/errors have to be corrected or mitigated before applying any post-processing technique, as otherwise the resultant information will certainly be corrupted or unusable. This pre-processing is normally based on filtering the signal to eliminate undesired components, high frequency noise in the case of the video methods or undesired harmonics in the case of radar. Filtering is also useful to eliminate the offset (base-line) of the acquired signals, as an example: in video based methods some base-line jumps can also appear due to sudden movements of the subject. In the literature multiple types of filter are found. The Butterworth filter is the one that is more commonly used in the literature, as this particular type presents a flat frequency response in the passband with the shortest transitory period. Also, depending on the wanted attenuation, the order of the filter can be increased/decreased without affecting the pass-band flatness.

Depending on the non-contact method, different cut-off frequencies are used, as the fundamental frequency of the respiratory signal is compressed between 0.18 Hz and 0.27 Hz [38], the typical cut-off frequencies that are found in the literature range from 0.05 Hz to 0.1 Hz for the high-pass frequency and from 1 Hz up to 15 Hz for the low-pass frequency. A good example of the different filters used in non-contact methods can be found in [50, 95, 121, 16].

After the pre-processing of the respiratory signal, the analysis is performed either in the time domain or in the frequency domain. Time domain methods are based on the extraction of the time between respiratory periods, also called breath to breath or respiration to respiration (RR). This time is expressed in seconds but as this metric is derived from the medical field, more precisely from the Pneumotachograph measurements, it can also be expressed in breaths per minute (BPM). Breaths per minute are computed as shown in Equation 2.6. From the breath to breath series the mean breaths per minute is extracted as well as the variability of the respiratory signal. Some examples of this type of time domain analysis can be found in [50, 121], this information is useful as it gives insight how the subject is breathing and the different events that take place during a test, for example apnoea periods or even rapid breathing due to stress [39].

$$BPM_i = \frac{60}{RR_i} \tag{2.6}$$

Frequency domain analysis on the other hand, makes use of the fast fourier transform (FFT), or other analogue methods, to obtain the spectral power density of the respiratory signal. Frequency domain methods also give information on how much variability the respiratory signal of the subject has had during a given period. This information is obtained by looking at the bandwidth of the signal, more bandwidth means more variability in the respiration of the subject. Other metrics to take into account when performing a frequency domain analysis are: the peak in the spectrum, as it coincides with the mean breathing rate, and the symmetry/ratio of the bandwidth before and after the peak, which determines if the subject has been breathing in a regular fashion or not. Examples of frequency domain analysis can be found in [121, 131, 136].

Overall the methods used to pre-process and post-process respiratory signals in the literature, are fundamentally the same that the ones used in any biological signal, and more precisely the ones found in ECG processing and heart rate variability analysis.

Chapter 3

Video-based pattern tracking for respiratory measurement

3.1 Introduction

As seen in the state of the art (chapter 2), multiple techniques can be applied to acquire the respiratory signal by the means of video analysis, which measures the displacement of the thoraco-abdominal wall due to the inhalation and exhalation processes. Regarding these methods, the ones based on video cameras are of special relevance for this chapter.

Most of the video camera-based methods found in the literature, need the interaction of the user in order to properly acquire the respiratory signal, this interaction is often required to determine a region of interest (ROI) in the area of the thorax or upper-torso. Without this ROI most of these algorithms could not extract the respiratory signal. Other methods, on the other hand, require for the user to adopt a static posture or even to lie on a bed.

Another challenge of these type of methods is that are often focused on obtaining the respiratory rate or the mean respiratory frequency, which limits the applicability of the methods in terms of respiratory rhythm extraction, or even to assess the variability of the respiratory signal. Moreover, there are only a few of them that are capable of acquiring the respiration of the subject in real-time.

In this chapter a new video-based method is presented, with the aim to verify the feasibility of using a consumer-grade camera to obtain the respiratory signal of a subject, and to improve the current state of the art in terms of respiratory rhythm measurement. For this reason, a new method has been presented based on tracking the movements of the thoracic cage by the means of a custom pattern placed on the thorax of the subject. Moreover, this method does not require any interaction with the user, and has been designed to perform in real-time.

The method presented in this chapter has been patented under the title *Respiratory Signal Extraction* with patent number WO2018/121861 [141]. Moreover, an article version of the present chapter has been submitted to the journal Biomedical Signal Processing and Control from Elsevier, currently being reviewed for publication.

3.2 Materials And Methods

The proposed algorithm uses distinctive patterns placed on top of the subject to obtain the respiratory signal. When the pattern is placed on top of the chest, by tracking the spacial evolution of the pattern inside the frame, the displacement of the thorax could also be obtained, hence the respiration of the subject [142]. In the next subsections, all the steps involved in detecting the pattern inside the frame and its posterior tracking is outlined.

3.2.1 Proposed Algorithm

A full diagram of the proposed algorithm is depicted in Figure 3.1. Three distinctive steps are defined in the block diagram, being the first the pattern detection, the second the feature tracking and the final step the signal extraction. The pattern detection can be further divided into multiple steps, which includes the feature extraction, feature matching and feature validation among others.

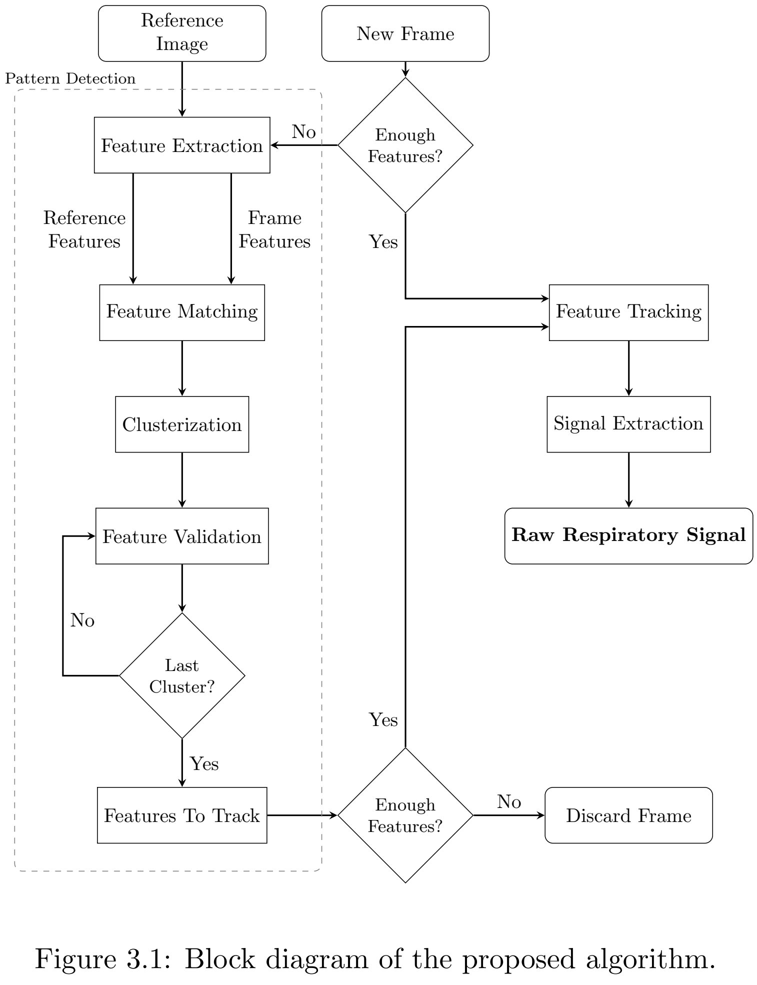


Figure 3.1: Block diagram of the proposed algorithm.

The proposed algorithm makes use of a distinctive pattern placed on top of the subject, and in order to detect this pattern inside the frame, a reference image is needed.

Reference Image

The pattern and reference image used by the algorithm, are composed of a series of vertical and horizontal lines combined with three squares on top and three squares on the bottom, with the figures drawn in black on top of a white background. This particular disposition was chosen to maximize the amount of corners (and possible features) that can be extracted, while maximizing the contrast between lines and background. An example of the designed pattern is shown in Figure 3.2a. While the generated pattern is designed to maximize contrast and corners, and as it is a computer generated image, when used as a reference image there is no ”texture” on it thus reducing the amount of features that can be extracted. To maximize the ”texture” and because the pattern placed on the subject (Figure 3.2c) is printed on paper, a picture of the printed pattern on top of a black background is used as the reference image (Figure 3.2b). This reference image has the same benefits (contrast and corners) of the original computer generated image, and the texture of the real printed pattern.

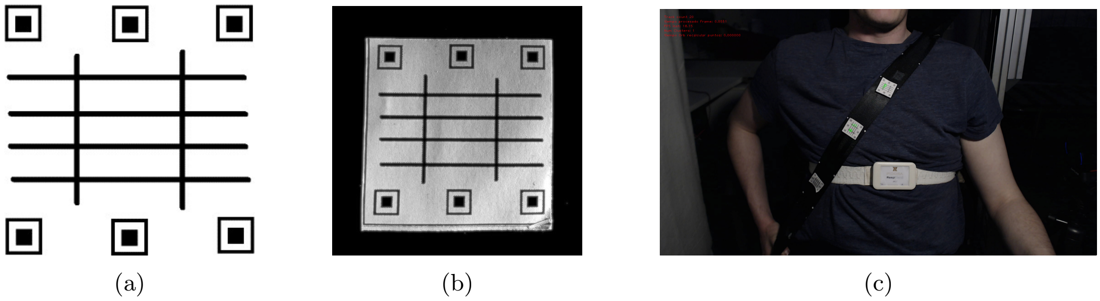

(a) (b) (c)

Figure 3.2: a) Depicts the generated pattern, b) shows the reference image using the generated pattern and finally,c) shows the patterns located on top of the seatbelt with a subject in the setup.

Pattern Detection

Once the pattern has been placed on top of the subject, the algorithm has to detect the pattern inside the frame in order to begin the tracking and the respiratory signal extraction. To perform the pattern detection a feature extraction, feature matching and feature validation among others, must be performed to acquire the location of the pattern inside the frame. All the processing steps involved are represented inside the dotted region in Figure 3.1.

Feature Extraction and Matching

Feature extraction from both the reference image and the frame from the camera is done by the means of the ORB algorithm [143]. ORB stands for Oriented FAST and Rotated Brief, which is an algorithm that allows the extraction of the most relevant features of an image, while being robust against rotations and illumination changes in comparison to other feature extractors such as Scale-Invariant Feature Transform (SIFT) [144]. The ORB algorithm has also been proven efficient in applications where low resources are available, and with proven computing efficiency in comparison with other methods, make the use of this feature extractor ideal for applications with real-time constraints. Using the ORB algorithm, a maximum of 1000 features where extracted from the frame, while a maximum of 400 features where extracted from the reference image. The difference in features is due to the difference in pixels between the reference image and the frame, while the frame has a resolution of 1080p the reference image only has 108x108 px, which justifies the lesser amount of features.

In order to match the computed features obtained from the frame and the reference image, a FLANN-based matching algorithm [145, 146] has been used. The Fast Library for Approximate Nearest Neighbours (FLANN) algorithm, uses the k-nearest neighbours [147] with an $l2$ norm to match each one of the features between each other. This step is crucial to form feature pairs between the reference image and the frame, this pairs will be latter used to track the evolution of the pattern inside the frame. The features of the frame that do not belong or match to the features of the reference image are automatically discarded.

Clusterization

As there it can be more than one pattern inside the frame, a clusterization algorithm has been used to detect which features belong to each pattern. The k-means algorithm [148, 149] has been used to group all the features in small clusters, as the features that belong to the same pattern have a higher probability to remain close to each other than features that belong to different patterns inside the frame. Finding the different feature clusters inside the frame allows for an individual tracking of each pattern, thus allowing multiple regions where the respiratory signal can be measured.

Feature Validation

To further validate the features inside each cluster, an homography algorithm [150] based on Random Sample Consensus (RANSAC) [147, 151] is used, as it finds true pairs among the features obtained from the ORB algorithm, taking into account changes of plane and distortions between the reference image and the frame. These algorithms are applied once for each cluster of features inside the frame.

Homography is defined as the projective transformation of the same feature that belong to the same plane between two images. The formal definition of homography, as defined by Vincent *et. al* in [150], is shown in Equation 3.1.

$$\begin{aligned} x'^{T}\mathbf{F}x &= 0 \\ \mathbf{x}' &= \mathbf{H}\mathbf{x} \\ \mathbf{x}' \times \mathbf{H}\mathbf{x} &= 0 \end{aligned} \tag{3.1}$$

where x'^T and x represent a pair of points inside two images $\mathbf{x}'$ and $\mathbf{x}$ respectively, $\mathbf{F}$ is the fundamental matrix that relates both points, and finally, $\mathbf{H}$ being the projective transform between both images, if the images belong to the same point world and plane.

In order to obtain true pairs of features, the basic RANSAC algorithm follows the steps as described by Vincent *et. al* in [150]:

- Variance normalized correlation is applied between features, whenever a pair has sufficiently high correlation is saved to form a set of candidate pairs.
- Four points are selected from the candidate pairs and the homography is computed (Equation 3.1)
- If the distance between $\mathbf{Hx}$ and $\mathbf{x}$ for a given set of pairs is below a certain threshold, the pairs are selected as valid.

Both the RANSAC and homography guarantee that all the obtained features are completely bijective between the reference image and each one of the clusters, removing any possible outlier that would have a very negative impact on the tracking stage and the respiratory signal extraction.

Finally, once all the features for each cluster have been validated using the previous algorithms, the remaining features are marked as features to track and passed onto the next section of the algorithm.

Feature Tracking

In order to track each feature, the pyramidal implementation of the Kanadae-Lucas-Tomasi (KLT) [125] optical flow algorithm has been used. Optical flow can be defined as the movement of a certain feature inside an image between two consecutive frames. A more formal definition of the algorithm can be found in chapter 2 (Equation 2.5). Given the definition of optical flow, two consecutive frames are needed in order to track the spacial evolution of each feature from one frame to the next. In the proposed algorithm these frames are: the previous frame from the camera and the current frame. In the initialization of the algorithm a copy of the first frame is used as a previous frame. To clarify the notation, and to illustrate how the optical flow is used in the proposed algorithm, the first frame to arrive will be named *PF* and the next frame of the camera will be renamed as *CF*.

The following steps describe the algorithm used to reduce the errors in the feature tracking, which is extracted from the *lk_track.py* OpenCV example [152]. To simplify the description of the algorithm, a supposition is made regarding that all the locations of the features inside *PF* are known.

- The first step is to update the location of each feature to the predicted location on *CF* by the means of the KLT algorithm. The aim of these steps is to obtain the estimate location from the features on *PF* to *CF*.

- To ensure that the updated locations are correct, the updated features on *CF* are used to predict its original location on *PF*. This step is used to prevent ”wandering” of the features, if the KLT algorithm updates a wrong location of a feature, this feature can wander through the image on consecutive iterations of the algorithm, thus producing aberrant results.
- The distances ($l2$ norm) between the original features and the ones obtained in the previous step are computed. If the distance between the original location and the predicted from the updated location onto the original image exceeds 1 px, the feature is automatically discarded. This distance was chosen empirically as it was the one that yielded better results for the proposed method.
- Finally, all the features that were not discarded in the previous steps are used in the following frames to continuously track the object.

One potential problem of the previous algorithm [152], is that eventually the number of features to track could decrease to an insufficient number due to the errors in the tracking. Whenever this is the case, and the number of features decrease below a certain threshold, the detection stage is triggered and new features are added to the tracking pool to ensure that the object is correctly tracked.

Another problem of the ”wandering” features, and the main reason why the previous algorithm is used, is that if a feature that has been ”wandering” throughout the image is used to obtain the raw respiratory signal, there is no guarantee that the feature is still located on a region where the respiration could be acquired (chest), thus multiple errors can be introduced into the signal. On the other hand, by applying the previous algorithm the location of each feature inside the region of interest (pattern) is guaranteed.

Signal Extraction

Once the features have been updated to its new location on the current frame, the respiratory signal can be extracted from each one of the patterns inside the frame. As the features have been clustered in previous steps, each cluster is processed individually to obtain a raw respiratory signal.

To obtain the evolution in the X-Y plane (image) of the patterns, a centroid is computed for each pattern (cluster) and by computing the distance between the centroid and the origin of the image (upper left corner), the spacial evolution of each pattern inside the image is obtained.

The centroid is computed by averaging the x and y coordinates of each feature inside the cluster (Equation 3.2), once the averaged x and y coordinates have been obtained, the $l2$ norm is used (Equation 3.3) to obtain the distance between the centroid and the origin of coordinates.

$$\begin{aligned} x_{avg} &= \frac{1}{N}\sum_{i=1}^{N} x_i \\ y_{avg} &= \frac{1}{N}\sum_{i=1}^{N} y_i \end{aligned} \tag{3.2}$$

$$Raw[k] = \sqrt{x_{avg}^2 + y_{avg}^2} \tag{3.3}$$

where N is the number of features, x_{avg} is the average of the x component for each feature and y_{avg} is the average of the y component of each feature. $Raw[k]$ represents the obtained sample for the given frame k.

Once a sample is obtained from a frame, the algorithm loops back to the beginning in order to obtain a new frame from the camera, and if there are enough features from the previous loop, the algorithm will compute the new location of each feature to obtain the next sample. If there is not enough features in any of the patterns inside the frame, the algorithm will perform a new pattern detection to extract new features for all the patterns inside the new frame.

A concatenation of all the samples obtained from each frame conform the raw respiratory signal. As there can be multiple patterns inside the frame, and each pattern will generate an individual signal, multiple regions of the subject can be monitored simultaneously. An example of the obtained raw respiratory signal from the central pattern is shown in Figure 3.3.

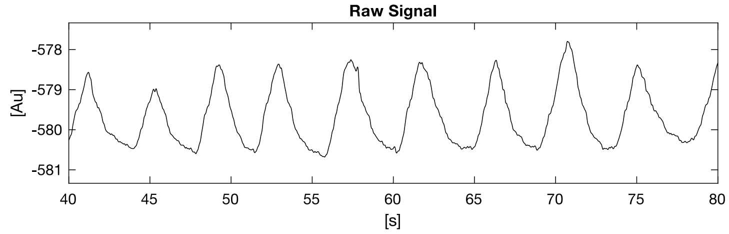


Figure 3.3: Fragment of the raw respiratory signal obtained with the proposed method.

3.2.2 Setup

The setup used was composed of a consumer-grade camera placed approximately at 70 cm of the subject. The subject was seated on a car chair in a comfortable position, and with the seat belt on. The used camera was a Logitech C920, configured at its native resolution of 1080p (Full HD) with MJPEG video codification. The camera was also configured to capture video at 15 fps, and although the camera can stream at 30 fps at this resolution, preliminary tests showed a systematic frame-rate drop which in return produced errors in the signal due to uneven sampling. The camera was also configured with automatic exposure disabled and blocked white balance in order to maintain the framerate as constant as possible. The FOV of the camera is 74.42° x 43.30° (H x V).

To maintain a constant illumination on the subject, an LED bulb from the Verbatim manufacturer (Ref: 52130), with warm white colour (CCT: 3000 K), with a luminous flux of 480 lm, a beam angle of 130 deg and a power consumption of 6.5 W was used. The light source was placed above the camera at an approximate distance of 70 cm from the subject and placed inside a parabolic light holder.

The reference system used to validate the method was the RespiBand$^{\text{TM}}$respiratory inductive plethysmograph (RIP) from BioSignalsPlux [94]. The system is comprised on a thoracic inductive band and a Bluetooth classic transmitter. The respiration of the subject is detected by the RIP sensor and transmitted to a computer via a Bluetooth classic serial port (SPP profile). The signal is sampled at 40 Hz with an ADC resolution of 12 bit, also the system has an embedded analogue band-pass filter with frequencies between 0.058 Hz and 0.9 Hz.

(a) (b)

Figure 3.4: a) Shows a schematic of the setup, b) depicts the actual setup. Figure b) extracted from: [29], Copyright Elsevier, all rights reserved.

Figure 3.4a shows the disposition of the camera inside the setup. On Figure 3.4b an image of the real setup is shown, were the light source and the camera can be appreciated.

Prior to any test, three patterns were placed on the seatbelt and inside the field of view of the camera. The placement of the patterns can be seen in Figure 3.2c. To ensure the same lighting conditions in all the tests and subjects, a pre-calibration of the white balance and exposure of the camera was conducted for each one of the subjects and reviewed for each test.

In order to make the study more reliable and repeatable, and although the algorithm is designed to perform in real-time, the video feed from the camera as well as the data from the RIP sensor were recorded using a laptop PC with a custom software based on the OpenCV Library (Version 3.0). All the signals were synchronized using the internal timestamp of the PC. Finally the laptop used to capture all the data had an intel i7-4710HQ processor, with an Nvidia GeForce GTX 850M graphics card and 8 GB of RAM.

3.2.3 Measurement Protocol

Twenty-one healthy subjects volunteered for the study, 11 of which were male and 10 female. Table 3.1 shows all the anthropometric data from the subjects which includes age, height and chest perimeter. Each subject gave their oral informed consent to freely participate in the study, and all the performed tests complied with the regulations of the Universitat Politècnica de Catalunya (UPC) and performed in accordance with the principles of the Declaration of Helsinki [153]. All the measurements were performed in a controlled environment with controlled lighting conditions.

Table 3.1: Anthropometric data from the subjects expressed as mean ± SD

Age [years]	Height [cm]	Chest Perimeter [cm]
26.6 ± 6.8	170.8 ± 7.4	88.4 ± 10.2

Prior to any measurement, every subject was asked to put on the RIP strap from the RespiBand system below the chest near the abdominal region, as it can be seen in Figure 3.2c. The subject was also asked to seat on the setup, and to fasten the seatbelt placed on the chair. All subjects were instructed to remain as still as possible during the tests.

Four tests were performed for each subject. Two of the tests involved breathing at a constant frequency, one consisted on breathing freely while the last test involved reading a text. The constant frequency tests were performed to test how the algorithm performs in the frequency domain, while the other two test were designed to test how the algorithm performs in real-life situations. The frequencies used in the constant frequency breathing test were 0.1 Hz and 0.3 Hz. To aid the subject to breath at these frequencies, a custom visual indicator that consisted on a moving bar was designed. To further ease the effort of the subject to breath at these frequencies, an asymmetrical breathing cycle was introduced, being 1/3 of the period for inhaling and 2/3 of its period to exhaling. Each test had a duration of 3 minutes, with a 30 second pause between tests. The total duration of the tests for each subject was approximately 15 minutes.

3.2.4 Signal Processing

The next subsection of this chapter describes all the signal processing steps used to normalise both respiratory signals, obtain the respiratory cycle series and finally how the error metrics are obtained.

Normalisation

Once the respiratory signals from both the reference method and the proposed method are obtained, a normalisation of the respiratory signal must be performed. This normalisation process is used to homogenise the sample rate between both methods and to filter undesired components in both signals. Another step in the normalisation process is the compression of both signals to a known scale, which allows a direct comparison between signals.

The signal processing steps taken to normalise both signals were the following:

1. The signals had to be interpolated at 80 Hz using a cubic spline, this step ensures that both signals have the same sampling frequency. Moreover in the case of the camera, the resampling compensates for any framerate loss that could result in an uneven sampling of the respiratory signal.

2. A 2nd order bidirectional Butterworth zero-phase bandpass filter, with cut-off frequencies between 0.05 Hz and 1 Hz was applied to both signals. This filter is designed to remove base-line drifts of the signal, specially in the case of the signal from the proposed method, and to remove undesired high frequency components.

3. In order to remove possible transitory periods in the respiratory signal obtained from the proposed method produced by rapid movements of the subject (that do not trigger the pattern detection stage), a moving median filter [154] was used. The filter had a window length of

three seconds. This window length had proven enough to smooth the signal, and at the same time shorter than a full respiratory period.

4. Finally, to compress both signals to a known amplitude comprised between -1 and 1, a non-linear function was applied. This function is based on the arctan properties [155] and it is described in Equation 3.4.

$$S_n[n] = \arctan \frac{S[n]}{\sqrt{\frac{\sum_{i=1}^{N}(S[i]-\bar{S})^2}{N-1}} \cdot \sqrt{2}} \tag{3.4}$$

where $S_n[n]$ is the resultant signal after the non-linear function (normalised respiratory signal), $S[n]$ is the filtered respiratory signal and $\bar{S}$ is the mean of $S[n]$.

Figure 3.5 shows a the normalised version of the signal depicted in Figure 3.3. It can be seen that the signal is now compressed between -1 and 1 and does not present any high frequency components that are present in the raw signal.

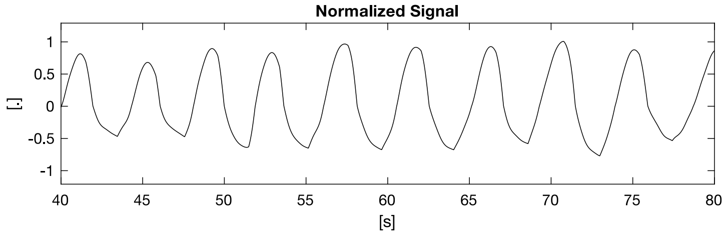


Figure 3.5: Normalized respiratory signal fragment of the signal in Figure 3.3

Respiratory Cycles Extraction

In order to compare both signals and to assess the error between them, the respiratory cycle (RC) series has been obtained from both signals. A respiratory cycle is defined as the time between two consecutive zero-crossing events with the same slope. The following steps were applied to both signals in order to obtain the respiratory cycles.

1. First, both respiratory signals had to be aligned and cropped so both signals contain approximately the same number of periods, but more important, to guarantee that both signals have the same number of samples. The signals were aligned using the Fisher ICC [156], iterating around one respiratory cycle to ensure a perfect alignment. Once the signals are aligned, the excess of each signal is cropped to ensure that both signals have the same length.

2. To avoid errors due to initial transitory periods on the signal obtained with the proposed algorithm, the first 10 seconds of both signals are removed.

3. The 65 percentile of each signal is computed to be used as a threshold. This value is used instead of the zero, as it yields better results with less errors produced by sudden slope changes of the signal.

4. The respiratory cycles are obtained by computing the time between positive slopes whenever the previous threshold is crossed.

5. The respiratory cycle series is build by concatenating all the obtained periods. To ensure that no erroneous periods are added to the series, only the cycles that are correctly detected on both signals are added.

Error Assessment

To further assess if the cycles detected on both respiratory signals have a correspondence between each other, or on the contrary there are cycles that are detected on one signal but not on the other, the sensitivity (SEN) and predictability of positive value (PPV) indicators have been computed. These indicators give information on how reliable the cycle detection is, and the validity of the detected cycles. In order to compute these two indicators, the number of correctly and incorrectly detected cycles must be assessed. To do so, the following confusion matrix has been computed.

1. true positive (TP): number of breath cycles that have been correctly detected in both the reference method and the proposed method.

2. false positive (FP): number of breath cycles that have been correctly detected in the reference method but not in the proposed method.

3. false negative (FN): number of breath cycles that have been correctly detected in the proposed method but not in the reference method.

In order to assess which cycles are labelled as TP, FP and FN, the following criteria has been taken into account based on the location of the detected cycles within both series: if the location of a given cycle from the proposed method, is within the margin of ± 33 % of the mean cycle length from the position of the corresponding cycle from the proposed method, the cycle is labelled as TP. If the location of detected cycle from the proposed method, is further than at least a 33 % of the mean cycle length from the location of the corresponding cycle from the reference method, this cycle is labelled as FN and discarded from both cycle series. Finally, if the location of the detected cycle from the proposed method, falls behind at least a 33 % of the mean cycle length from its corresponding cycle from the reference method, this cycle is labelled as FP and discarded from both cycle series.

The SEN indicator has been defined as ratio between TP and the total amount of cycles (TP + FN), and the PPV has been defined as the ratio between TP and the summation of TP and FP.

Once the cycle detection accuracy has been computed, both RC signals can be compared. To assess the error between RC series, a cycle to cycle comparison has been performed using the statistical methods shown in the following equations.

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

$$MAE_k = \frac{1}{N}\sum_{i=1}^{N}|e_k[i]| \tag{3.5}$$

$$MAPE_k = \frac{\frac{1}{N}\sum_{i=1}^{N}|e_k[i]|}{\frac{1}{N}\sum_{i=1}^{N}G_k[i]} \cdot 100 \tag{3.6}$$

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

$$SDE_k = \sqrt{\frac{\sum_{i=1}^{N}(e_k[i] - \bar{e}_k)^2}{N-1}} \tag{3.7}$$

where mean absolute error (MAE) (Equation 3.5) stands for mean absolute error, mean absolute percentage error (MAPE) (Equation 3.6) stands for mean absolute percentage of the error, and SDE (Equation 3.7) stands for standard deviation of the error.

In the aforementioned expressions, S represents the RC series obtained with the proposed method, while G represents the RC series obtained with he reference method, N is the total amount of breathing cycles per subject and k represents the subject being analysed.

Finally, to further compare the obtained cycles from both methods, and to analyse if there are any systematic errors in the cycle detection, a Bland-Altman (BA) [157] plot has been computed for each one of the tests. To compute the BA, no distinction between subjects has been made.

3.3 Results

In the next section, the results obtained from the analysis of the respiratory signals obtained with the proposed method and the reference method are shown.

3.3.1 Signals

The obtained respiratory signals for the proposed method and the reference method are shown in Figure 3.6. The Respiratory cycle series is also depicted in Figure 3.6. The signals from the proposed method are depicted with a solid line while the signals from the reference method are depicted with a dashed line.

As it can be seen, both respiratory signals are on top of each other, which indicates a high concordance between the obtained signal from the proposed method and the one from the reference method. The same can be said for the RC series, as it can be seen both signals are superimposed with practically overlapping between each other.

3.3.2 Performance

In this subsection the performance results of the proposed method are shown. Two types of performance are described, the computational cost and the statistical performance.

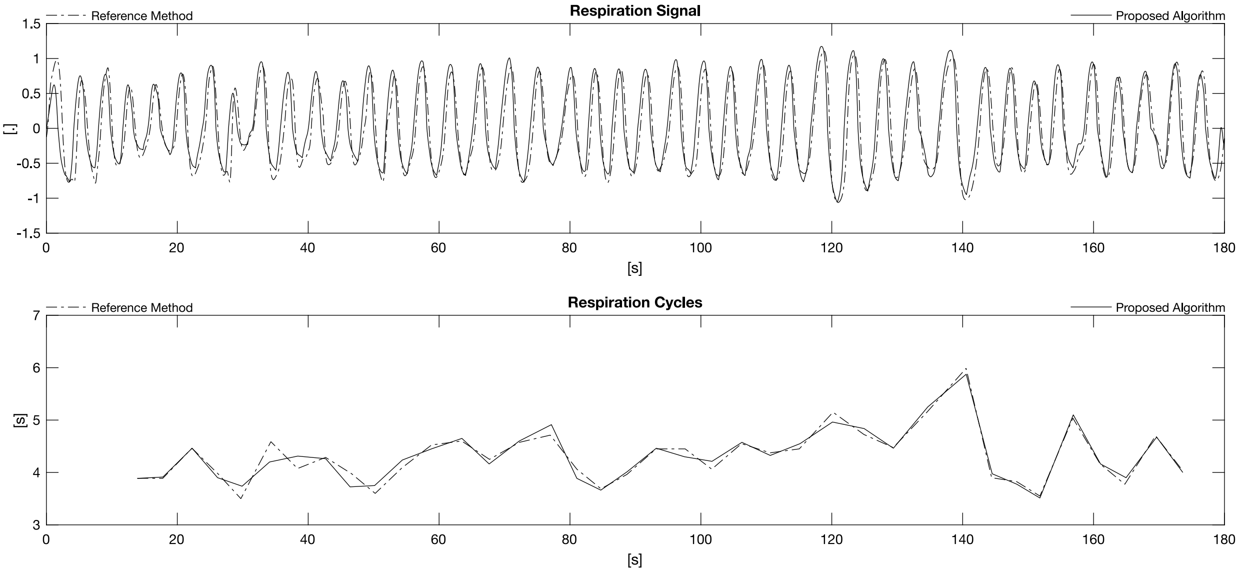


Figure 3.6: Comparison between the respiratory signals and respiratory cycles obtained with the reference method and the proposed method.

Computational Cost

In order to extract the computational cost of the algorithm, two different stages must be taken into account: the pattern detection and the combined stage of tracking and signal extraction. As the pattern detection is only executed when there is no available features, the computational cost of this stage can only be characterized when the code is executed, otherwise the cost is zero. On the other hand, for the feature tracking and respiratory signal extraction, the computational cost of the algorithm will be nearly constant and it will have a constant impact on the execution timings. The timing results for the pattern detection stage were 254.7 ± 2.8 ms, and for the combined feature tracking and respiratory signal extraction results were 12.4 ± 0.54 ms. As the performance bottleneck of the algorithm is this last combined stage, the results can be translated into a maximum of 80 frames per second at 1080p.

Error assessment

Three types of results are reported in this subsection, the results regarding the accuracy of the proposed method versus the reference method in the cycle detection, the results of the cycle by cycle comparison between methods and finally the BA plots and its associated results. To further clarify the presented results, a naming convention has been adopted for this section. The constant breathing tests have been renamed to "0.1 Hz" and "0.3 Hz" for its respective tests, the free breathing test has been renamed as "Free" and finally the test where the subject was asked to read a test has been renamed as "Reading".

In Table 3.2, the results regarding the accuracy of the cycle detection are shown.

Table 3.2: Performance indicators of the cycle detection, SEN and PPV

Test	TP	FP	FN	SEN [%]	PPV [%]
0.1 Hz	350	7	12	96.69	98.04
0.3 Hz	1040	5	13	98.77	99.52
Free	817	31	40	95.33	96.34
Reading	818	95	52	94.02	89.59

As it can be seen in Table 3.2 the SEN for all the tests is greater than 94 % being the test that yields the lowest value the Reading test. Regarding the PPV the constant breathing tests and the free breathing test show results greater than 95 % being the Reading test the only one that yields below 90 %.

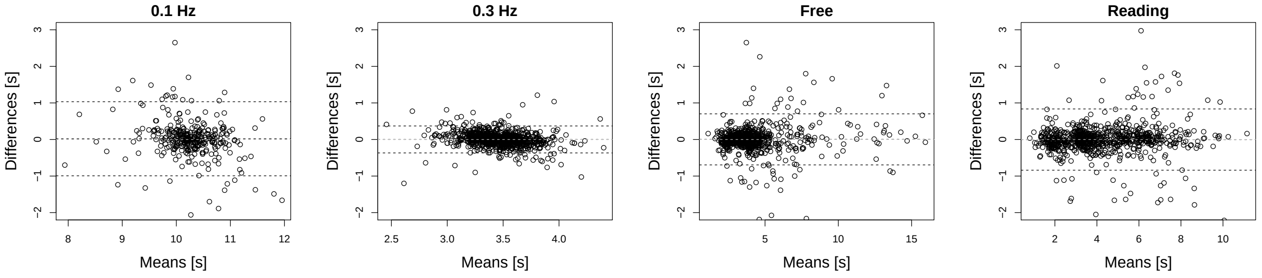
The next table (Table 3.3) presents the results for the cycle by cycle comparison between the proposed method and the reference method, along with the mean correlation between the cycle series of both methods for all the subjects.

Table 3.3: Correlation between methods, mean ± SD of MAE, MAPE and SDE.

Test	Corr. Signal (ICC)	MAE [s]	SDE [s]	MAPE [%]
0.1 Hz	0.945	0.331 ± 0.209	0.461 ± 0.289	3.319 ± 2.342
0.3 Hz	0.931	0.123 ± 0.046	0.179 ± 0.071	3.532 ± 1.286
Free	0.867	0.27 ± 0.282	0.359 ± 0.368	4.638 ± 2.691
Reading	0.850	0.275 ± 0.168	0.418 ± 0.274	5.627 ± 1.918

In Table 3.3, the correlation results for all the tests is greater than 0.85, being the correlations of the constant frequency breathing test greater than 0.93. Regarding the results for the error between methods, the MAE results show that the 0.3 Hz test is the one with lowest error being the 0.1 Hz the one with the greatest. The same can be applied to the SDE results being the 0.3 Hz test the one that yields lowest numbers. For the MAPE results the test that yielded the best accuracy was the 0.1 Hz test being the Reading test the one with lowest MAPE results. In order to assess the statistical interactions between tests, a paired t-test has been performed for all the possible combinations, being the only relevant result the case where the 0.3 Hz test was compared with the other tests. The t-test result for this case showed significant differences ($p < 0.05$) where the other results all showed non-significant differences.

Finally Figure 3.7 depicts four Bland-Altman (BA) plots, each one corresponding to one test. In the presented BA no distinction between subjects has been made, and all the cycles have been aggregated and compared in a cycle by cycle fashion. Regarding the BA representation, four dashed lines are depicted for each plot, the middle line depicts the mean of the Bland-Altman (BA) differences, while the two remaining dashed black lines represent the limits of agreement. The grey dashed line represents the zero-mean difference value inside the BA plot. The limits of agreement have been computed as the mean ± the 95% of the standard deviation of the differences.



(a) (b) (c) (d)

Figure 3.7: Bland-Altman of the computed periods for each frequency: a) 0.1 Hz, b) 0.3 Hz, c) Free and d) Reading.

It can be noted that all the BA plots are centred at the zero mean, with narrow limits of agreement. In Table 3.3, the differences for the four BA plots are shown. The four tests present almost near zero mean, being the 0.1 Hz test the one that presents the higher mean with 9 ms. Regarding the limits of agreement, the test that shows the greatest standard deviation is also the 0.1 Hz test. The 0.3 Hz test on the other hand is the one that presents the lowest mean and standard deviation, agreeing with the results shown in Table 3.2.

Table 3.4: Mean ± SD of the differences in the Bland-Altman plot in Figure 3.7

0.1 Hz [ms]	0.3 Hz [ms]	Free [ms]	Reading [ms]
9 ± 534	0± 189	1 ± 364	0 ± 431

3.4 Discussion

In this chapter a new method to extract the respiratory signal of a subject has been presented. The method makes use of a consumer-grade camera, and it is based on tracking the thoracic cage movements by the means of tracking a custom pattern inside the frame. The proposed method is able to perform at 80 Hz in its tracking stage, making its application suitable for real-life environments where real-time measurements are required. In Figure 3.6 an example of the respiratory signal obtained through the proposed method and the reference method is depicted. It can also be seen that both signals are practically superimposed, implying a high concordance between methods. Given the results shown in Table 3.3, it can be seen that the minimum correlation between methods is 0.85 for the reading test, with correlations higher than 0.93 for both the constant breathing tests. This difference between the uncontrolled tests (free and reading) and the constant breathing tests in terms of correlation, could be due to the involuntary movements of the subject in the uncontrolled tests, as well as the artefacts produced while reading out loud a text. This small artefacts reduce the correlation between the signal obtained with the proposed method and the reference method.

Previous to the error analysis, the accuracy of the respiratory cycle detection must be assessed. Regarding the cycle detection accuracy results shown in Table 3.2, it can be seen the difference in the obtained cycles depending on the test, for instance, the 0.1 Hz test only has 350 true positive (TP) periods, while the 0.3 Hz test has 1040 TP. As both the constant breathing frequency test have approximately the same number of false negative (FN) and false positive (FP), any misalignment between the RC series will produce a higher error in the 0.1 Hz test than in the 0.3 Hz as the later has more respiratory cycles. As for the free breathing and reading test, both have similar number of cycles as the subjects where breathing naturally without constraints.

In Table 3.2 the sensitivity (SEN) and predictability of positive value (PPV) results are also shown, being the lowest SEN and PPV the one for the Reading test. Taking into consideration the number of cycles, the SEN results for the 0.1 Hz test (96.60%) can be explained due to the low number in TP if compared with other tests. For the reading and free breathing tests, a respective SEN of 94.02 % and 95.33 % can be appreciated, the difference in results for this tests can be explained if it is taken into account that while reading, multiple artefacts appear in the respiratory signal, and while both tests have similar TP, the number of FP is greater in the reading test than in the free breathing test. Finally, for the 0.3 Hz test a SEN of 98.77 % can be seen, which is the best results due to the the high number of TP. Overall, and given te results in Table 3.2, a worst case scenario of more than a 90 % sensitivity is obtained for the proposed method. Given the PPV results, the worst case is the reading test with a 89.59 %, while the other cases show results greater than 95 %. This indicate a a very accurate cycle detection for the proposed method when compared with the reference method.

Relative to the error results shown in Table 3.3, for the MAE it can be seen that the 0.3 Hz test is the one that yields better results with the lowest mean and standard deviation. This can be explained as in this test the cycles ar shorter than on the other tests, hence the error in the estimation of the respiratory cycles is lower if compared with the other methods. On the contrary, the 0.1 Hz test is the one with the highest MAE mean and standard deviation. This can also be explained if it is taken into account that the 0.1 Hz test has less cycles than the other tests, as the cycles are longer. This can also be seen in Table 3.2 for the TP and FN between tests. For the free and reading tests, the MAE results for both tests are very similar between each other, being the free test the one with higher standard deviation. This can be explained if it is taken into account that both tests have approximately the same number of respiratory cycles.

For the SDE results, the 0.1 Hz test is the one with the highest error, being the 0.3 Hz the one with the lowest. This has the same interpretation as for the MAE results, as the cycles in the 0.1 Hz are shorter, any error in the signal will contribute more than in signals with shorter cycles. In general all the tests show a MAE and SDE results below 0.5 s, indicating a low error between the proposed method and the reference method.

Regarding the MAPE results, the 0.1 Hz and 0.3 Hz tests yield very similar results, with the 0.1 Hz test being the one with the smallest mean, implying a high accuracy in both methods. On the other hand, the free and reading test present a higher mean and standard deviation for the MAPE results, being the reading test the one with the highest values. This last two tests, as the nature of the respiratory signal is more turbulent, specially for the reading test, the MAPE results are worst than the ones presented for the constant breathing tests. It can also be noted in Table 3.3, that the MAPE results are worst whenever the correlation between signals is lower, and better results are reported when the correlation is high. This can be explained as the MAPE indicator is penalized by the same factors than the signal correlation. One important aspect to remark for the MAPE results, is that all the tests show a MAPE below 10 %.

Finally, for the BA results in Figure 3.7 and Table 3.4, a near zero mean can be observed for all the tests, being the highest mean 9 ms for the 0.1 Hz test. As for the constant breathing frequency tests, the 0.3 Hz test is the one with narrower limits of agreement and the one that presents a higher concentration of cycles within the limits. For the 0.1 Hz test, on the contrary, is the one with the wider limits of agreement (higher standard deviation), which can be explained if it is taken into account the low number of cycles in this particular test. The BA results for this two tests are in agreement with the MAE and SDE results shown previously, where the 0.3 Hz presents lower MAE and SDE than the 0.1 Hz test. Regarding the free and reading tests, both show a higher dispersion than the 0.3 Hz test, being the reading test the one that presents a higher standard deviation between the two of them. This increased standard deviation is due to the turbulent nature of the respiratory signal while reading. Moreover, no bias error can be found in any of the four tests, with all tests showing a high agreement between the proposed method and the reference method.

There where several limitations to this study, being the first one the number of recruited subjects for this study. Due to errors in the acquisition stage, produced by repeated occlusions of the pattern and excessive movement of the subject, or due to errors on the signal extraction stage, as a consequence of the pattern detection stage being triggered multiple times due to the aforementioned occlusions and movements, only 21 of the 23 measured subjects could be used. The second limitation to this study was that all the tests were performed in a laboratory setup, with a very controlled environment. For this reason, the proposed algorithm was not tested in light changing conditions or in the presence of external vibrations, and as a consequence, it cannot be assured that the proposed algorithm would have the same performance in real-life conditions. To assess this last limitations, the proposed algorithm will be tested against light changing environments and in the presence of external vibrations.

3.5 Conclusions

In this chapter a non-contact video-based method has been presented, with the aim to acquire the respiratory signal by the means of a consumer-grade camera and a pattern located on the chest of the subject. The proposed method consists on detecting a custom designed pattern inside the frame, and once its position has been obtained, the respiratory signal can be extracted by tracking its position throughout the scene. 21 subjects were measured and four test were performed under different breathing constraints. The reference method used was a commercial inductive plethysmograph.

The proposed method shows a high correlation (≥ 0.85) with the reference method, with high accuracy (SEN ≥ 94 %) in the respiratory cycle detection and low error (MAE < 0.34), indicating a good agreement between the proposed method and the reference method.

Chapter 4

Comparison of video methods for respiratory rhythm measurement

4.1 Introduction

In the previous chapter a new video-based method was introduced with the aim of obtaining the respiratory signal of a subject without contact. This method serves as a starting point to validate the feasibility of using video analysis to obtain the respiratory signal of a subject.

In chapter 2 section 2.3.2 multiple video methods have been disclosed, as they have been proposed in the literature to obtain the respiratory signal. Three major types of camera used for this purpose can be found, which includes depth cameras, RGB/IR cameras and thermographic cameras. Each one of these methods has its advantages and limitations, for example, RGB video cameras even though are cheap and practically found in every device, can have issues with the illumination of the scene or even sudden movements of the subject, which difficult the acquisition of physiological variables through this method. Depth cameras, on the other hand, have the advantage of being able to directly measure the distance between the subject and the camera, hence being able to directly obtain the thoraco-abdominal displacement of a subject. As a downside, this type of hardware is not as ubiquitous as an RGB/IR camera, which produces a dependence between the proposed algorithm and the type of camera used. For instance, if a mono depth camera is used (chapter 2 section 2.3.2), the computed depth map will have certain characteristics that will differ from the one obtained with stereo depth camera. Moreover, as a limitation of mono depth cameras, the scene needs to be perfectly illuminated by the IR projector, otherwise the depth map cannot be computed. Finally, thermal cameras can also be used to perform a direct measurement of the temperature changes at the nostril or mouth region, which is affected by the respiration, at a distance. Even though the hardware has improved in the recent years, and become more accessible, the required sensor resolution for this type measurement is scarcely available and expensive. Moreover, these type of cameras require an extensive calibration to compensate for the temperature drifts at the sensor.

All the aforementioned methods, have been used in the literature to obtain the respiration of a subject at a distance. Nevertheless, most of the methods found in the literature only focus on the mean respiratory frequency or in measuring the respiratory rate.

The aim of this chapter is to provide a comparison of the three aforementioned methods in terms of respiratory signal measurement and respiratory rhythm analysis, and to characterise the error between these methods and a reference method in terms of the instantaneous frequency of the respiratory signal. Moreover, in this comparison, for the RGB camera the algorithm proposed in the previous chapter is used to assert the performance of RGB-camera based methods against the other two devices, and to assess if the other two methods could be viable candidates for future implementations.

The contents and findings of this chapter have been published as a journal article [29] with the following reference:

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, 2019.
doi: 10.1016/j.bspc.2019.02.004

Available at the editor's web: https://linkinghub.elsevier.com/retrieve/pii/S1746809419300400.

4.2 Materials and Methods

4.2.1 Setup

(a) (b)

Figure 4.1: a) Depicts the position of the different cameras within the setup, b) depicts the real setup. Figure b) Extracted from: [29], Copyright Elsevier 2019, all rights reserved.

The setup was composed of a car seat, emulating a car cockpit, and three cameras that where located in front of the subject as depicted in Figure 4.1a, moreover, a light source was used to illuminate the scene. In Figure 4.1b an actual image of the cameras, the light source and the car seat used in the setup is shown.

To validate the measurements, an inductive plethysmography system was used as a reference method. The same commercial system used in chapter 3, the RespiBand RIP sensor from BioSignalsPlux^{TM} [94], was used in this study.

Table 4.1: Specifications of the used cameras

Type	Model	Resolution	FPS	FOV (H x V)
RGB	Logitech C920	1920x1080	15	70.42°x43.30°
Depth	Kinect V1.	640x480	10	58.5°x46.6°
Thermal	NEC InfReC G100	320x240	25	32°x24°

Table reproduced from [29], Copyright Elsevier 2019, all rights reserved.

Table 4.1 describes the models and characteristics of the three cameras used in this study. The three cameras used were a Logitech^{TM} C920 webcam, a Microsoft Kinect^{TM} depth camera and a Nec^{TM} InfReC G100 thermal camera.

The RGB camera was configured identically than in chapter 3 section 3.2.2, where the sensor was configured to record at 1080p, 15 fps and with the white balance and exposition set to a constant value. The framerate setting was chosen at 15 fps, as in preliminary studies where the camera was configured to record at 30 fps, a systematic loss of frames was observed.

The Microsoft Kinect^{TM} is comprised of two cameras, an RGB camera and a mono-depth camera. The depth camera has a build in IR projector that projects a pattern to the target, and an IR camera which registers the scene. By computing the differences between the projected pattern in the scene with the theoretical pattern (as described in chapter 2 section 2.3.2), the distance between the sensor and the measured object can be obtained. The depth camera returns a point-cloud matrix containing all this distances [158]. For this camera, both the RGB and Depth video feeds where obtained and recorded, but due to the low resolution of the RGB feed, this later one was not used in this study.

The thermal camera on the other hand, provides a grey-scale image representing the temperatures inside the FOV of the camera, as described in chapter 2 section 2.3.2. In this study the temperature range of the sensor was set between 28 °C to 38 °C), being the highest temperature represented by the white colour, and the lowest by the black colour. The used camera did not have a digital interface, and for this reason an analogue video recorder was used to acquire the composite video output of the camera.

Figure 4.1a is a representation of the position of the three cameras inside the setup. Figure 4.2, on the other hand, shows the view that the three cameras have of the subject, being each one positioned to maximize the amount of features of interest that are needed to obtain the respiratory signal. In Figure 4.2a the FOV of the depth camera is depicted where the thorax region can be appreciated, in Figure 4.2b the FOV of the thermal camera is shown where the nostril area can be seen, and in Figure 4.2c the FOV of the RGB camera is illustrated, where the thorax of the subject can be seen as well as the different patterns that will be used to obtain the respiratory signal. Finally, in Figure 4.2d a detailed picture of the used pattern is shown.

(a) (b) (c) (d)

Figure 4.2: a) Shows a frame from the depth camera, b) depicts the view from the thermal camera, c) shows a frame from the RGB camera and finally, d) shows the pattern placed on the subject. Source: [29], Copyright Elsevier 2019, all rights reserved.

A warm white LED bulb was used to illuminate the scene as a constant source of light. The light-source was the same used in Chapter 3 Section 3.2.2. The LED bulb had the following specifications, it was manufactured by Verbatim with reference number 52130, the light temperature was CCT: 3000 K, with 6.5 W of power consumption, a luminous flux of 480 lm and a beam angle of 130°. The temperature of the room was set to 24 °C ± 1 °C.

Prior to any test, a calibration of all the cameras was performed. The exposure and white balance of the RGB camera was fixed to the lighting conditions of the setup, while for the thermal camera, a warm-up period was required, as well as thermal bias correction to avoid drifts in the measured temperature over time.

Even though all the respiratory extraction algorithms were designed to perform in real-time, in order to make the study more reliable, all the video feeds and the RIP sensor where simultaneously recorded using the same PC, using a custom software based on the OpenCV Library (Version 3.0), being each video feed saved as a separate video file. The signals where synchronized upon processing using the timestamp of the PC.

4.2.2 Measurement Protocol

The same measurement protocol described in chapter 3 section 3.2.3 was used, where the same 21 subjects where measured. Table 4.2 summarizes the anthropometric data for all the subjects.

Table 4.2: Mean ± SD of the anthropometric data from all the subjects.

Age [years]	Height [cm]	Chest Perimeter [cm]
26.6 ± 6.8	170.8 ± 7.4	88.4 ± 10.2

The subjects were asked to perform four tests, in two of them the subject was asked to breath at a constant frequency (0.1 Hz and 0.3 Hz), in the third test the subject was asked to breath freely, and finally, in the fourth test the subject was asked to read out-loud a text.

4.2.3 Signal Extraction

Before any comparison between methods, the raw respiratory signal from each one of the video sources had to be extracted. For this reason two custom algorithms were developed and the method presented in Chapter 3 were used to acquire the respiratory signal from each source.

As each test had a duration of 180 s but each feed had a different sampling frequency (Table 4.1), a different number of frames were extracted for each video feed. 2700 frames, 1800 frames and 4500 frames were respectively extracted from the RGB camera, the depth camera and the thermal camera.

RGB Video

In order to obtain the raw respiratory signal from the RGB video, the method described in chapter 3 section 3.2.1 has been used. The timestamp from the video was also extracted in order to synchronize all the video feeds and the reference method.

Depth Point cloud

In order to process the depth point cloud (Figure 4.3a) to obtain the raw respiratory signal, the distance between the subject and the camera was identified, and a maximum and minimum distance was obtained. From the obtained distances comprised between the values 500 to 900 mm (Figure 4.3b), the evolution of the average distance between the subject and the camera was computed [141]. By computing this distance, the movement of the thorax of the subject can be obtained, hence the respiration of the subject [137]. By concatenating all the obtained distance for each frame, the respiratory signal can be obtained.

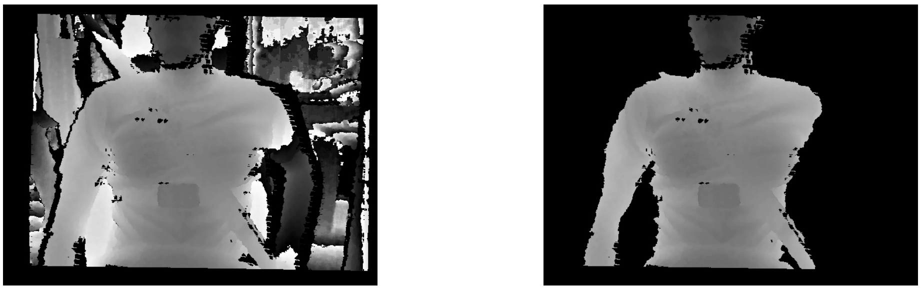

(a) (b)

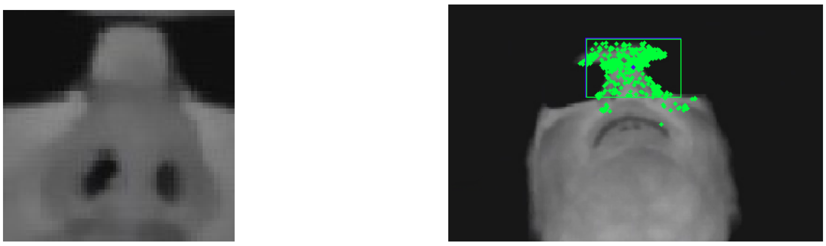
Figure 4.3: a) Point cloud obtained from the Kinect camera, b) cropped point cloud between the distances 500 and 900 mm.

Thermal Video

The respiratory signal from a thermographic camera is obtained by tracking the changes in the temperature at the nostrils. The respiratory signal extraction principle from the thermal video is explained in chapter 2 section 2.3.2. Three different steps were performed to obtain the temperature changes at the nostril region: the first step was to train a HAAR cascade classifier [159, 160], to automatically detect and track the nostril region within the frame. To train the classifier, one frame was extracted every 18 seconds from the 0.1 Hz breathing test, being a total of 10 frames extracted from each subject. After the frame extraction, the nostril region was manually selected and cropped to train the classifier. Figure 4.4a depicts an example of a selected nostril region.

Once the classifier was trained, the second step involved using the classifier to obtain for each subject and test, an approximate region containing the nostrils. To accurately detect the nostril region, a first generic portion of the image was manually selected and passed to the classifier, with the aim to detect the exact location of the nose region. Once the location was found, the features inside the nose region where obtained using the Shi-Tomasi "goodFeaturesToTrack" algorithm of the OpenCV library [161], as by using this features in conjunction of an KLT Optical flow algorithm, the movements of the nose can be tracked throughout the video, and the temperature at the nostril region can be obtained more accurately. Figure 4.4b shows an example of the detected nostril region (green rectangle) and the features obtained in the nostril region in order to be tracked by the means of the optical flow.

Finally, as described in the state of the art (section 2.3.2), by averaging the pixels inside the ROI the raw respiratory signal of the subject can be obtained.



(a) (b)

Figure 4.4: a) Detail of a nostril region manually selected from a given frame, b) detected nostril region (green rectangle), with the obtained features (green crosses) used for the optical flow tracking.

4.2.4 Signal Processing

In order to compare the respiratory signals obtained by the different video methods and the one obtained from the reference method, the raw respiratory signals had to be normalized. The processing steps to perform the normalization of each one of the respiratory signals are the same as described in chapter 3 section 3.2.4. All the processing steps were performed using the Matlab software.

The normalization process can be summarized using the following steps.

- The signals were interpolated at 40 Hz using a cubic spline.

- A second order bidirectional Butterworth band-pass filter between 0.05 Hz and 1 Hz was used.
- The result of a moving median filter [154] was subtracted to the signal to remove the transitory periods produced by the rapid movements of the subject.
- Finally, to compress the signal between -1 and 1, a non-linear function based on the arctangent [155] Equation 4.1 was applied.

$$S_n[n] = \arctan \frac{S[n]}{\sqrt{\frac{\sum_{i=1}^{N}(S[i]-\bar{S})^2}{N-1}} * \sqrt{2}} \tag{4.1}$$

where $S_n[n]$ is the resultant signal after the non-linear function (normalised respiratory signal), $S[n]$ is the filtered respiratory signal and $\bar{S}$ is the mean of $S[n]$.

Figure 4.5 shows an example of a raw signal from the reference method and its equivalent normalized signal. where the effect of the non-linear compression can be appreciated.

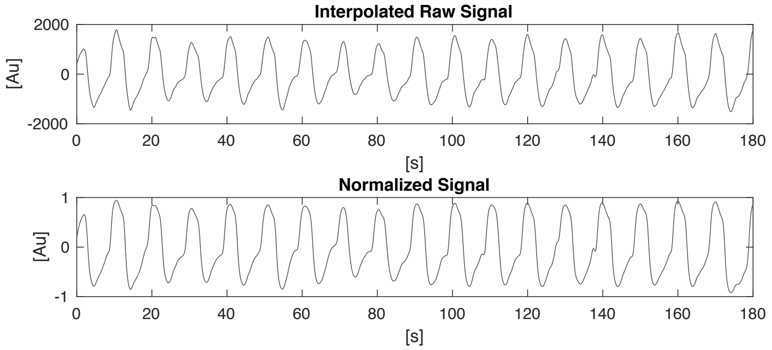


Figure 4.5: The top row depicts the interpolated raw signal, while the bottom row depicts the normalized signal. Both signals have been obtained from the 0.1 Hz test. Source: [29], Copyright Elsevier 2019, all rights reserved

Hilbert Instantaneous Frequency

Once the respiratory signals for all methods have been normalized, in order to characterize the error between them, the Hilbert IF has been computed for all methods. The IF is a time-frequency representation where the frequency evolution in time of a signal is shown. This approach has been taken to accurately compare all the obtained signals from the different methods, moreover, time-frequency representation has been traditionally used to extract respiration from ECG [162, 163, 164] or even in baseline wander removal [165, 166]. Furthermore, this technique can also be used to analyse respiratory signals.

The Hilbert IF is defined as the first derivative of the instantaneous phase of the respiratory signal [167]. The computation of the Hilbert IF is possible only if the signal is periodic, and as the respiratory signal can be approximated as a quasi-periodic signal, the IF can be computed. Moreover, to remove higher frequency harmonics from the signal, a narrow-band pass-band filter is applied, thus allowing to express the respiratory signal by the means of the analytic signal.

The first step to obtain the IF signal for all the methods, is to compute the Hilbert transform from the normalized respiratory signal (equation 4.2). Once the Hilbert transform is computed, the instantaneous frequency is computed using the following equations:

$$H_s[n] = Hilbert(S_n[n]) \tag{4.2}$$

$$\phi[n] = \arg\left(H_s[n]\right) \tag{4.3}$$

$$\phi_n[n] = \begin{cases} \phi[n], & \text{if } |\phi[n]| < \pi \\ \phi[n] \pm 2\pi, & \text{if } |\phi[n]| \geq \pi \end{cases} \tag{4.4}$$

$$\omega[n] = \phi_n[n] - \phi_n[n-1] \tag{4.5}$$

$$IF[n] = \omega[n]\frac{Fs}{2\pi} \tag{4.6}$$

where $H_s[n]$ is the Hilbert transform of the normalized respiratory signal $S_n[n]$, $\phi[n]$ is the argument of the Hilbert transform, $\omega[n]$ is the instantaneous frequency in radians, Fs is the sampling frequency and finally, $IF[n]$ represents the instantaneous frequency in Hz. All the operations are performed in the discrete-time domain, where n represents a certain sample.

Equations 4.3 and 4.4 are used to obtain the argument of the Hilbert transform and unwrap the phase in 2π increments to ensure continuity. This is the first step to obtain the IF from the Hilbert transform. Equation 4.5 computes the difference of adjacent angles, this operation is used to obtain the discrete first derivative of the phase thus obtaining the instantaneous frequency in radians. Finally Equation 4.6 is used to convert the previous result to Hz.

To ensure that no abrupt peaks are present in the computed IF due to the phase correction, a Hodrick-Prescott [168] filter is applied to the IF with a smoothing factor of $2 * 10^6$.

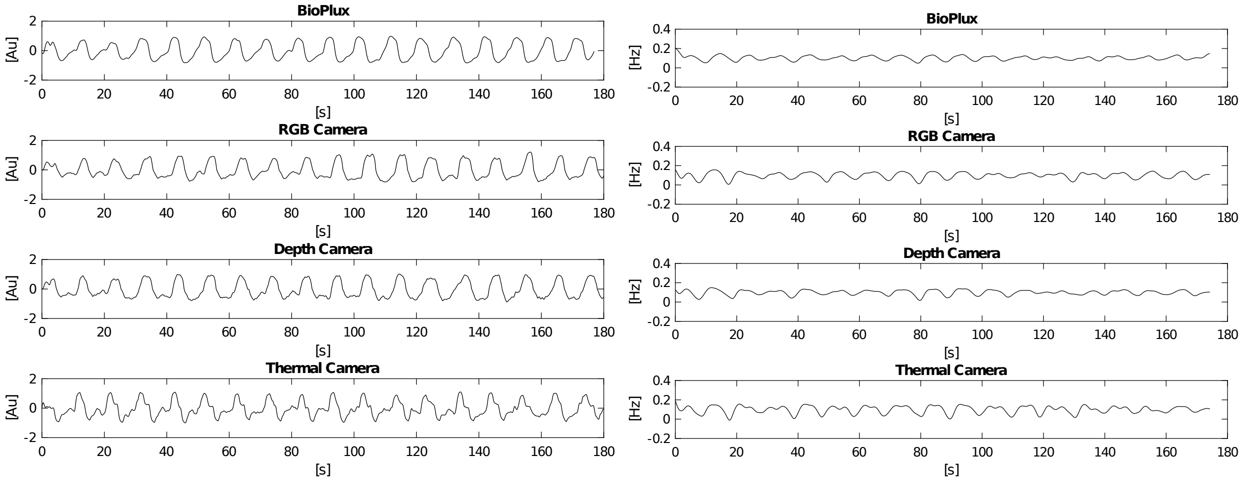
Finally, to ensure that all the IF from all the methods are properly aligned with each other, the IF signals have been aligned by maximizing the Fisher's [156] ICC by iterating each signal around one respiratory period. Once all the signals were aligned, the signals were cropped so each signal had the same length.

4.2.5 Error Characterisation

The standard deviation of the error (SDE) has been used to characterise the error between the IF from the video methods and the IF of the reference method. The SDE was computed using the error between methods on a sample by sample basis.

4.3 Results

4.3.1 Signals



(a) (b)

Figure 4.6: a) Depicts the normalised respiratory signals from each one of the studied sources, b) depicts respectively, the instantaneous frequency obtained from each one of the respiratory signals in (a). Source: [29], Copyright Elsevier 2019, all rights reserved.

Figure 4.6 shows an example of both the normalized signals for each method and the obtained IF side by side. Figure 4.6a depicts an example of the resulting normalized signals for the video methods and the reference method, corresponding to the 0.1 Hz test. Figure 4.6b, on the other hand, shows the obtained IF for each one of the signals represented in Figure 4.6b. It can also be noted in Figure 4.6b that the mean value of the obtained IF, as the test that is shown corresponds to the 0.1 Hz constant breathing test, is approximately 0.1 Hz for all the methods.

4.3.2 Performance

In order to improve the readability of the results, the names of the tests and methods have been modified. In this regard, the signals from the reference method are renamed as ”Plux”, the signals from the RGB camera are renamed as ”Pattern”, the depth camera metod as ”Depth”, and finally, the Thermal camera signals are renamed as ”Thermal”. Also, the tests have been renamed in the following way: the constant breathing tests have been renamed as ”0.1 Hz” and ”0.3 Hz” respectively, the free breathing test as ”Free” and the Reading test has been renamed as ”Reading”.

Table 4.3 presents the mean ± standard deviation (SD) of the mean IF for all the tests and methods. It can be appreciated that for the constant breathing tests, the mean ± SD for the reference method is 0.1 ± 0.022 Hz for the 0.1 Hz test and 0.293 ± 0.017 Hz for the 0.3 Hz test. As the subjects were instructed to breath at a constant frequency, the SD for these two tests is small, while the mean is practically the target frequency for each test. On the other hand, the Free and Reading tests yield similar mean results of 0.242 Hz and 0.234 Hz respectively, as in both tests each subject was breathing at its own natural frequency, the mean frequency is in concordance with the range of 14 to 20 breaths per minute at rest [38] of an adult subject.

Table 4.3: Mean ± SD of the instantaneous frequency for all methods.

	Plux [Hz]	Depth [Hz]	Pattern [Hz]	Thermal [Hz]
0.1 Hz	0.100 ± 0.022	0.098 ± 0.029	0.098 ± 0.036	0.169 ± 0.093
0.3 Hz	0.293 ± 0.017	0.289 ± 0.018	0.289 ± 0.018	0.267 ± 0.072
Free	0.242 ± 0.053	0.245 ± 0.065	0.236 ± 0.059	0.227 ± 0.094
Reading	0.234 ± 0.09	0.233 ± 0.095	0.237 ± 0.098	0.234 ± 0.125

Table reproduced from [29], Copyright Elsevier 2019, all rights reserved.

Bland-Altman Analysis

Two Bland-Altman (BA) plots have been obtained in order to quantify the agreement between the different video methods and the reference method, as well as a two statistical t-test analysis for each BA plot. The first t-test analysis consisted on comparing the differences in the BA to the null-bias hypothesis, while the second t-test was performed to compare if the BA differences for each method where statistically independent between each other. All the statistical tests performed were computed using the R software.

Prior to the Bland-Altman (BA) analysis, a series of outliers had to be removed using the following criteria, that if the differences in the mean IF, for a given subject and test, between methods exceeded two times the standard deviation of the whole test, then the given subject was deemed as an outlier and discarded from that test. The number of removed outliers was the following: four subjects were removed from the 0.1 Hz test, three subjects from the 0.3 Hz test, five subjects from the Free test and finally three subjects were removed from the Reading test, from a total of 21 measurements each. To test if the removal of outliers from the BA analysis could alter the statistical results, a random removal of subjects was performed to the BA for each test, which yielded no significant differences in the statistical results. The removal of these outliers was necessary to compensate for the errors performed in the measurement stage due to two primary sources of errors: the consecutive loss in frames during acquisition and due to the excessive movement of certain subjects which had a negative impact in the obtention of the respiratory signal and by consequence the estimation of the IF.

Figure 4.7 shows the BA analysis between the reference method and the different video methods studied in this chapter of their mean IF. A total of 12 BA are shown, where the columns represent each one of the four performed tests, and the rows represent each one of the video methods being compared. Figure 4.8 on the other hand, represents the BA analysis of the standard deviation of the IF for each test and method, distributed in the same fashion as Figure 4.7.

In each BA plot, four lines are shown, the first line, depicted as dashed grey, represents the zero-mean value of the differences, being the three remaining dashed lines: the mean value of the differences (central line), and the limits of agreement of each BA plot. This last two lines are computed as the mean value plus/minus 1.96 times the standard deviation of the differences. By observing the distance between the grey dashed line (zero-mean) to the central line, the bias of the plot can be assessed, as well as to assess if the zero-mean value is within the limits of agreement.

Table 4.4 presents the mean and standard deviation of the differences for each one of the BA plots in Figure 4.7, as well as the t-test results for the null-bias hypothesis and the comparison between BA plots. Table 4.5 shows the mean and standard deviation of the differences and the t-tests results, for the BA in Figure 4.8.

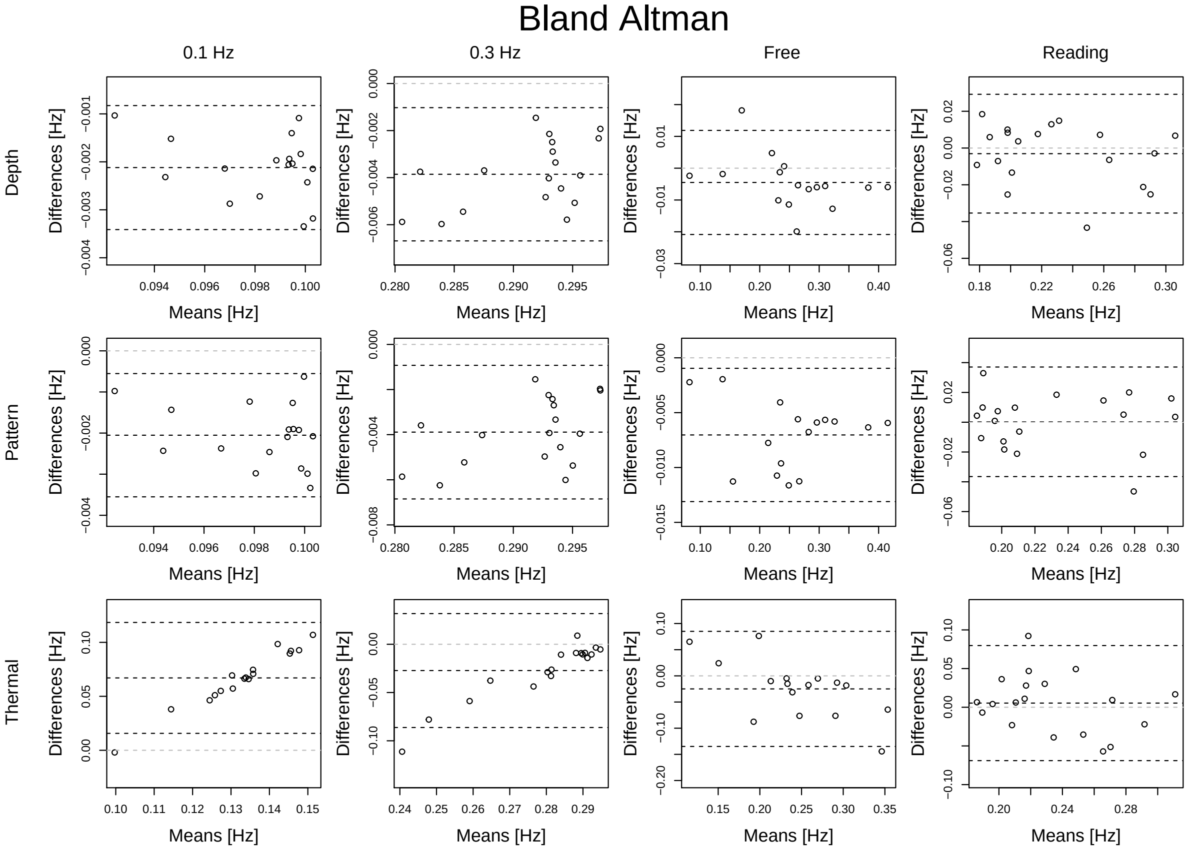


Figure 4.7: Bland-Altman plots of the mean IF, where the rows correspond to the method being compared, and the columns to each test. Source: [29], Copyright Elsevier 2019, all rights reserved.

Table 4.4: Numerical results of the differences from the BA plots in Figure 4.7.

	Depth [mHz]	Pattern [mHz]	Thermal [mHz]	Significance
0.1 Hz	-2.119 ± 0.660	-2.052 ± 0.765	67.045 ± 26.213	$**; \#\#; \vee\vee; NS_\dagger; \ddagger\ddagger; \diamond\diamond$
0.3 Hz	-3.857 ± 1.444	-3.888 ± 1.511	-27.247 ± 30.109	$**; \#\#; \vee; NS_\dagger; \ddagger\ddagger; \diamond$
Free	-4.483 ± 8.348	-7.032 ± 3.100	-24.933 ± 56.106	$*; \#\#; NS_\vee; NS_\dagger; NS_\ddagger; NS_\diamond$
Reading	-3.067 ± 16.482	0.275 ± 18.761	5.366 ± 37.945	NS for all tests

Significance is represented as: ”NS_*” for $p > 0.05$, ”$*$” for $p < 0.05$ and ”$**$” for $p < 0.001$. For the null bias t-test: ”$*$” corresponds th the Depth method, ”#” corresponds to the Pattern method and ”$\vee$” corresponds to the Thermal method. For the t-test between methods: ”†” corresponds to the Depth vs Pattern, ”‡” corresponds to the Depth vs Thermal and finally ”$\diamond$” corresponds to the Pattern vs Thermal. Table reproduced from [29], Copyright Elsevier 2019, all rights reserved.

The results in Table 4.4 for the Depth and Pattern methods, show very similar mean and standard deviation values for the 0.1 Hz and 0.3 Hz test, being the mean in the range of mHz and with practically no standard deviation. Regarding the significance of the null-bias hypothesis, both test for both methods present very significant differences ($p < 0.001$). Regarding te results of the Free test, an increased mean and standard deviation in respect from the constant breathing test can be appreciated for both methods. The Depth method, in one hand, presents significant differences ($p < 0.05$) for the null-bias hypothesis, while the pattern method presents very significant differences ($p < 0.001$). For the Reading test results, both methods present an increased standard deviation, up to two times greater than in the Free test for the Depth method and up to six times greater for the Pattern method, being the Depth method the one that presents the highest mean. For this particular test, both methods do not present significant differences for the null-bias hypothesis.

The t-test comparing the BA results for the Depth method and the Pattern method does not yield significant differences.

Given the results for the Thermal method in Table 4.4, it can be observed that either the mean and standard deviation for the constant and Free breathing tests present higher values than the respective tests for the Depth and Pattern method. The Reading test on the other hand, presents similar values than te ones in the Pattern and Depth methods with an slightly increased standard deviation. It is important to remark that for the 0.1 Hz test results, the mean value is almost one order of magnitude greater than the mean value for the same test of the other two methods. The null-bias hypothesis statistical analysis results, show very significant differences for the 0.1 Hz test ($p < 0.001$), and significant differences ($p < 0.05$) for the 0.3 Hz test. The Free and Reading tests do not present significant differences for the null-bias hypothesis.

The results for the t-test analysis between BA show very significant differences between the Thermal method and the other two methods for the 0.1 Hz test. For the 0.3 Hz test, there are very significant differences between the Depth and the Thermal methods, and significant differences between the Pattern and the Thermal methods. For the Free and Reading tests, no significant differences can be appreciated between the Thermal method and the other two methods.

The BA representation in Figure 4.8 analyses the standard deviation of the IF between the studied methods and the reference method. The relevance of Figure 4.8 lies on the study of the variability of te IF between methods, hence the frequency range of the reference method versus the frequency range of the studied methods.

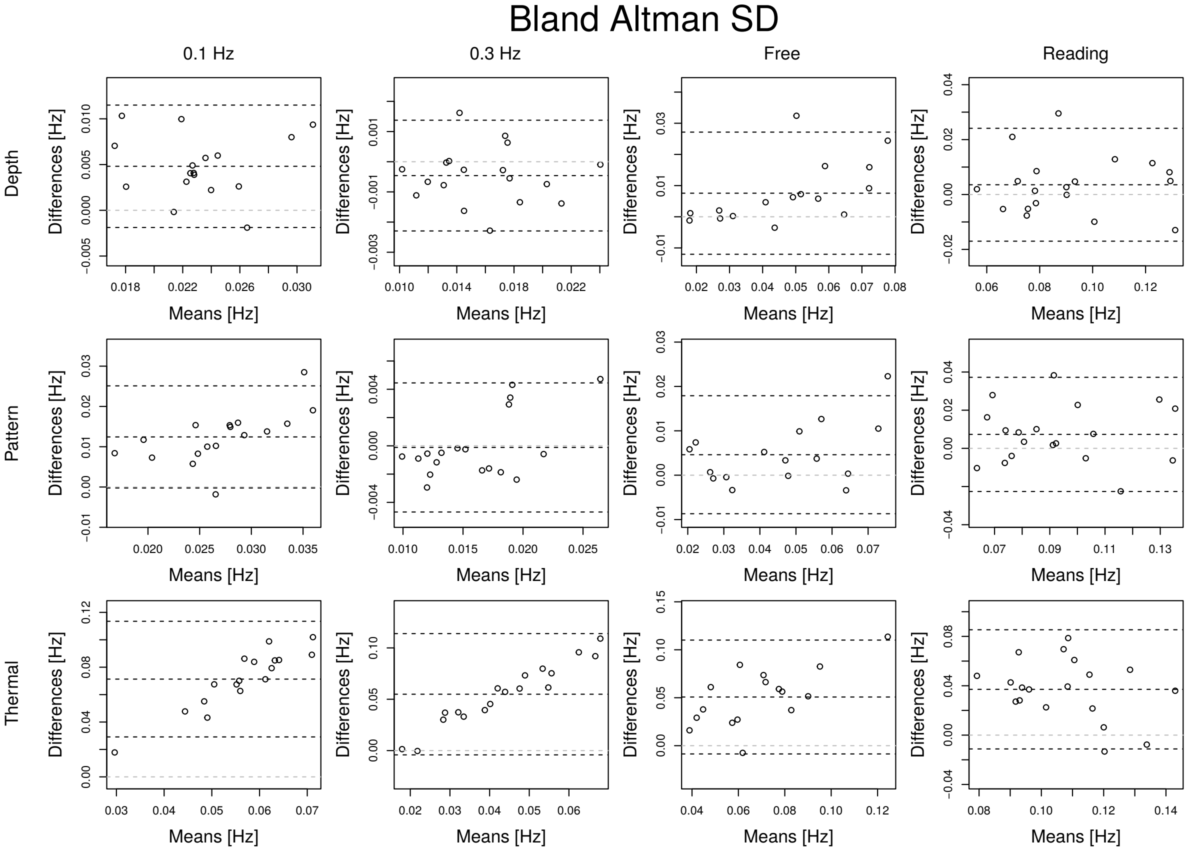


Figure 4.8: Bland-Altman plots of the Standard Deviation IF, where the rows correspond to the method being compared, and the columns to each test. Source: [29], Copyright Elsevier 2019, all rights reserved.

Regarding the results in Table 4.5, the Depth and Pattern methods show similar results in all the tests, being the greatest difference in the 0.1 Hz test, where the mean value for the Pattern method is three times greater tan the one for the Depth method. As for the Thermal method, both mean and standard deviation presents values that are almost an order of magnitude greater than the ones for the other two methods for all tests. As for the statistical results of the null-bias hypothesis, the Depth method presents very significant differences ($p < 0.001$) for the 0.1 Hz test and significant differences ($p < 0.05$) for the Free test, while no significant differences are found for the 0.3 Hz and Reading test. Regarding the Pattern method, the same statistical results are obtained. The

Table 4.5: Numerical results of the differences from the BA plots in Figure 4.8

	Depth [mHz]	Pattern [mHz]	Thermal [mHz]	Significance
0.1 Hz	4.81 ± 3.413	12.432 ± 6.469	71.308 ± 21.521	$**; \#\#; \vee\vee; \dagger\dagger; \ddagger\ddagger; \diamond\diamond$
0.3 Hz	-0.46 ± 0.936	-0.115 ± 2.329	54.887 ± 30.135	$NS_*; NS_\#; \vee\vee; NS_\dagger; \ddagger\ddagger; \diamond\diamond$
Free	7.562 ± 9.997	4.612 ± 6.783	50.753 ± 30.282	$*; \#; \vee\vee; NS_\dagger; \ddagger\ddagger; \diamond\diamond$
Reading	3.554 ± 10.488	7.314 ± 15.262	37.079 ± 24.641	$NS_*; NS_\#; \vee\vee; NS_\dagger; \ddagger\ddagger; \diamond\diamond$

The significance results are represented as in Table 4.4. Table reproduced from [29] 2019, Copyright Elsevier, all rights reserved.

Thermal method on the other hand, shows very significant differences ($p < 0.001$) for all the test. The statistical results for the comparison between tests show very significant differences ($p < 0.001$) between the Depth and Pattern method for the 0.1 Hz test, with no significant differences on the other tests. The statistical results for the comparison between the Thermal method and the other two methods, show very significant differences ($p < 0.001$) in all the tests.

4.3.3 Error Characterisation

Once the SDE for all methods was computed, a Shapiro-Wilk test [169] was performed to each SDE values grouped by methods. The null-hypothesis for this test was to determine if the SDEs could be approximated by a normal distribution. The test results show that two of the three error groups (Plux vs Depth and Plux vs Pattern) do not show a normal distribution. For this reason The SDE values are reported as median and interquartile range, and a the Friedman non-parametric test [170] was used to compare the different methods and tests. Finally, a Wilcoxon-Nemenyi-McDonald-Thompson [170] post-hoc test was also applied to characterise the relationships between the error groups.

Table 4.6 shows the results for the standard deviation of the error, where similar results can be seen for the Pattern and Depth methods for all tests. Regarding the Thermal method, the SDE results are four times and five times greater than the ones for the other two methods in the 0.1 Hz and 0.3 Hz test respectively. For the Free test, the SDE is two times greater than in the other two tests, while for the Reading test the results are similar with the other methods.

For the post-hoc test, the interaction between SDE Depth and SDE Pattern shows no significant differences in any test, while the interaction between SDE Depth and SDE Thermal, and SDE Pattern and SDE Thermal shows significant ($p < 0.05$) and very significant differences ($p < 0.001$) in all the tests.

Table 4.6: Results of the SDE expressed as median and IQR [25 ; 75] and post-hoc analysis of the interactions between methods.

	Depth [Hz]	Pattern [Hz]	Thermal [Hz]	Significance
0.1 Hz	0.018 [0.016 - 0.019]	0.024 [0.019 - 0.027]	0.096 [0.084 - 0.102]	$NS_*; \#\#; \vee\vee$
0.3 Hz	0.011 [0.008 - 0.014]	0.013 [0.009 - 0.018]	0.074 [0.049 - 0.092]	$NS_*; \#\#; \vee\vee$
Free	0.038 [0.014 - 0.052]	0.021 [0.015 - 0.034]	0.106 [0.073 - 0.114]	$NS_*; \#; \vee\vee$
Reading	0.091 [0.073 - 0.117]	0.096 [0.079 - 0.126]	0.14 [0.131 - 0.155]	$NS_*\#\#; \vee$

Significance is represented as: "NS_*" for $p > 0.05$, "$*$" for $p < 0.05$ and "$**$" for $p < 0.001$. The interactions are represented as follows: "$*$" for the Plux vs Depth and Plux vs Pattern, "#" for the Plux vs Depth and Plux vs Pattern, and finally, "$\vee$" for the Plux vs Pattern and Plux vs Thermal. Table reproduced from [29], Copyright Elsevier 2019, all rights reserved.

4.4 Discussion

Given the results of the BA plots in Figure 4.7 and in Table 4.4, it can be seen that for the Depth and the Pattern methods, the constant breathing tests exhibit very narrow limits of agreement, with low values either in the mean and the standard deviation. This indicates a high agreement between the studied methods and the reference method for this particular tests. Nevertheless, in both methods a minimal bias can be appreciated, even though the values are very low if compared to the fundamental frequency in both tests. Moreover, the statistical results indicate non-significant differences between methods thus indicating a high agreement with the reference method. Regarding the t-test results for this two particular tests and methods, the results indicate very significant differences between the studied method and the reference method. This can be explained as it is taken into account the low values of the mean and standard deviation, in special the later, as it is close to zero.

Free breathing on the other hand, for both methods, present wider limits of agreement with an increased mean and standard deviation, if compared to the constant breathing tests. This increase can be caused by the non-periodic nature of free breathing, and also by the fact that each subject was breathing at their natural frequency, thus producing more dispersion between subjects from the same test. Regarding the mean and standard deviation of the differences for this particular test, it can be appreciated that the Depth method presents a higher standard deviation and a lower mean than the Pattern method, while the Pattern method presents a higher mean and lower standard deviation. This can be interpreted if the statistical results for this particular test are taken into account, as the t-test between methods yield non-significant differences a possible explanation could be statistical fluctuation of the data. As for the null-bias hypothesis statistical results, the same explanation as in the constant breathing tests applies.

As for the Reading test for the Depth and the Pattern, it can be seen similar mean results to the Free test with an increased standard deviation for both methods. This increase can be caused due to the nature of the test, as the subjects were asked to read out loud a text, thus producing unique variations to the respiratory signal in function of how the subject was reading. Moreover, the statistical analysis for this test shows non-significant differences in either the null-bias hypothesis or the comparison between methods.

Given all the aforementioned results, and the different statistical analysis between the Depth and the Pattern methods. It can be observed that even though the results for both methods are very similar, even the point-cloud distribution of the BA plots, the Depth method yields slightly better results in the free breathing test.

Finally, for the Thermal results of the BA analysis in Figure 4.7 and Table 4.4, for the constant breathing frequency tests a gain error can be clearly seen. One explanation for the gain error could be a possible bias between the sampling frequency used to obtain the data, and the one used to process it. Moreover, the null-bias hypothesis also shows very significant and significant differences for the 0.1 Hz and 0.3 Hz test respectively, which are clearly direct consequence of the gain error. In the Free and Reading test on the other hand, no gain error can be appreciated. In these two test an increased standard deviation can also be seen if compared with the other two methods for the same test. The statistical results for these tests show non-significant differences in the null-bias hypothesis and in the comparison between methods.

Taking into account the results presented in Figure 4.8 and Table 4.5, for the 0.1 Hz test the mean value of the differences for the Pattern method almost doubles the one presented for te Depth method. This implies that the Pattern method has more variability in the IF estimation than the Depth method when compared with the reference method. Nevertheless, both methods presents narrow limits of agreement in the 0.1 Hz test. In the statistical analysis for this test, very significant differences can be seen for both methods, both in the null-bias hypothesis and the method comparison, which can be due to the fundamental frequency of the test being close to one of the cut-off frequencies of the band-pass filter. For the 0.3 Hz test on the other hand, both the Depth method and the Pattern method yield similar results, with small differences in the standard deviation. For the 0.3 Hz test, no statistical differences in the null-bias hypothesis or in the method comparison can be found for these methods. The Free and Reading tests, for both methods, yield an increased mean and standard deviation results if compared to the constant frequency breathing tests, which can be a direct consequence of the unconstrained breathing and reading as mentioned before. Regarding the statistical results for these two tests, in the null-bias hypothesis test, significant differences for the Free breathing in both methods can be found, while non-significant differences are found for the Reading test and for the comparison between methods. Moreover, for these two tests, the Depth method yields slightly better results than the Pattern method, indicating a higher agreement with the reference method in terms of IF variability.

The Thermal method results on the other hand, present a clear gain error for the 0.1 Hz and 0.3 Hz test. For all the tests an increased mean and standard deviation can also be seen if compared with the other two methods, which can be interpreted as a low agreement between the reference method and the Thermal method in terms of IF variability. The statistical results for this method also show very significant differences, either in the null-bias hypothesis and in the method comparison, thus indicating a low agreement between te reference method and the studied method.

For the results shown in Table 4.6 regarding the SDE, it can be seen that both the Depth and Pattern method show low median SDE and narrow IQR range for all test, thus indicating a very good agreement with the reference method. The Thermal method on the other hand shows an increased SDE median value, almost doubling the one presented for the other two metods, with also an increased IQR range for all the tests. This indicates a low agreement between the Thermal method and te reference method, and moreover, that the Thermal method presents higher IF variability than the reference method.

Given the SDE results, on the other hand, a clear improvement can be seen in either the median and the IQR range, whenever the frequency of the respiratory signal increases for all methods. This improvement can be due to the fundamental frequency of 0.1 Hz being to close to one of the cut-off frequencies of the band-pass filter used in the processing stage, whereas the 0.3 Hz does not present this type of issues. Regarding the Free and Reading test, an increase in the median and IQR can also be appreciated for all tests, which can be caused by the unconstrained nature of the tests.

Regarding the statistical analysis, and post-hoc tests, between the SDE of the different methods, non-significant differences can be seen for the interaction between the Depth and Pattern methods. The results for the interactions between the Depth and Thermal and the Pattern and Thermal, show very significant and significant differences respectively in the post-hoc analysis, which implies that the Thermal method yields different results than the other two tests, thus having a low agreement with the reference method.

There were several limitations to the study, being the first limitation the number of subjects that could be included in the study. Only 21 from the 23 subjects were used, as the two that were not included presented unrecoverable errors in the acquisition stage, thus producing errors in the signal extraction, specially for the Thermal camera. The second limitation was the signal extraction of the Thermal camera, as there were several errors introduced in the thermal signal, for example due to occlusions of the nostrils or sudden movements of the face. Even though the movements of the subject were mitigated by using a tracking algorithm of the nostril region, these movements still posed as the major source of error to this particular method. Moreover, as the resulting signal was not quasi-periodic, this posed as a problem in the IF estimation.

The same limitations regarding the measurement setup found in Chapter 3 apply to this study, as the measurements where performed in a very controlled environment.

4.5 Conclusions

In this chapter a comparison between three different video-based methods has been performed, in terms of respiratory rhythm and respiratory variability. The Depth and Pattern methods yield very good results when compared with the reference method, with small differences in the order of mHz in both the BA analysis and the SDE. For the Thermal method, even though the results show a poor performance if compared with the other two methods, it cannot be discarded directly due to the aforementioned limitations.

Chapter 5

Infrared-depth camera-based method for respiratory rhythm measurement

5.1 Introduction

From the previous chapter, it was concluded that the Pattern method, presented in chapter 3, yielded very good results when compared to the reference method, while it was also concluded that the method based on the depth camera performed slightly better than the Pattern method. The depth method was based on using a mono depth camera (Microsoft Kinect^{TM}) to obtain the point-cloud map of the subject. As described in the previous chapter in section 4.2.3, the respiratory signal of the subject was measured by obtaining the mean value of the depth point-cloud within the range 500 to 900 mm.

As it can be seen in chapter 2 section 2.3.2, in the literature there are multiple methods that make use of this type of cameras to obtain the physiological variables of a subject can be found in the literature. Nevertheless, in the same section the limitations of such methods are also described, ranging from the use of predefined ROIs, to tests that have been performed in such controlled conditions that the subject had very limited freedom of movement. Moreover, another limiting factor of such methods, and the depth method used in the previous chapter, was the use of mono depth cameras which require perfectly illuminated scene to properly work. In this chapter, the aforementioned limitations are corrected, by replacing the mono depth camera with an stereo depth camera to overcome the illumination limitation that present mono depth cameras.

The aim of this chapter is to present a new non-contact camera-based method to measure the respiratory signal of a subject using an hybrid architecture. The method uses synchronized IR and depth point-cloud feeds, obtained with the same camera, to track the movements of the subject and to compute a suitable ROI on demand using the IR feed, while the obtained ROI is used in the depth point-cloud to extract the respiratory signal of the subject as described in the previous chapter. By using an IR feed to obtain the movement of the subject, the presented method is more robust to illumination changes than a regular RGB camera, while being able to recompute the ROI automatically whenever the subject has changed its position inside the frame. This also makes the proposed method more robust to postural changes or sudden movements as it does not depend on a prefixed ROI. In this regard, no interaction with the subject is required to obtain its respiratory signal.

The contents and findings of this chapter have been published as a journal article [30] with the following reference:
Mateu-Mateus, M., Guede-Fernandez, F., Garcia-Gonzalez, M. A., Ramos-Castro, J., Fernandez-Chimeno, M. (2019). Non-contact Infrared-Depth Camera-based Method for Respiratory Rhythm Measurement While Driving. IEEE Access, 1–1. https://doi.org/10.1109/ACCESS.2019.2947759 doi: 10.1109/ACCESS.2019.2947759

5.2 Materials And Methods

The method proposed in this chapter is based in the simultaneous acquisition of both the infrared (IR) video feed and the depth point-cloud feed from a depth camera. The camera used, is a stereo depth camera which uses two IR video cameras and an IR projector (as described in chapter 2 section 2.3.2), to obtain the depth point-cloud. To obtain both video feeds from the camera, the provided software development kit from the manufacturer (librealsense library version 1.12.1) was used, while the proposed method was based on the OpenCV and Dlib [171] libraries (version 3.4 and 19.16 respectively). The code for the proposed algorithm was written entirely in Python (Version 3.7 with cython).

5.2.1 Proposed Algorithm

In Figure 5.1, a flow diagram of the proposed algorithm is shown, which includes both the IR and Depth video feeds, as well as all the processing steps necessary to obtain the respiratory signal from the subject.

Infrared Frame

The IR feed from the camera is used for two purposes, the first is to obtain the position of the face of the subject, which location and width will be used to obtain a ROI that will be later used on the Depth feed, and to detect whenever the subject is moving. To compute the movement of the subject, an optical flow algorithm has been used based on the Kanadae-Lucas-Tomasi (KLT) implementation, for this reason, two consecutive frames are needed from the IR feed. The following steps describe how the movement of the subject is obtained:

- In order to increase the performance of the optical flow computation, and as the aim is to obtain an approximate movement of the subject, the frame (current and previous) is first decimated by a factor of 4 by scaling the image to 1/4 of its original size. This value has been obtained empirically as it reduces the computational cost of the optical flow, without compromising the accuracy of the method.

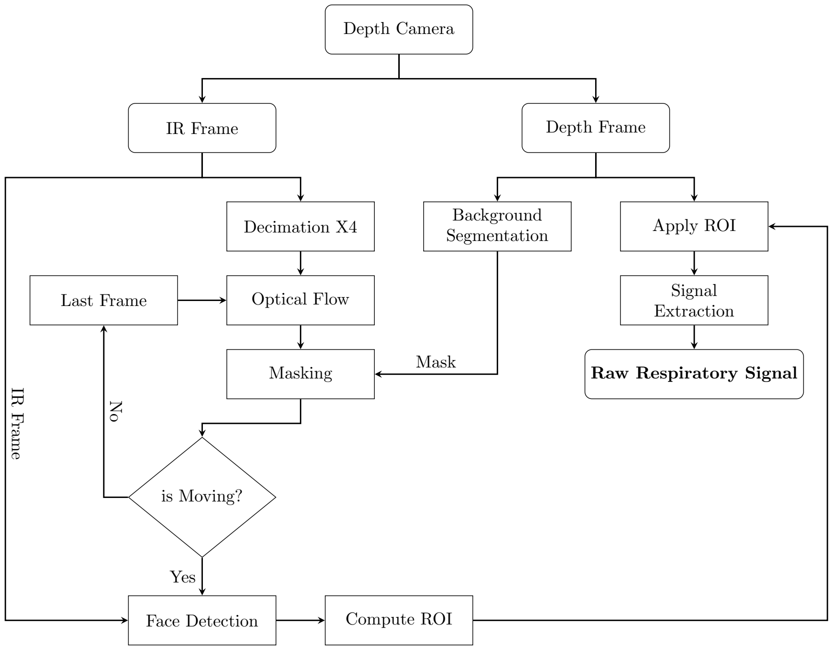


Figure 5.1: Block diagram of the proposed method. Source: [30] (CC BY 4.0: https://creativecommons.org/licenses/by/4.0/)

- The optical flow is computed using the current and previous decimated frames from the camera. The algorithm used for this step, is the pyramidal implementation of the Kanadae-Lucas-Tomasi (KLT) [125] described in chapter 2 section 2.3.2. For the first frame, a set of features are obtained using the Shi-Tomasi [161] good features to track algorithm. Then, the new location of these features is obtained in the current frame using the KLT algorithm.

- The background removal mask obtained from Depth point-cloud processing branch, is applied to prevent errors due to background movements.

- The last step, is to obtain the mean modulus ($l2$ norm) of the masked results of the optical flow. If the modulus is below a certain threshold, the previous frame is updated with the current frame. Otherwise, the face detection stage is triggered.

Whenever the algorithm detects that the subject is moving, the face detection stage is performed to obtain a new ROI location for the Depth feed. To obtain the position of the face from the subject, the last IR frame (without decimation) is used, in order to increase the spacial resolution of the face detection algorithm.

In order to detect the face inside the IR frame, the Dlib library [171] was used. The algorithm presented in this library uses the Histogram Oriented Gradients (HOG) features along with linear classifiers to obtain the position of human faces inside the frame. Whenever a face is detected inside the IR frame, a ROI is obtained following Equation 5.1. An example of the obtained ROI is depicted in Figure 5.2a.

$$\begin{aligned} X_{roi} =& X_i + w/2 \\ Y_{roi} =& Y_i + h * 2 \end{aligned} \tag{5.1}$$

where (X_{roi}, Y_{roi}) define central coordinates of the computed region, (X_i, Y_i) define the upper left corner of the rectangle that defines the face location within the frame, and finally, w and h are respectively the width and height of previous rectangle. The origin of coordinates is located on the upper left corner of the frame, as it can be seen in Figure 5.2a.

As it can be appreciated from Equation 5.1, the central coordinates for the ROI are obtained by taking into account the length and width of the rectangle that defines the face inside the frame. It can also be noted that the centre of the ROI is located two times the height and half the width of the face below the rectangle as it can be seen in Figure 5.2a. From this central point, the ROI is conformed by a 32 x 32 pixel square. The square width has been empirically obtained from previous experiments, where different widths and shapes were tried. Moreover, the 32 x 32 pixel square was the one that yielded better results in terms of respiratory signal, as well computational cost. Wider ROIs produced comparable results but had a higher computational cost, while smaller ROIs yielded poor results.

Depth Point-cloud

The depth point-cloud feed is also used to perform two tasks, the first is to obtain a mask that will be used to segment the background of the IR feed, while the second task is to extract the respiratory signal by the means of the obtained ROI from the previous steps.

The background segmentation mask is obtained by thresholding the point-cloud map between 0.5 m and 1 m. This distance margin represents the most probable range where the subject would be located respectively from the camera, according to the position of the subject and the camera within the measuring setup. These thresholds are applied in the following way, if the depth is greater than 1 m or lower than 0.5 m, the value is replaced by a zero. The mask is obtained by replacing the values within the desired range to the maximum value. Once the mask is generated, it is used in the IR stage to filter the optical flow computation.

Whenever the ROI is obtained from the IR stage, the respiratory signal from the subject can be extracted from the depth point-cloud by averaging all the points within the ROI. As average distance between the camera and the subject is modulated by the movement of the thorax, the respiration of the subject can be obtained [137]. A superimposed image of the obtained ROI within the depth point-cloud is shown in Figure 5.2b. The extraction of the respiratory signal from the subject is based on the method used in the previous chapter (chapter 4), and from the one presented on the patent Respiratory Signal Extraction [141] published in 2018 with number: WO/2018/121861.

(a)

(b)

Figure 5.2: a) Depicts the IR frame obtained from the camera, b) depicts the depth point-cloud. In (a) the different metrics used by the method are shown, and referenced to the location of the face. In (b) the computed ROI from the previous metrics is depicted as a blue square. The images correspond to a preliminary study and are not the actual setup. Source: [30] (CC BY 4.0).

5.2.2 Measurement Setup

The stereo depth camera RealSense ZR300 from Intel [133] was used to simultaneously acquire the IR frame and the depth point-cloud. The camera was configured via the librealsense library to acquire both feeds at 30 fps with a resolution of 492 x 372 pixels. This specific resolution was chosen, as it is the maximum resolution at which the camera is able to obtain the point-cloud for distances as short as 0.5 m. The camera was located inside the setup approximately between 0.6 - 0.8 m away from the subject, whereas the real distance between the camera and the subject was determined by the analysis of the point-cloud.

The same reference system used in the previous chapters, the RespiBand RIP sensor from BioSignalsPlux$^{\text{TM}}$ [94], was used to validate this method. The communication between the sensor and the computer was performed by the means of a Bluetooth Classic serial port.

In order to record the video feeds from the camera while simultaneously acquire the signal from the reference method, a custom computer program based on the Robotic Operative System (ROS) [172] was used. A *.bag binary file format was used to store all the video feeds and the signal from the reference method, with their respective timestamps so a posterior analysis could be performed. The *.bag ROS binary format allows to capture and emulate the exact time when the frames are captured from the camera. By emulating the time between frame arrivals, the proposed method has been tested for framerate loss due to the computational cost of the algorithm, hence testing the performance of the proposed method in real-time environments. The ROS version used in this study was the Kinetic Kame running on an Ubuntu 16.04 LTS.

5.2.3 Measurement Protocol

Twenty healthy subjects, 10 males and 10 females volunteered for the study. The anthropometric data for the volunteers is summarized in Table 5.1.

Table 5.1: Anthropometric data from the subjects expressed as mean ± SD

Age [years]	Height [cm]	Weight [kg]	BMI [kg/m^2]
37.55 ± 6.66	171.48 ± 10.38	76.99 ± 18.02	26.04 ± 4.87

Table reproduced from [30] (CC BY 4.0).

The measurements for this study were performed at the facilities of the Biomechanics Institute of Valencia (IBV). The measurement setup was composed of a full-size car adapted to perform as a simulator, and a highway route projected on a wall screen in front of the setup. Pictures of the setup, provided by the IBV, are shown in Figure 5.3.

(a) (b)

Figure 5.3: Pictures of the car simulator from the facilities of the Biomechanics Institute of Valencia, photos courtesy from the IBV, Source [30] (CC BY 4.0)

All the measurements and the protocol to perform them, was dictated and performed by the Biomechanics Institute of Valencia (IBV). Each subject was asked to wear the RIP sensor from BioSignalsPlux, and was asked to behave and drive normally within the capabilities of the simulator, while following the road indications. Moreover, each subject was able to adjust the car seat at convenience. All the measurements were provided as is, and only the data was processed to obtain the respiratory signal.

5.2.4 Signal Processing

Once the raw respiratory signal has been extracted by the means of the proposed method, the signal has to be normalized in order to be compared with the signal obtained from the reference method. All the signal processing was performed using Matlab version 2018b, while the statistical analysis was performed using R version 3.6.0.

Signal Normalisation

The signal processing techniques that have been used to process and normalise the raw respiratory signal, were the same ones that were applied in the previous chapters of this thesis. A summary of the signal processing steps undertaken to normalise the respiratory signal are detailed.

- The signals are interpolated at 40 Hz using a cubic spline.
- A second order bidirectional Butterworth band-pass filter between 0.05 Hz and 1 Hz is used on the interpolated signals.
- The result of a moving median filter [154] is subtracted to the signal to remove the transitory periods produced by the rapid movements of the subject.
- Finally, to compress the signal between -1 and 1, a non-linear function based on the arctangent [155] Equation 5.2 is applied.

$$S_n[k] = \arctan \frac{S[k]}{\sqrt{\frac{\sum_{i=1}^{N}(S[i]-\bar{S})^2}{N-1}} * \sqrt{2}} \tag{5.2}$$

where $S_n[n]$ is the resultant signal after the non-linear function (normalised respiratory signal), $S[n]$ is the filtered respiratory signal and $\bar{S}$ is the mean of $S[n]$.

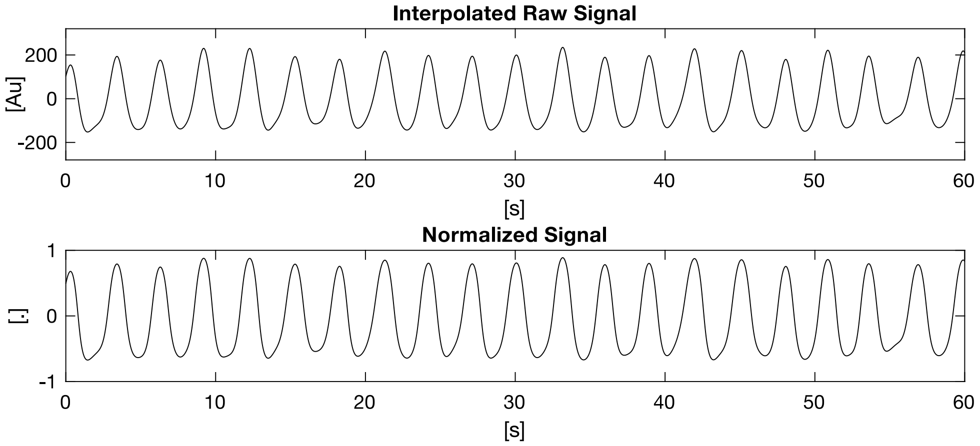


Figure 5.4: The first plot depicts the interpolated raw respiratory signal from the reference method, while the second plot corresponds to its normalized version. Source: [30] (CC BY 4.0).

Figure 5.4 shows an example of the raw respiratory signal and the normalised respiratory signal obtained from the reference method.

Opportunistic piece selection

Once the signals have been normalised, as the aim of the proposed method is to measure respiratory rhythm in real-life conditions, each test has been divided into pieces based on an opportunistic approach. Each piece has a variable length with a minimum duration of 60 s. Multiple pieces selected by visual inspection have been obtained from each subject. Each piece has been selected based on the amplitude of the signal, where a piece is deemed valid if the amplitude is at least 0.5 and the total length of the piece is equal or greater than 60 seconds.

5.2.5 Performance Characterisation

Statistical performance

To compare the proposed method with the reference method, two types of evaluation have been performed: the respiratory cycle series analysis and the Hilbert IF signal comparison between methods, from each one of the opportunistic pieces independently from the subject that where taken.

In order to prevent errors when comparing pieces from the reference method and the proposed method, the signals have been aligned using the Fisher's intra-class correlation (ICC) [156], by iterating around one respiratory period. Because the maximum displacement between methods will not be greater than one respiratory period, using one period length to slide the signals between each other ensures that the ICC correlation is maximised, and guarantees a perfect alignment between the signal obtained from the proposed method and the one from the reference method.

The respiratory cycle series has been obtained using the same procedure than in chapter 3. The following steps summarise the procedure described in chapter 3 section 3.2.4.

- A threshold based on the percentile 65 is computed in both signals.
- Using the previous threshold, the intersection with positive slopes is obtained from both methods.
- Finally, the distance between consecutive slopes is computed to form the RC series.

To compute the Hilbert IF from both methods, the same procedure that has been described in chapter 4 section 4.2.4 has been used. The following steps summarize how the Hilbert IF is obtained.

- The Hilbert transform is obtained from both signals.
- The phase of the Hilbert transform is obtained and unwrapped with 2π increments to ensure continuity.
- The difference between adjacent samples is obtained to conform the IF in radiants and then converted to Hz.
- Finally, a Hodrick-Prescott [168] with a smoothing factor of $2 * 10^6$ is applied to the signal to remove the effects of the previous phase correction.

Detection Assessment

To characterise the performance in the respiratory cycle detection, the same techniques and criteria presented in chapter 3 section 3.2.4 have been used. A confusion matrix using: true positive (TP), false positive (FP), false negative (FN), sensitivity (SEN) and predictability of positive value (PPV) indicators has been computed for each piece.

The same error indicators from chapter 3, have also been used to characterise the error between respiratory cycle series obtained from the proposed method and the reference method. The used indicators are the mean absolute error (MAE) and mean absolute percentage error (MAPE). Moreover, the Fisher's ICC for all the computed cycles, without making a distinction between subjects or pieces, has been obtained.

Finally the standard deviation of the error (SDE) has been used to characterise the error in a sample by sample basis either from the respiratory cycle series and the IF signal between methods.

Error Assessment

In order to compare each respiratory cycle from the proposed method with its counterpart from the reference method, a Bland-Altman (BA) analysis [157] has been performed. For this analysis each piece has been used independently from the test or the subject that it belongs. Regarding the IF, a BA analysis has been performed as well, taking into account the mean IF and the SD IF. For these two last tests, each piece has been used independently of the subject for this analysis. Prior to performing the BA analysis, an Anderson-Darling test (ADT) [173] has been performed to ensure that the BA differences had a normal distribution. For this reason, the median and percentiles 2.5 and 97.5 (95 % reference interval) has been used instead of the traditional mean and limits of agreement for the IF BA, as the samples do not present a normal distribution.

The Anderson-Darling test (ADT) was also applied for the SDE results, either for the respiratory cycle results and for the IF results, as well as to the length of each piece, to verify if the data had a normal distribution or not. The results of the test show that for both the cycles SDE and the length of each piece, there is no significant differences, which indicate that the data presents a normal distribution. The IF SDE on the other hand, and similarly to the BA IF, does present very significant differences, which discards the null hypothesis that the data presents a normal distribution.

Regarding the opportunistic piece selection, in order to test the statistical dependence between the length of each piece and the SDE, for either the respiratory cycles and for the IF, an ANOVA test [174] between the length of each piece and the mean SDE for each piece has been performed, as well as Kruskal-Wallis test [175] between the length of each piece and the SDE IF for each piece. An ANOVA test has also been performed between the cycle SDE and the PPV for each piece to assert the relationship between increased FN and an increase in the cycle SDE.

5.3 Results

5.3.1 Signals

Table 5.2 summarises the number of pieces obtained from each subject, the piece length for all subjects expressed as median and IQR, the number of cycles for each piece, and finally the length of each cycle for each piece from the reference method and the proposed method.

Table 5.2: Opportunistic piece selection results. RM stands as Reference Method, PM stands as Proposed Method.

Piece/Subj.	Piece Len. [s]	Cyc./Piece	Cyc.Len/Piece RM [s]	Cyc.Len/Piece PM [s]
6.53 ± 3.22	167.5 [108 ; 271.38]	47 [27 ; 77]	3.45 [2.98 ; 4.55]	3.45 [2.98 ; 4.63]

Table reproduced from [30] (CC BY 4.0).

Figure 5.5 shows an example obtained from one of the pieces, where a comparison is shown between respiratory signals obtained from the proposed method and the reference method, as well as the respiratory cycle signals and the IF signals from both methods.

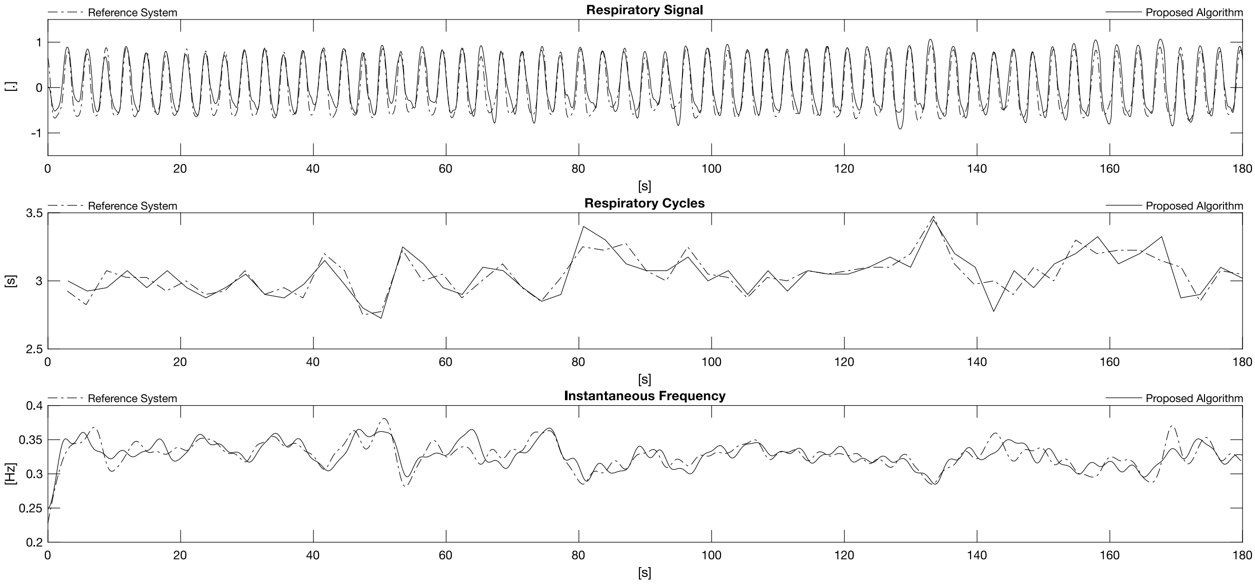


Figure 5.5: Example of the obtained signals from both methods for a given piece. The first plot depicts the normalised respiratory signals, the second plot corresponds to the RC series and the third plot to the IF. Source: [30] (CC BY 4.0).

On Figure 5.5, it can be seen that the respiratory signals from both methods are very similar between each other, indicating a high concordance between methods. For the respiratory cycle plot and the IF signal plot, it can also be seen that both signals for both methods follow the same temporal and amplitude evolution, with small differences between signals.

As an example of the concordance between the cycle series and the IF signal, if it is taken into account the interval between 40 s and 60 s from the previous figure, it can be seen that whenever the respiratory cycle series increases, an equivalent decrease in the IF can be observed. This can be explained as whenever the respiratory cycles increase in length, the instantaneous frequency for that given cycle decreases. The same case can be observed as whenever the respiratory cycles decreases, the IF signal increases.

5.3.2 Performance

Table 5.3 show the results for the accuracy assessment of the respiratory cycle series. Regarding the TP, FP and FN results, no distinction between subjects or pieces has been done. Moreover, a global results are shown for the SEN and PPV, computed using the aggregated TP, FP and FN.

Table 5.3: Performance indicators of the cycle detection, where the SEN and PPV indicators are expressed as mean ± SD.

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 stands for Global SEN and G. PPV stands for Global PPV. Table reproduced from [30] (CC BY 4.0).

Table 5.4 shows the results for the mean correlation between cycle series for all pieces, as well as the MAE and MAPE results.

Table 5.4: Fisher ICC for all the computed cycles, MAE and MAPE error indicators expressed as mean ± SD.

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

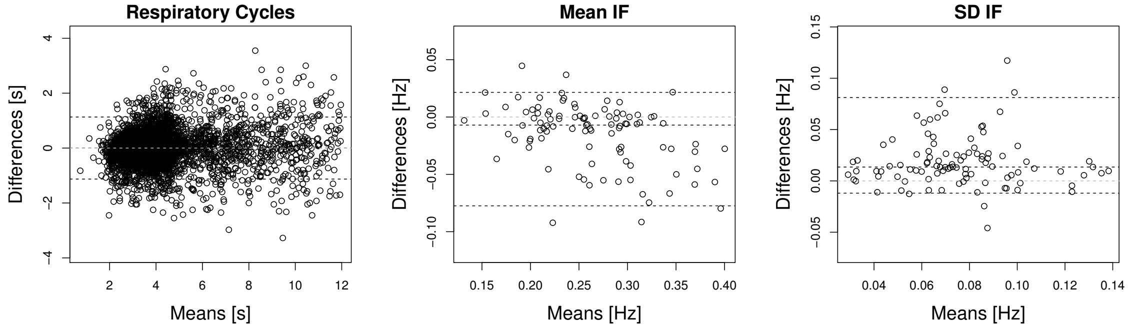
Table reproduced from [30] (CC BY 4.0).

Regarding the MAE and MAPE results, both indicators have been obtained from the TP cycle series, moreover, as the MAPE indicator is influenced by the length of the analysed signal, to avoid errors due to the differences between lengths for each piece, the weighted mean and standard deviation has been used for this indicator.

Error Characterisation

Prior to the Bland-Altman (BA) analysis, a series of outliers had to be removed. The applied criteria was: if the cycles where at least 2 times or greater than the standard deviation of the group, those cycles were removed. Regarding the IF, the criteria was that if the differences for a piece were greater or lower than the 95 % reference interval, those pieces were discarded. A total of 4 pieces were discarded for the IF, and a total of 100 cycles were removed.

Figure 5.6 shows the computed BA plots, being the first plot (Figure 5.6a) the comparison between respiratory cycles from the reference method and the proposed method, were all the cycles have been used independently. The two remaining BA plots (Figures 5.6b, 5.6c), show respectively the mean IF and the SD IF for each piece between the proposed method and the reference method, where no distinction between subjects has been made. The same representation as in the previous chapters of the mean and limits of agreement as well as the zero-mean has been performed for the Figure 5.6a, being the grey dashed line the zero-mean, while the other three lines represent the mean and limits of agreement. For the plots in Figures 5.6b and 5.6c, the grey dashed line also represents the zero-mean, while the other three lines represent the median and the 95 % reference interval.



(a) (b) (c)

Figure 5.6: Bland-Altman representation where a) analyses all the detected cycles, b) the mean IF and c) the SD IF. Source: [30] (CC BY 4.0).

Table 5.5 shows the results for the BA differences of Figure 5.6, being the differences for the cycle BA plots represented as mean and SD, and the differences for the IF BA plots represented as median and 95% reference interval.

Table 5.5: Numerical results of the BA plots in Figure 5.6, expressed as mean ± SD for Figure 5.6a and median and 95 % reference interval [2.5 ; 97.5] for Figures 5.6b and 5.6c.

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 reproduced from [30] (CC BY 4.0).

The results for the SDE of the respiratory cycle series and the IF signal are summarised in Table 5.6, as well as, the results for the statistical analysis regarding the influence of the length to the cycle SDE and the comparison between the cycles SDE and the PPV. This last results show non-significant differences for the influence of the length to the cycle SDE, while significant differences ($p < 0.05$) can be found for the cycle SDE and PPV interaction.

Table 5.6: Results for the SDE and statistical analysis. The SDE is expressed as mean ± SD for the RC series and median and IQR [25 ; 75] for the IF.

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

The significance results, either for the ANOVA and the Kruskal-Wallis test, are presented as: "$*$" for the interaction between length vs cycle SDE, "#" for the interaction between length and IF SDE and finally, "†" for the interaction between cycle SDE vs PPV. The significance is represented as follows: "NS_{*}" for $p > 0.05$, "$*$" for $p < 0.05$ and "$**$" for $p < 0.001$. Table reproduced from [30] (CC BY 4.0).

5.4 Discussion

Given the results shown in Table 5.3 for the accuracy of the cycle detection, a global SEN and PPV (77.21 % and 80.69 %) are reported. If each piece is taken into consideration, the SEN and PPV yield slightly better results with mean values of 80.12 % and 80.74 % respectively with both presenting a standard deviation below 15 %. This results cannot be considered as high, but it has to be taken into account that the subject had no restriction whatsoever inside the cabin, which is translated into multiple movements (even within the same piece) that produce abrupt baseline shifts in the obtained signal from the proposed method. Even after the normalisation of the obtained respiratory signal, and even though most of the aforementioned errors are greatly minimised, those errors still have a very negative impact in the cycle detection, thus decreasing the detection accuracy. For this reason, different results are also yielded when taking into account all the detected cycles or only the ones detected from each piece.

Moreover, another relevant factor to take into consideration is the variability between pieces within the same subject, as it has a clear impact in the SEN, PPV and error results. This factor can be seen in the statistical results of the interaction between the SDE and the PPV for each piece (Table 5.6), which show significant differences ($p < 0.05$) in the interaction between these two indicators. The rationale behind this result is that whenever the PPV decreases, as the PPV is the relationship between TP and FP, an increased number of FP can be seen. This also implies that the signals present more disparities between each other, which will be translated into a higher SDE between methods. If the PPV indicator increases, then the SDE will decrease as well. The errors which produce a higher detection of FP and the increase of the SDE, are greatly increased due to the freedom of movement of the subject and the errors associated to these movements as explained before.

Regarding the results shown in Table 5.4, a high cycle correlation (0.96) between the proposed method and the reference method can be seen. For the MAE and MAPE results, low values for the mean and standard deviation can be appreciated for both indicators. For the MAE results, if compared to the mean duration of a respiratory cycle, this results only represents roughly the 10 % of the median length. The MAPE results, on the other hand, show values similar to the ones reported for the MAE indicator, with a weighted mean and standard deviation of 8.49 % and 2.99 % respectively. This results can be interpreted as a good agreement between methods, if it is taken into account the aforementioned constraints associated with the tests.

To further compare the proposed method with the reference method, taking into account the movements of the subject and without the limitations of the reliability of the cycle detection for these constraints, the Hilbert IF signal for each one of the methods has been obtained. By using the IF, as it is obtained without the need of using a threshold, the comparison between methods does not need to take into account piece-related information such as the length of the piece.

For the results of the BA analysis in Figure 5.6a and Table 5.5, it can be seen a high agreement between methods, with mean and standard deviation for the differences of 10.13 ms ± 576.56 ms, where no apparent bias can be seen. Regarding the BA plots on Figures 5.6b and 5.6c, even though it cannot be directly appreciated by visual inspection that there is no bias present (in both figures the zero value is within the 95 % reference interval), as the samples do not present a normal distribution, it cannot be discarded either. For the results of these two BA plots in Table 5.5, the mean IF BA plot presents a low median value with a narrow 95 % reference interval, which can be interpreted as a good agreement between methods. For the SD IF BA plot, a low median and narrow 95 % reference interval can also be observed, which also indicates a good agreement in the variability of the IF for both methods.

Given the standard deviation of the error results presented in Table 5.6, the cycle SDE shows relative low values with a mean and SD of 0.62 s and 0.32 s respectively, which indicate a high agreement between the obtained cycles for the proposed method and the reference method. Regarding the SDE IF results, a low median value and increased 95 % reference interval can be seen. This results are in consonance with the results in the BA analysis, moreover, if it is taken into account that the subject had no constraints during the test, the increased 95 % reference interval can be a direct consequence of the movements inside the setup, and the unique respiratory patterns for each subject.

Regarding the statistical analysis results in Table 5.6, non-significant differences can be found either for the ANOVA results and the Kruskal-Wallis tests, regarding the interaction between the length of the pieces and cycle SDE and the length of the pieces and the IF SDE respectively. This results can be interpreted as there is no statistical interaction between the length of each piece and the error associated with it (either for the cycles or the IF). Moreover, this results confirm that using variable-length pieces does not have an impact on the error between the proposed method and the reference method, which implies the viability of using variable-length opportunistic piece selection in real-life systems, to compare alternative methods that measure respiratory rhythm.

Several limitations were present in this study, being the first limitation the number of subjects that could be used for this study. Only 17 of the 20 participants were used, as for the three discarded subjects the depth point-cloud was corrupted beyond recovery due to errors in the generation of the point-cloud. This errors where produced because the subjects were more close to the camera than the minimum required distance for the camera to generate the point-cloud map.

The second limitation of this study is regarding the test conditions, even though the test were performed in a full size car without imposing any constraint to the subject, the lightning conditions of the setup were more close to a night environment without illumination changes. Moreover, no external vibrations were present during the test. For these reasons, it cannot be assured that the same performance would be achieved in presence of external vibrations or in changing light conditions (daylight), as both conditions would certainly have a negative impact in the respiratory signal extraction.

The last limitation of this study was the lack of an automatized process to perform the piece selection based on the quality of the signal. An algorithm to assess the quality of the signal, would have provided an objective tool to select the length of each piece or even to provide a template to select the pieces by visual inspection. As each piece was selected by visual inspection, inevitably some errors were introduced to the pieces. To compensate for these errors, some outliers had to be removed from the results.

5.5 Conclusions

In this chapter a new hybrid non-contact IR-Depth method has been presented, with the aim of measuring the respiratory signal of a subject without any intervention from the user. The proposed method uses the IR feed to obtain the movements of the subject, as well as the position of the face (used to compute the ROI), and the depth point-cloud feed which is used to perform a background segmentation to prevent errors, and by using the previous ROI the respiratory signal of the subject is extracted. The method has been validated using a commercial inductive plethysmograph. Moreover, 20 subjects were measured in a full car simulator without any constraints. The obtained results show a high correlation between the obtained cycles from both methods, with low error results either for the respiratory cycle series and the IF signal, indicating a good agreement between the proposed method and the reference method, moreover, no bias could be found in the BA analysis. The statistical results show that no statistical interactions could be found regarding the length of the opportunistic pieces and the error for each piece, which validate the hypothesis that variable-length opportunistic pieces can be obtained based on the quality of the signal, without having a negative impact on the analysis and characterisation of the method.

Chapter 6

Respiratory rhythm measurement from a lateral perspective

6.1 Introduction

As presented in the state of the art section 2.1, the respiratory mechanics are not only confined to the torso of the subject, as it includes the abdomen as well [34, 32]. Even though there are plenty of examples in the literature that used the thorax displacement to obtain the respiratory signal by the means of video-based methods (as well as the methods presented in the previous chapters), there are only a few that measure breathing from a lateral perspective [176, 33]. As an example, in the study presented by *Nozoe et al.* [33], the wall volume variation during tidal breathing is studied by the means of multiple optical markers located around the thorax and abdominal region of the subject. Even though the method cannot be directly compared with the proposed method in this chapter, as the study does not measure breath to breath variability, the study concludes that it is feasible to measure chest wall volume variation during tidal breathing from a lateral perspective in healthy subjects. Using this measuring point instead of a frontal approach, can have a series of benefits which are not present in more traditional methods. For instance, as the method would no longer record the face of the subject or the upper torso, the subject could benefit for an improved privacy if compared to traditional video-based methods. Moreover, as the measuring point is located at the side of the thorax, the occlusions produced by the movement of the arms can be greatly attenuated.

Given the two previous methods presented in chapters 3 and 5, even though both methods yield good results in terms of respiratory rhythm measurement, both methods still have several limitations that could be improved. For instance, both require advanced computer vision techniques to acquire the respiratory signal. In the case of chapter 3, a pattern needs to be located inside the frame, while for chapter 5, the location of the face needs to be obtained in order to compute the necessary ROI for the respiratory signal extraction. Moreover, as both methods measure the respiratory signal from a frontal perspective, they are less immune to occlusions. While this last example had little effect on the Pattern method (chapter 3), it definitely had an impact on the

Depth method as the subjects where given total freedom to move inside the measuring setup. Nevertheless, regarding this last issue, the Depth method (chapter 5) had the ability to compute the movements of the subject, in order to recompute the used ROI to obtain the respiratory signal.

From the idea exposed in the previous chapter of using the modulus of the dense optical flow in order to obtain the movements of the subject, and given the feasibility of measuring the respiratory signal from a lateral perspective, we propose in this chapter a new method that makes use of a dense optical flow techniques at the side of the thorax to obtain the respiratory signal of the subject.

The proposed method uses dense optical flow techniques to measure the displacement of the intercostal and abdominal muscles produced by breathing, by the means of a video camera located at the side of the thorax, while the subject is seated. Furthermore, to prevent errors in the respiratory signal extraction (occlusions, movements, etc.) a method to evaluate the respiratory signal quality is proposed. The quality indicator is based on the same idea of tracking the movement of the subject from the previous chapter, thus allowing the automated detection of quality respiratory signal pieces. Good quality has been defined as when the only source is the respiratory signal, while bad quality is defined when the respiratory signal is embedded with noise or errors produced by occlusions, movement or even talking. Finally, the proposed method has been designed with the aim of being extremely simple to implement, thus minimizing the possible bottlenecks and guaranteeing real-time performances. Moreover, the method has also been designed to make use of the full frame instead of a ROI, while constraining the necessary resolution from the camera.

The contents and findings of this chapter have been published as a journal article [31] with the following reference:
Mateu-Mateus, M., Guede-Fernández, F., Á. García-González, M., Ramos-Castro, J. J., Fernández-Chimeno, M. (2020). Camera-Based Method for Respiratory Rhythm Extraction From a Lateral Perspective. IEEE Access, 8, 154924–154939. https://doi.org/10.1109/ACCESS.2020.3018616 doi: 10.1109/ACCESS.2020.3018616

6.2 Materials And Methods

6.2.1 Respiratory Mechanics

As stated in chapter 2 section 2.1, the muscles involved in the respiratory mechanics (intercostal muscles and diaphragm) are the main responsible of the expansion and contraction of the upper torso, which favours the gas exchange at the lungs. Nevertheless and as stated in section 2.1, there are other contributing factors to the respiratory mechanics such as the abdominal movement [32, 33], that have a clear impact on the combined respiratory effort [34]. As an example, Nozoe *et. al* performed an study of the viability of measuring the respiratory volume evolution of the chest wall and abdomen using optical markers [33], which has proven crucial to this study as [33] demonstrates the feasibility of acquiring the respiratory signal by observing the evolution of the thoracic-abdominal region from a lateral perspective.

6.2.2 Proposed Algorithm

The proposed method measures the displacement of the rib cage and the intercostal and abdominal muscles by the means of dense optical flow tracking, from a lateral perspective while seated. The proposed method has been written in Python (Version 3.7 with Cpp bindings like cython) and makes use of the OpenCV (Version: 3.4) library to perform the different image processing steps, as well as the acquisition of the video feed from the camera.

The following pseudo-code (Algorithm 1), summarizes all the basic operations that are needed to compute the respiratory signal from the video signal, as well as to clarify the different components that are necessary for the algorithm to work.

Algorithm 1 Proposed Algorithm, Source: [31] (CC BY 4.0)

```
procedure RESPIRATORY SIGNAL ACQUISITION
    $frame \leftarrow$ getGreyFrame()
    $(X, Y) \leftarrow$ calcOpticalFlow($frame$, $preFrame$)
    $\langle r, \angle\theta \rangle \leftarrow$ cart2polar(X,Y)
    $mod \leftarrow$ mean($r$)
    $ph \leftarrow$ mean($\theta$)
    quality_index $\leftarrow calcSignSkew(mod)$
    raw_respiration $\leftarrow ph$
    $preFrame \leftarrow frame$
```

As it can be seen in Algorithm 1, the proposed method has a very simple architecture that only relies on the computation of the dense optical flow to obtain the raw respiratory signal from the phase, while the quality index is extracted from the modulus. This simplicity allows to infer possible bottlenecks in the performance of the method and to characterise the latency of the respiratory signal extraction. As the optical flow algorithm is the limiting factor in terms of performance of the method, by choosing an efficient implementation of this method would directly translate into low latency in the respiratory signal extraction.

The chosen optical flow algorithm for this study has been the Dense Inverse Search (DIS) by Kroeger *et. al* [127], which is part of the OpenCV library. The next subsection presents a more in detail exposition of the proposed method, and how the respiratory signal and the quality index are obtained.

Optical Flow & Signal Extraction

The first step of the algorithm is to acquire the frame from the camera. As the algorithm is intended to work even if an IR camera is used, and as a grey-scale image reduces the computational cost of the algorithm (instead of 3 channels only one is used), a conversion between the obtained frame to the grey-scale colour-space is performed.

Optical flow algorithms compute the motion (image displacement) between two consecutive frames. There are two types of optical flow algorithms: sparse and dense. Sparse optical flow algorithms such as the Kanadae-Lucas-Tomasi (KLT) [125] only compute the new location of a given set of features from the last frame to the current frame, while dense optical flow algorithms such as the DIS track the motion in a pixel by pixel basis. To compute the optical flow two consecutive frames from the camera are needed, if only the current frame is available from the camera, a copy of the current frame is used as the previous frame.

Regarding the DIS algorithm by Kroeger *et. al*, its principal characteristic is how the algorithm computes the motion between two frames by searching for patch correspondences for a given point between the two frames. Also the algorithm uses the inverse search method to find this correspondences and finally the algorithm computes the dense flow field of every patch. The full mathematical definition of the DIS algorithm can be found in [127]. The main advantage of this algorithm is that it has been optimized to perform in real-time with high accuracy in the optical flow estimation.

After the optical flow has been computed using the DIS algorithm implemented in the OpenCV library, the function returns two matrices containing the x coordinates (**X**) and y coordinates (**Y**) of the motion (flow field) between the two consecutive frames. After the coordinates of the flow-field are obtained, by computing the phase and the modulus of the flow field, the raw respiratory signal and the raw quality index are computed.

Raw Respiratory Signal

The following equations (6.1, 6.2) summarize how the phase of the flow-field is obtained from the **X** and **Y** coordinate matrices.

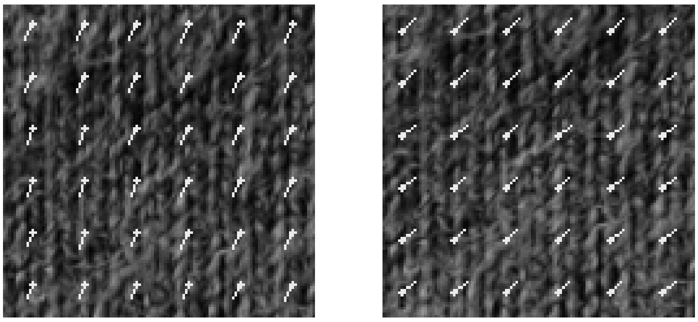
$$\boldsymbol{\theta} = atan2(\mathbf{X}/\mathbf{Y}) \tag{6.1}$$

$$R(k) = \frac{1}{N}\sum_{i=1}^{N}\theta_i \tag{6.2}$$

where $\boldsymbol{\theta} \in (-\pi, \pi]$ is a matrix containing all the angles (phase) for each set of (x, y) coordinates. $R(k)$ is the average value of all the phase values for a given frame, which in return is the kth sample of the raw respiratory signal. Finally N represents the total number of points in the flow field.

As it can be seen in Equation 6.1, the phase of the flow field is computed using the arctangent formula, particularly the $atan2()$ implementation has been used instead of the classical $atan()$ function. One of the particularities of the $atan2()$ implementation, is that it takes into account the real quadrant of the phase which it translates into a continuous range between $(-\pi, \pi]$. Once the phase for each set of (x, y) coordinates is computed, to obtain the raw respiratory sample for the given frame an average of all the phase samples is computed (Equation 6.2).

Figure 6.1 depicts an example of how the flow field looks like during inhalation (Figure 6.1a) and exhalation (Figure 6.1b), whereas the cross represents the current position of the characteristic on the frame, while the tail represents the position of the cross in the previous frames.



(a) (b)

Figure 6.1: 120 px wide square images of the thoracic-abdominal region, where a) shows the movement during inhalation and b) during exhalation. The images where extracted using the OpenCV library. Source: [31] (CC BY 4.0).

As it can be appreciated in Figure 6.1a, the cross-head arrows of the flow field are oriented to the first quadrant, which will yield at the maximum inhalation values comprised between $[0, \frac{\pi}{2}]$. On Figure 6.1b on the other hand, it can be appreciated that the cross-head arrows of the flow field are oriented to the third quadrant yielding at the maximum exhalation values between $(-\pi, -\frac{\pi}{2}]$. This phenomena is explained as whenever the thorax rises during inhalation, the ascending displacement of the thorax orients the flow field to the first quadrant, while when the thorax deflates during exhalation, the flow field is oriented in the opposite direction, hence the third quadrant. By computing the evolution of the angle of the flow field during the respiratory cycle (at each frame), and taking into account the previous phenomena where the flow field is oriented depending on the movement of the thorax, the raw respiration of the subject can be obtained through the phase of the optical flow.

Quality Index

As stated in chapter 5, the modulus of the optical flow changes whenever the subject moves, yielding higher amplitudes whenever the subject moves abruptly or is talking, and smaller amplitudes whenever the subject is at rest. The rationale behind the hypothesis that a quality index can be extracted from the modulus of the optical flow is: as the respiratory signal is a slow frequency signal, the frame by frame movement and hence the modulus, will yield small values when the subject is breathing normally, whereas when the subject is talking, or in the presence of abrupt or sudden movements (as used in chapter 5 to recompute the ROI) the modulus will yield higher values. Taking into account this behaviour, by analysing the statistical behaviour of the modulus (third moment estimator or Skewness), a quality index can be proposed.

The modulus of the flow field (Equation 6.3) for each set of (x, y) coordinates is computed using the $l2$ norm (euclidean distance). To obtain a raw quality index sample, an average of all the modulus points for each set of coordinates of the frame is computed (Equation 6.4), conforming the modulus signal.

$$\mathbf{M} = \sqrt{\mathbf{X}^2 + \mathbf{Y}^2} \tag{6.3}$$

$$S(k) = \frac{1}{N}\sum_{i=1}^{N} M_i \tag{6.4}$$

where the matrix $\mathbf{M}$ contains all the modulus for each set of (x, y) coordinates of the flow field, $S(k)$ represents the averaged modulus and the kth sample of the modulus signal. Finally N represents the total number of points in the flow field.

To compute the quality index, three steps are required: the sliding window conformation, the skewness computation (Equation 6.5) and finally to smooth the quality index, a recursive filter (Equation 6.6) is used.

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

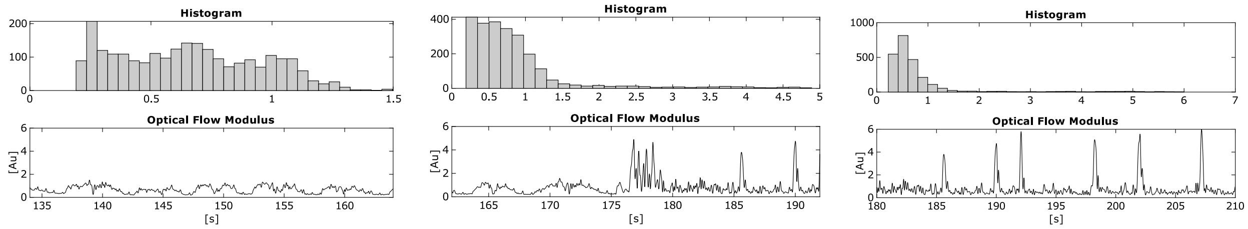
$$m_3(k) = E\left[\left(\frac{\mathbf{s}-\mu}{\sigma}\right)^3\right] \approx \frac{\frac{1}{n}\sum_{i=1}^{n}(s_i - \bar{s})^3}{\left(\frac{1}{n-1}\sum_{i=1}^{n}(s_i - \bar{s})^2\right)^{\frac{3}{2}}} \tag{6.5}$$

$$Q(k) = 0.9\ Q(k-1) + 0.1\ m_3(k) \tag{6.6}$$

where $\mathbf{s}$ represents the sliding window which contains n samples of the modulus, $\bar{s}$ represent the mean value of $\mathbf{s}$, n is the number of samples of the sliding window in seconds multiplied by the sampling frequency of the camera, $m_3(k)$ represents the kth sample of the computed skewness of $\mathbf{s}$, and finally $Q(k)$ is the kth sample of the quality index being $Q(0) = 0$ as the initial value of the recursive filter.

To conform the sliding window from the raw modulus samples, a First In First Out (FIFO) queue has been used. FIFO queues are efficient and easy to implement, moreover, by using this type of queue the amount of memory used is limited to the size of the queue. If the classical FIFO window is implemented, whenever a new modulus sample is available, the oldest sample in the queue is automatically discarded. An analysis of the influence of the window length in the quality index is performed in the results section, with the aim to assess which number of samples yield better results.

Figure 6.2 shows three histogram plots from three different snapshots of sliding windows of 30 second length each, containing three different types of modulus from three different respiratory patterns. Figure 6.2a contains 30 seconds of free breathing at a regular peace, Figure 6.2b contains a transition between free breathing and reading out loud a text, and finally Figure 6.2c contains 30 seconds of the subject reading a text. As it can be seen in Figure 6.2, depending on the nature of the respiratory signal (free breathing or reading) the symmetry of the statistics of the sliding window change. For this reason as a second step to compute the quality index, the 3rd statistical moment (skewness) has been used to quantify the symmetry of the statistic inside the sliding window.



(a) (b) (c)

Figure 6.2: Histogram of the modulus given three different scenarios: a) free breathing, b) transition between free breathing and reading and c) reading. Source: [31] (CC BY 4.0).

As an example, in Figure 6.2a as the subject is breathing freely and in a regular manner, the statistics of the modulus are practically symmetrical which will yield a low skewness value. On the other hand, as it can be seen in Figure 6.2b whenever the subject starts to read a text, the statistics shift to the left becoming more asymmetrical which in return will produce an increase in the computed skewness. Finally Figure 6.2c shows a completely asymmetrical histogram that will produce a high skewness value. As the respiration becomes more regular, the statistics of the sliding window will become more symmetrical, thus the value of the skewness will decay to low values. This example is still valid if the source of the error is not an alteration of the respiratory signal, but sudden movements of the subject or occlusions.

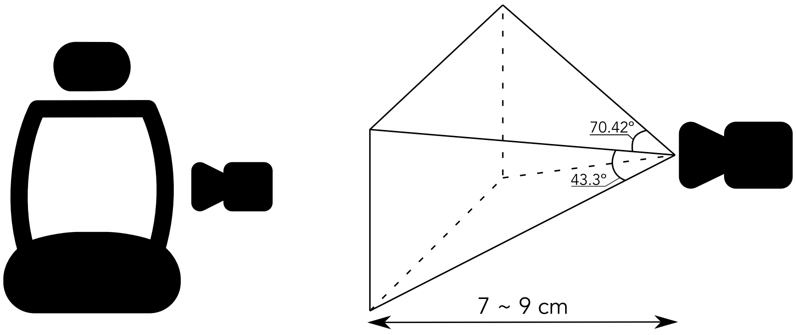
Finally, to smooth the quality index and to remove undesired components, a recursive low-pass filter (Equation 6.6) has been used. The multiplying factors of the filter are 0.9 for the previous sample and 0.1 for the current one. This two multiplying factors have been chosen empirically based on preliminary results.

6.2.3 Measurement Setup

The camera used to acquire the lateral image of the thorax, was the same that was used in chapter 3. The camera was a Logitech^{TM} C920, configured with its native resolution of 640x480 px at 24 fps with a YUY2 colour-space. For this study the focus and the exposure were blocked to a fixed value, being the exposure fixed to prevent frame loss due to the internal auto-exposure algorithm of the camera. The remaining parameters of the camera were set to its default value.

The same reference system used in the previous chapters was used in this study to validate the proposed method. The system was the RespiBand^{TM} wearable RIP system from BioSignalsPlux [94]. The respiratory signal from the reference method was acquired at 40 Hz and 12 bit resolution via Bluetooth, using a Serial Port Profile (SPP).

On Figure 6.3a, the disposition of the camera within the setup can be appreciated, where the camera was mounted on a tripod stand and placed approximately at a distance of 7-9 cm away from the thorax of the subject. Moreover, the camera was placed at an approximate height between the 7th and the 10th rib of the subject.



(a) (b)

Figure 6.3: a) Position of the camera within the setup, b) FOV of the camera and distance with the subject. Source: [31] (CC BY 4.0).

Due to the camera being placed within a few centimetres of the subject, it is crucial to take into account the FOV of the camera in order to characterise the effective area of vision. The FOV has 78° x 70.42° x 43.3°, of diagonal, horizontal and vertical angles respectively as it can be seen in Figure 6.3b. Given the FOV of the camera, it must be noted that this FOV is specified for a resolution with a 16:9 aspect ratio, but as the used resolution has a 4:3 aspect ratio, the horizontal field of view is automatically cropped to match the used resolution. Given the FOV, the aspect-ratio and the distance between the subject and the camera in Figure 6.3b, the effective area of vision of the thorax of the subject is comprised between 41.17 cm^2 and 68.07 cm^2, with a ratio of 86.37 px/cm and 67.18 px/cm.

Regarding the reference method, the RIP sensor was strapped around the thorax and placed beneath the pectoral muscle of each subject. Finally, in order to facilitate the performance of the tests, a desk was placed in front of the setup for the subject to rest its arms, and to comfortably perform the last test.

All the obtained data from the reference method and the video feed was stored in a similar fashion than in the previous chapter. For this reason, a custom ROS [172] program was used to save all the combined data from the RIP sensor and the camera into a *.bag binary format. As stated in the previous chapter, this format allows to synchronise the obtained data and to emulate the exact time when the data was captured (from either sensor), allowing a posterior validation that no data or frame was lost during acquisition. Moreover, by using the ROS emulated time, the latency of the algorithm can be assessed as well. The ROS version used in this chapter was the Melodic Morenia under Ubuntu 18.04 LTS. The laptop used to record and process the video feed from the camera and the data from the RIP sensor, had an Intel i7-4710HQ processor, an Nvidia GeForce GTX 850M graphics card and 8 GB of RAM.

6.2.4 Measurement Protocol

Fifteen healthy subjects, 8 male and 7 female volunteered for the study, with the anthropometric data for the volunteers summarized in Table 6.1. Each subject gave their written consent to participate in the study, which was performed in accordance with the principles of the Declaration of Helsinki [153], and all the measurements complied with the regulations of the Universitat Politècnica de Catalunya (UPC). All the measurements were performed in a controlled environment.

Table 6.1: Anthropometric data from the subjects expressed as mean ± SD.

Age [years]	Height [cm]	Weight [kg]	BMI [kg/m^2]
26.6 ± 3.54	170.47 ± 9.26	64.6 ± 11.97	22.09 ± 2.72

Table reproduced from [31] (CC BY 4.0).

Prior to any test, the subjects were asked to seat on the setup chair and to rest their arms comfortably on the desk located in front of the setup. The subjects were also asked to wear the strap with the RIP sensor around the chest. The distance between the subject and the camera, as well as the height of the tripod, was calibrated once for each subject in order to maintain approximately the same distance and height among all the measurements and subjects.

Four tests were performed in this study with the aim to validate the method under different constraints, two of the tests were aimed to validate the method against known respiratory frequencies, one test was performed with the aim to validate the method under free breathing conditions, and finally, the last test was aimed to validate the hypothesis of the quality index obtained from the modulus of the optical flow.

The two test aimed to validate the respiratory signal against known frequencies, involved the subject breathing at 0.1 Hz and 0.3 Hz for a duration of 3 minutes each. To help the subjects maintain the respiration at the target frequency, a custom visual aid similar with the one used in chapters 3 and 4 was developed. The visual aid consisted on a moving bar with 2/3 of the period for inhaling (ascending movement) and 1/3 for exhaling (descending movement).

Regarding the free breathing test, the test had a duration of 10 minute with no respiratory constraints imposed to the subject (the subject could breath freely). To guarantee that the subject was breathing at its natural respiratory frequency (without any external interference), each subject was asked to watch a documentary while breathing normally.

Finally, the last test was designed to test whenever the quality index could detect if the subject was talking or performing an activity, for this reason three different tasks where performed during the test: free breathing, reading out loud a text and solving a sudoku puzzle. The total length of the test was 10 minutes, while the duration of the different categories was split as depicted in the timetable in Table 6.2. To verify that the quality index could correctly determine whenever the subject was reading a text or performing a sudoku, a 2 minute period of free breathing was introduced between the two activities. The only constraint introduced in this test was that the subject had total freedom of movement while solving the sudoku puzzle, but had to remain silent and breath normally.

Table 6.2: Tasks performed and timetable of the quality index test

3 min	1 min	2 min	4 min
Free Respiration	**Read a Text**	**Free Respiration**	**Solve a Sudoku**

Table reproduced from [31] (CC BY 4.0).

In order to verify the alignment between the obtained respiratory signal from the video and the one obtained from the reference method, each subject was asked to perform an apnoea prior to each test. All the four tests were performed sequentially, being the first test the 0.1 Hz and the last the quality index test. Between each test a 3 minute interval was given to each subject to relax.

6.2.5 Signal Processing

After the raw respiratory signals have been obtained from the video feed by the means of the proposed method, and in order to compare the obtained signal with the one from the reference method, a series of signal processing steps have been applied using the Matlab (Version 2019b) and R (Version 3.6.0) software.

Respiratory Signal Normalisation

To compare both the respiratory signals obtained with the proposed method and with the reference method, either in terms of respiratory cycles or instantaneous frequency, a normalization of both signals had to be performed. The signal processing steps applied on both respiratory signals are the same ones that were applied in the previous chapters of this thesis, with slight modifications.

- First, in order to normalize the sampling frequency, both signals are interpolated at 80 Hz using a cubic spline.
- The interpolated signals are filtered using a zero-phase 2nd order Butterworth bandpass filter, with cut-off frequencies between 0.05 Hz and 0.6 Hz. This filter has different function depending on the source of the respiratory signal, in the case of the proposed method, the filter is used to remove high frequency noise produced by the optical flow computation. Regarding the reference method, the filter is used to remove artefacts from the thoracic band.
- Finally, to normalize the amplitude of both signals, a compression between -1 and 1 using a non linear function based on the arctangent [155] is applied. Equation 6.7 shows the mathematical definition of the applied transformation.

$$S_n[k] = \arctan \frac{S[k]}{\sqrt{\frac{\sum_{i=1}^{N}(S[i]-\bar{S})^2}{N-1}} * \sqrt{2}} \tag{6.7}$$

where $S_n[k]$ is the resultant signal after the non-linear function (normalised respiratory signal), $S[k]$ is the filtered respiratory signal and $\bar{S}$ is the mean of $S[k]$.

In Figure 6.4 the previous normalization steps can be appreciated, as it depicts the raw respiratory signal obtained with the proposed algorithm and the signal obtained after the normalization process.

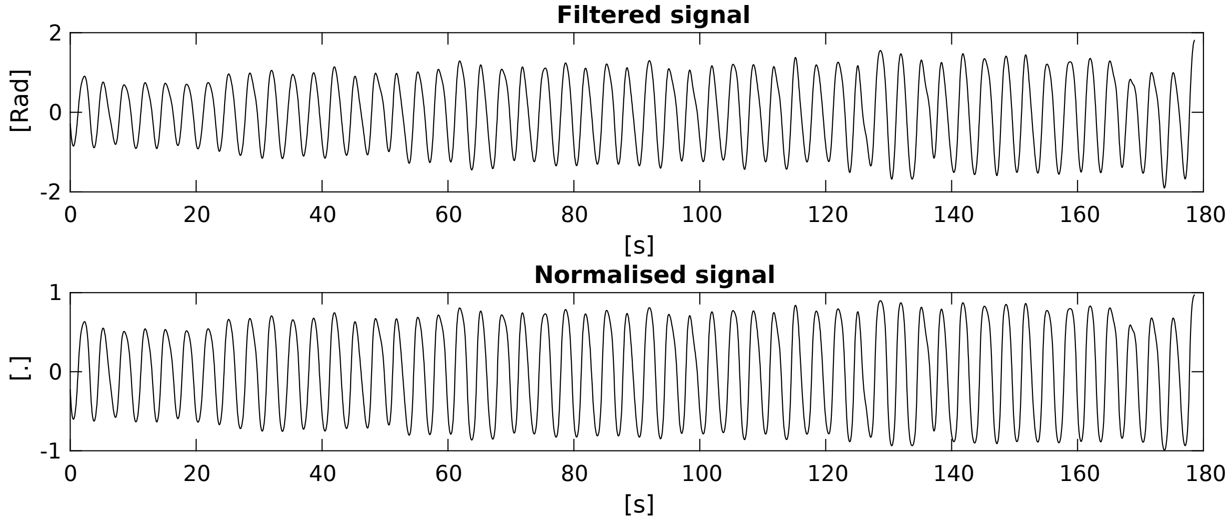


Figure 6.4: The first plot depicts the filtered respiratory signal from the proposed method, while the second plot corresponds to its normalized version. Both signals have been extracted from the 0.3 Hz test. Source: [31] (CC BY 4.0).

Respiratory Instantaneous Frequency Estimation

For the instantaneous frequency estimation, in this chapter a different approach has been used instead of the Hilbert transform, as the synchrosqueezing transform [177] has been used instead to compute the IF. The synchrosqueezing transform has been extensively used to extract the IF from different physiological signals such as ECG and respiration [178, 179, 180], as well as used for more advanced signal processing techniques such as to separate components from non-stationary signals [181]. One of the main benefits from using this transformation instead of the Hilbert transform, is that it does not require the signal to be quasi-periodic, moreover, the synchrosqueezing transform presents higher resolution in the time-frequency domain than the Hilbert transform. The synchrosqueezing transform was defined by Daubechies *et al.* [177] in 2010, as a frequency relocation method based on the continuous wavelet transform (CWT) to obtain the IF of arbitrary waveforms. A more in-depth explanation, as well as an exposition of the different mathematical definitions needed to compute the synchrosqueezing transform, is performed in the next chapter of this thesis (chapter 7, section 7.2.1).

As the synchrosqueezing transform is a 2D time-frequency representation, and the aim of this chapter is to compare the IF obtained from the respiratory signal for both methods, the component which presents the maximum spectral power (ridge) has been selected as IF. All the synchrosqueezing transform computation and the ridge selection has been performed using the tools provided by the Matlab software, more specifically the functions *wsst* and *wsstridge*. To ensure that no "hops" between adjacent frequency components with similar spectral power do not occur and add errors to the obtained IF, a penalty of 20 in the ridge computation has been put in place.

Respiratory Cycle Extraction

The respiratory cycle (RC) has been obtained using the same procedure than in the previous chapters (3 and 5). The following steps summarise the procedure described in chapter 3 section 3.2.4.

- A threshold based on the percentile 65 was computed in both signals.
- Using the previous threshold, the intersection with positive slopes is obtained from both methods.
- Finally, the distance between consecutive slopes is computed to form the RC series.

To prevent errors in the RC series, and as performed in the previous chapters, only the cycles that have been detected in both methods have been added to the series.

6.2.6 Performance Characterisation

In this study three different tests have been performed, and for this reason, and in order to characterise and evaluate the performance and error of each test, three different methodologies have been proposed as well. For the constant breathing tests, involving the 0.1 Hz and 0.3 Hz tests, only the instantaneous frequency has been evaluated in terms of error and by the means of two Bland-Altman (BA) plots [157] for each frequency. For the free breathing test, both the RC series and the IF have been obtained, and have been evaluated in terms of accuracy in the cycle detection, error between series and BA analysis for both the RC series and the IF. Finally for the quality index test, a set of statistical methods have been used to asses its performance, as well as to stablish the relationship between the sliding window length and the performance of the quality index. The following subsections summarize the different indicators used and how they have been obtained.

Detection Assessment

In the free breathing test, in order to characterise the accuracy of the respiratory cycle detection on both methods, the same procedures and methodologies presented in chapter 3 section 3.2.4 have been used. For the cycle by cycle comparison the sensitivity (SEN) and predictability of positive value (PPV) indicators have been computed as well. In this study nevertheless, even though the same definition of TP, FP and FN from the previous chapters are used, those indicators are only used for the computation of the SEN and PPV and are not reported in the analysis.

Error Assessment

For the free breathing test, the same error definition, as well as the error indicators, from chapter 3 have been used (MAE and MAPE) to characterise the error between the respiratory cycle series obtained from the proposed method and the one obtained from the reference method.

Regarding the Bland-Altman (BA) analysis, for the constant breathing test, two BA plot have been computed either for the mean IF and the SD IF for each one of the two tests. On the other hand for the free breathing test, three BA plot have been obtained: one for the respiratory cycles, where as well as in the previous chapter, no distinction between subjects has been made and all the individual cycles have been used, and two additional BA plots have been computed for the mean IF and SD IF. Regarding the BA involving the IF, an Anderson-Darling test (ADT) [173] was performed to verify that the differences had a normal distribution. As the results from the test showed that the differences did not present a normal distribution, the median and the 95 % interval (percentiles 2.5 and 97.5) has been used in the IF BA analysis.

To further characterise the error between methods, either for the two constant breathing tests and the free breathing test, the SDE has been computed on a sample by sample basis for the IF signal and the RC series in the case of the free breathing test. As in the case of the BA plot analysis, an ADT test has been applied to the SDE for both the IF and RC signals. The SDE of the IF did not present a normal distribution with a $p < 0.05$ discarding the null hypothesis. On the other hand, the SDE of the RC series showed a $p > 0.05$ indicating that the samples adjust to a normal distribution.

Finally, the Spearman correlation has been computed for the RC series and the IF signal from the free breathing test, between the proposed method and the reference method on a sample by sample basis .

Statistical Performance

Prior to performing any statistical test on the quality index signal, the different pieces that conform the test have been manually cropped into three different categories: Free breathing, Reading and Activity. To encompass all transitions from one category to the other, an offset has been added to the beginning and the end of each piece. Taking into account this offset, the free breathing piece is comprised between the 0.5 minute and 2.5 minute mark, the Reading section between the 2.5 and 4.5 minute mark and finally the Activity category is encompassed within the 5.5 and 10 minute mark. It has to be noted that while the Free breathing and the Reading category both have a 2 minutes length signal, the Activity piece has 4.5 minutes of signal length. The difference in the length between categories is due to the unpredictable nature of the movements inside the Activity piece, hence, to ensure that at least two movements are contained within the piece, its length is increased.

To perform the different statistical tests, once each category has been correctly cropped and labelled, the mean amplitude of the quality index of each piece is obtained. In order to verify which statistical test is more suited for the analysis of the different categories, an ADT test was performed to assert if the categories presented a normal distribution or not. The results showed that the Free breathing category did not present a normal distribution, hence non-parametric statistical tests have been used. The quality index pieces for each category have been compared between each other, by the means of the Kruskal-Wallis non-parametric test [182] and the Nemenyi post-hoc test with a Tukey distribution. The statistical tests have been performed with the aid of the PMCMR R package [183] under the R software.

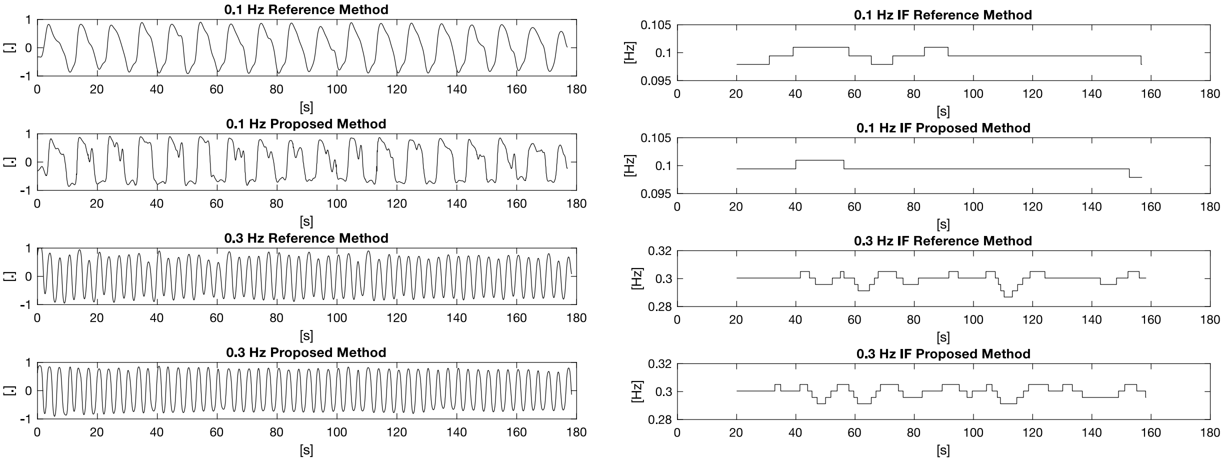
Finally, to test the ability of the quality index to determine wherever the subject is breathing normally, reading or performing an activity, an area under the curve (AUC) has been computed using the *perfcurve* Matlab function. The *perfcurve* Matlab function uses a cross-validation algorithm to asses the ratio between true and false positives, for this reason, the data has been relabelled as ”1” for the free breathing pieces of the test, where the signal quality is good, and ”0” for the rest of the two categories where the signal quality is bad.

6.3 Results

6.3.1 Constant Breathing Tests

The following results show the analysis in terms of the estimation of the instantaneous frequency of the constant breathing tests (0.1 Hz and 0.3 Hz), between the proposed method and the reference method.

Signals



(a) (b)

Figure 6.5: Respiratory signals (a) and its respective IF (b) from both the reference method and the proposed method for the constant breathing tests. Source: [31] (CC BY 4.0).

Figure 6.5 depicts a comparison between the obtained respiratory signals from both methods (Figure 6.5a) and its respective IF (Figure 6.5b) for both 0.1 Hz and 0.3 Hz. Moreover, as it can be seen in Figure 6.5b, for both tests the IF is centred around 0.1 Hz and 0.3 Hz respectively for both methods. It can also be seen in Figure 6.5 a high concordance between methods, either for the respiratory signals and the IF.

For the 0.1 Hz test (Figure 6.5a), the respiratory signal from the proposed method exhibits small ripples (high frequency components) if compared with the reference method. The source of the ripples could be due to the subject not able to continue inhaling or exhaling air, as the respiratory frequency of 0.1 Hz produces a very long respiratory period.

Table 6.3: Mean IF and SD IF expressed as median and IQR [25 ; 75] for the constant breathing tests.

Test	Mean IF RM [mHz]	SD IF RM [mHz]	Mean IF PM [mHz]	SD IF PM [mHz]
0.1 Hz	99.79 [99.64 ; 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]

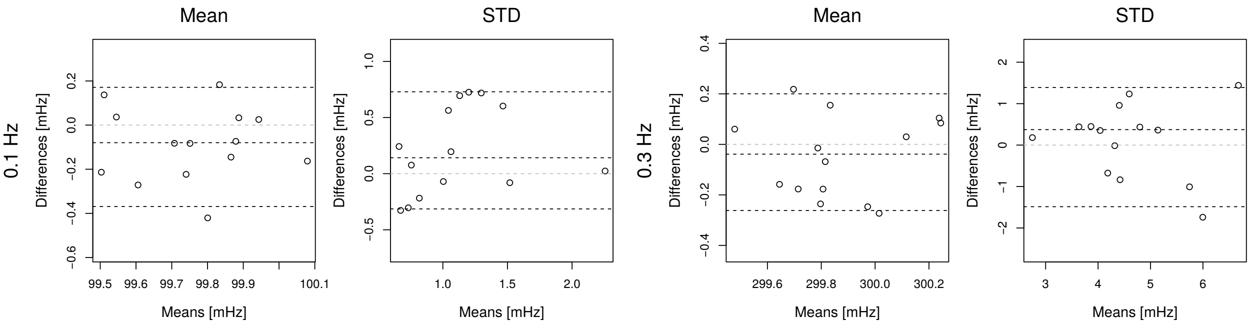
RM stands for Reference Method and PM stands for Proposed Method. Table reproduced from [31] (CC BY 4.0).

Table 6.3 shows the mean and SD of the IF obtained from the proposed method and the reference method for all subjects. As the IF results do not present a normal distribution, either the mean and the SD IF are presented as median and interquartile range (IQR). The results on this table show the base frequency for both methods (either in mean and SD) of the obtained IF. It can also be noted that both methods present practically the same median for both tests either in mean and SD IF.

Performance

Figure 6.6 depicts the IF BA plots for constant breathing tests. Figure 6.6a shows the BA plot for the 0.1 Hz test, where the first plot shows the comparison between the mean IF obtained with the reference method versus the proposed method, and the second plot shows the comparison between the SD IF obtained from the reference method versus the proposed method.

Analogously, Figure 6.6b shows the BA plot for the 0.3 Hz test. It has to be noted that each point in the BA plots represents the IF comparison (mean or SD) for a given subject, where no segregation between subjects has been made. Nevertheless, one outlier was removed from each one of the tests with the following criteria: if the differences between methods exceed 1.5 times, smaller or greater, the 95 % reference interval, this sample is removed from the BA plot. This criterion is established to correct the errors produced in the IF computation by the synchrosqueezing transform, when excessive jumps in the base frequency appear.



(a) (b)

Figure 6.6: Bland-Altman plots of the Mean and SD IF for the constant breathing tests, where a) corresponds to the 0.1 Hz test and b) to the 0.3 Hz test. Source: [31] (CC BY 4.0).

Table 6.4: SDE results expressed as median and IQR [25 ; 75], and numerical results of the BA plots in Figure 6.6 expressed as median and 95 % reference interval [2.5 ; 97.5], for the constant breathing tests.

Test	SDE [mHz]	BA Mean [mHz]	BA SD [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]

Table reproduced from [31] (CC BY 4.0).

Table 6.4 contains the standard deviation of the error of the IF between the reference method and the proposed method expressed as median and IQR [25 ; 75], as well as the median and 95 % reference interval of the different BA plots in Figure 6.6, for both the constant breathing tests.

6.3.2 Free Breathing Test

The results in this subsection are regarding the performance in terms of respiratory cycle extraction and IF estimation of the free breathing test.

Signals

Figure 6.7 depicts an example of the obtained signals from the free breathing test, where the normalised respiratory signals, RC series and IF signals can be appreciated. On the first plot regarding the respiratory signals from both methods, it can be seen that both signals are practically overlapping with small differences between each other, indicating a high concordance between methods. In the same fashion, the RC series present the same behaviour with both signals on top of each other, following the same temporal evolution with only small discrepancies in the number of periods. Regarding the IF signal on the last plot, in the same way than the RC series, the IF from both methods are completely on top of each other, following the same temporal evolution.

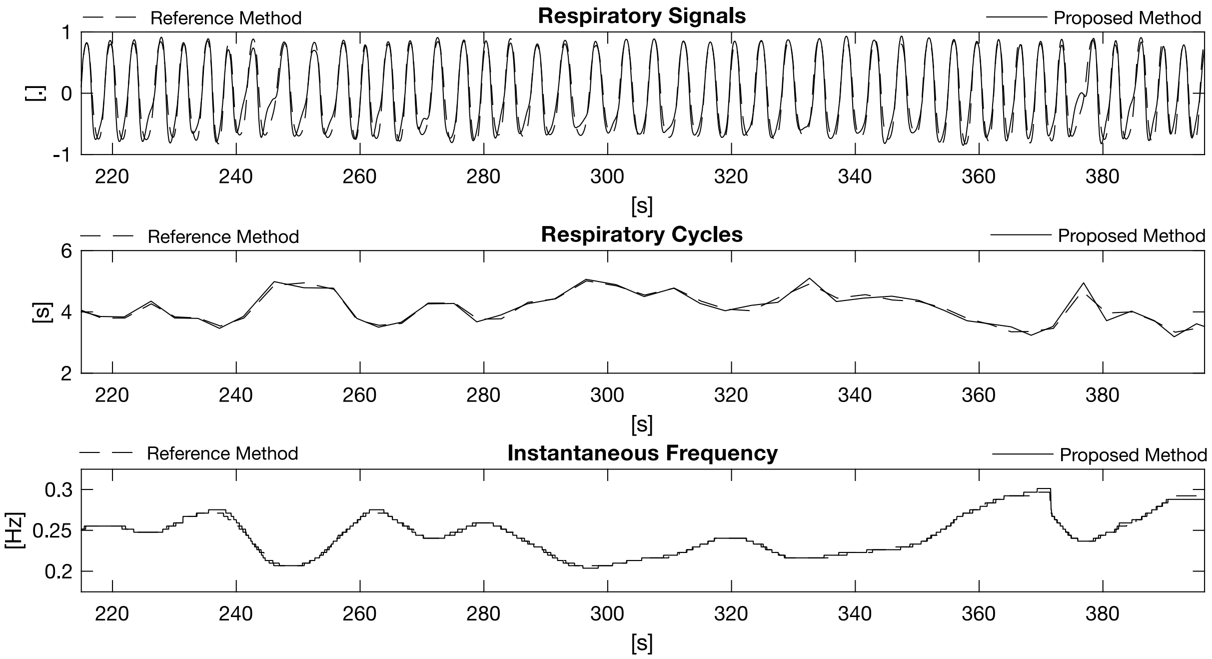


Figure 6.7: Example of the obtained signals for the free breathing test. The first plot depicts the normalised respiratory signal, the second plot corresponds to the RC series and the third plot depicts the IF. Source: [31] (CC BY 4.0).

Table 6.5: Mean IF and SD IF expressed as median and IQR [25 ; 75] for the free breathing test.

Mean IF RM [mHz]	SD IF RM [mHz]	Mean IF PM [mHz]	SD IF PM [mHz]
238.42 [201.09 ; 263.89]	36.52 [22.98 ; 44.88]	241.44 [212.49 ; 268.3]	38.10 [22.6 ; 45.68]

RM stands for Reference Method and PM stands for Proposed Method. Table reproduced from [31] (CC BY 4.0).

Table 6.5 shows the mean and SD IF results for the free breathing test for both methods expressed as median and IQR [25 ; 75]. It can be seen that the median for both methods is approximately 0.24 Hz, which is within the range of 14 to 20 breaths per minute corresponding to the typical respiratory frequency at rest in adult population [38]. As the free breathing test was performed so each subject could breath at his/her own base frequency, it can be seen that the IQR of the mean IF increases if compared to the constant breathing tests. Moreover, the same increase can be seen in the median of the SD IF. Finally, it must be noted that both methods yield similar values either for the mean IF and the SD IF.

Performance

Regarding the performance of the respiratory cycle series, Table 6.6 shows the obtained results for the cycle detection, while Table 6.7 summarizes the SDE and correlation results.

For the cycle detection results, the SEN and PPV indicators (Table 6.6) have been computed for each subject in order to determine the accuracy of the proposed method in the RC detection. The SEN results are presented as median and IQR range as a result of very few false negatives in the cycle detection, which in return caused that the aggregated SEN series, one SEN value per subject, to not present a normal distribution. The PPV values on the other hand, are presented as mean ± SD. The MAE and MAPE results are also shown in Table 6.6, and in the same fashion as the SEN and PPV indicators, they have been also computed once for each subject on a sample by sample basis. Both results are reported as mean ± SD.

Table 6.6: Performance and error indicators of the RC series for the free breathing test.

SEN (%)	PPV (%)	MAE [s]	MAPE (%)
100 [99.18 ; 100]	94.7 ± 4.16	0.22 ± 0.1	4.87 ± 1.57

The SEN is expressed as median and IQR [25 ; 75], while PPV, MAE and MAPE are expressed as mean ± SD. Table reproduced from [31] (CC BY 4.0).

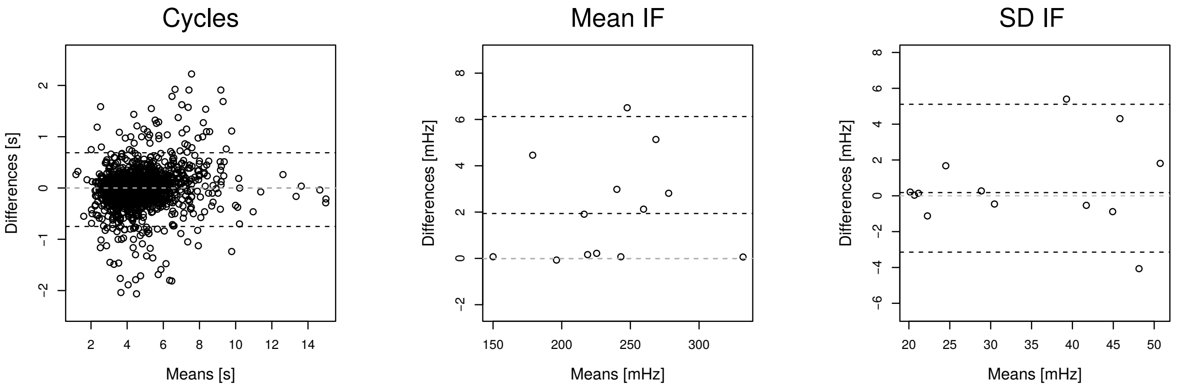
Given the SDE and correlation results in Table 6.7, the SDE of the RC series results are expressed as mean ± SD while the SDE FI results are expressed as median and IQR as they do not present a normal distribution. The Spearman correlation results on the other hand are both expressed as mean ± SD for either the RC and the IF. It must be noted that the correlation results, either for the RC series and the IF signal, are greater than 0.9 indicating a high correlation between methods.

Table 6.7: SDE and Spearman correlation results of the free breathing test.

SDE Cycles [s]	SDE IF [mHz]	Corr. Cycles	Corr. IF
0.34 ± 0.16	12.95 [4.38 ; 29]	0.94 ± 0.03	0.9 ± 0.09

The RC SDE and both correlations are expressed as mean ± SD, while the SDE IF is expressed as median and IQR [25 ; 75]. Table reproduced from [31] (CC BY 4.0).

The first BA plot in Figure 6.8 shows a cycle by cycle comparison between the obtained respiratory cycles with the reference method and the ones obtained with the proposed method, where all the individual cycles have been taken into account and no distinction between subjects has been made. A total of 2007 cycles have been analysed in this BA. The two remaining plots analyse the mean and SD IF respectively estimated from the reference method and the proposed method, where each sample represents a subject and no segregation between subjects has been made. Two outliers from the IF BA plots have been removed using the same criteria than in the constant frequency tests. No outliers were removed from the Cycle BA plot.



(a) (b) (c)

Figure 6.8: Bland-Altman analysis of the free breathing test, where a) depicts all the detected cycles, b) the mean IF and c) the SD IF. Source: [31] (CC BY 4.0).

Finally, Table 6.8 summarises the numerical results of the BA plots in Figure 6.8. The differences of the cycle BA are expressed as mean ± SD, while the differences for the IF BA plots are expressed as median and 95 % reference interval, as they do not adjust to a normal distribution. Regarding the mean of the cycle BA differences, it has to be noted that the cycle differences presented a negligible mean of the order of μs, which could be attributed to statistical fluctuations, hence, 0 s is reported instead.

Table 6.8: Numerical results of the BA plots in Figure 6.8, where the cycle BA differences are expressed as mean ± SD and both the Mean IF and SD IF differences as median and 95 % reference interval [2.5 ; 97.5].

Cycle [s]	Mean IF [mHz]	SD IF [mHz]
0 ± 0.35	1.94 [0 ; 6.13]	0.17 [-3.14 ; 5.11]

Table reproduced from [31] (CC BY 4.0).

6.3.3 Quality Index

The following subsections show the results of the quality index evaluation, as well as the results of the study between the relationship of the sliding window length and its performance.

Signals

In Figure 6.9, the first two plots show the respiratory signal obtained from the reference method and the respiratory signal obtained from the proposed method. The third plot shows the obtained quality index with the different phases of the test superimposed, where the first 3 minute piece corresponds to free breathing, the next minute the subject was instructed to read out loud a text, followed by 2 minutes of free respiration and finally 4 minutes of solving a sudoku puzzle (perform an activity involving movement from the subject). It must be noted that for the statistical analysis, as previously stated, the different pieces have been manually labelled to have 2 minute each for the free breathing and reading and 4 minutes for the activity piece.

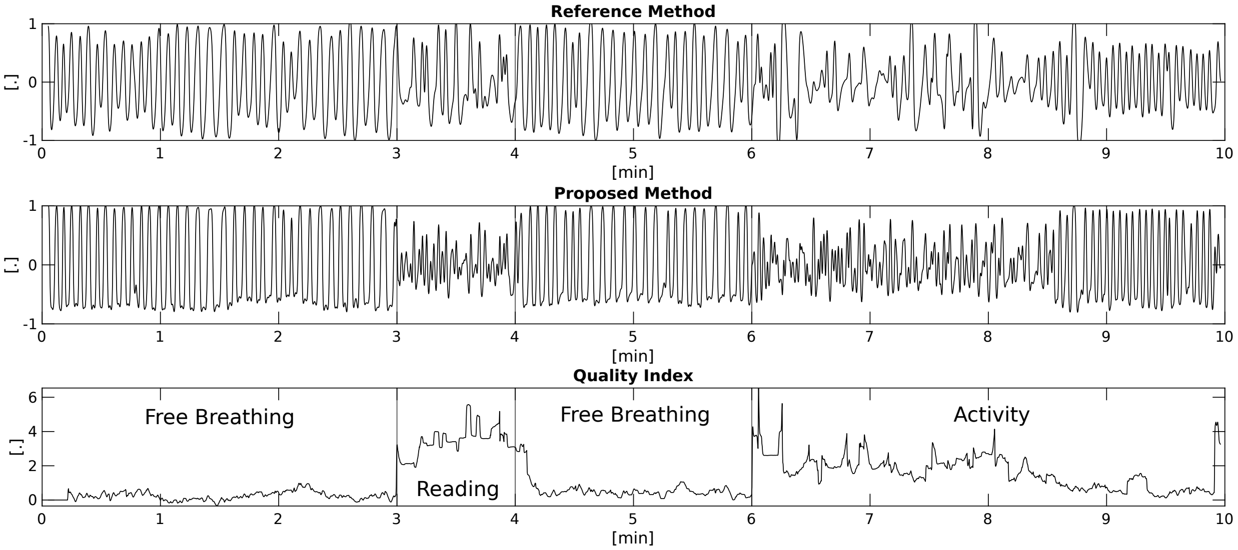


Figure 6.9: The first two plots depict the respiratory signal obtained respectively from the reference method and the proposed method. The third plot depicts the obtained quality index (using a 10 s window), with the different regions of the test labelled on top. Source: [31] (CC BY 4.0).

If the first two plots in Figure 6.9 are compared with the quality index, it can be clearly seen that when the subject is instructed to read a text, the respiratory signal in both methods becomes more chaotic and looses its periodicity, on the contrary, when the subject stops reading the text, the respiratory signal falls back to its periodic state. Regarding the quality index, low values can be appreciated when the respiratory signal is periodic. On the other hand, when the subject starts to read, the values of the quality index increase indicating a degradation of the respiratory signal. Another characteristic that can be seen in the quality index plot in Figure 6.9, is the abrupt base-line shift produced by the different transitions between free breathing and reading, and free breathing and performing an activity.

Performance

Table 6.9 shows the statistical test results (post-hoc) of the analysis between sections of the quality index test, where different lengths for the sliding window have been used. The results of the Kruskal-Wallis test comparing all the pieces between each other, show that for window lengths between 5 s to 30 s the p-values are inferior to $7.748 \cdot 10^{-}5$, which is the maximum p-value corresponding to the 30 s window. For the 1 s window, on the other hand, the returned p-value was $3.899 \cdot 10^{-}4$.

Table 6.9: Post-Hoc results and AUC for the quality index test.

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. Table reproduced from [31] (CC BY 4.0).

The post-hoc results in Table 6.9, which evaluate the statistical interaction between the different categories, show very significant differences $p < 0.001$ in the interaction between free breathing and reading for all window lengths, significant differences $p < 0.05$ in the interaction between free breathing and performing an activity for the window lengths higher than 5 s, with no significant differences for the window length of 1 s, and finally for the interaction between reading and performing an activity, no significant differences can be seen for all windows except for the 5 s length which shows a $p < 0.05$. Regarding the area under the curve analysis, it can be seen that all window lengths present results higher than 0.8, being the maximum result the one yielded by the 5 s window.

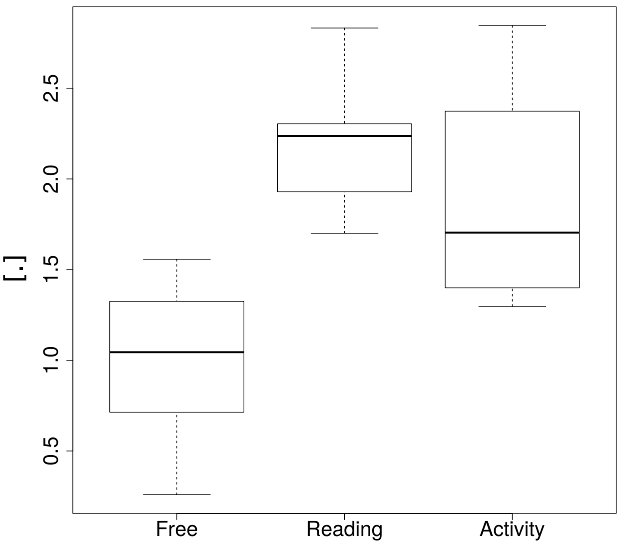


Figure 6.10: Boxplot representation of the free, reading and activity pieces from the quality index test (using a 10 s window). Source: [31] (CC BY 4.0).

Figure 6.10 shows a box plot corresponding to the 10 s window quality index. This window length was chosen as the default, as the results in Table 6.9 yield for this window length the lowest p-value in the free breathing versus activity in the post-hoc test, moreover a high value of AUC is also yielded for that given window length.

6.4 Discussion

6.4.1 Constant Breathing Tests

Given the signals showed in Figure 6.5a for the constant breathing tests, small ripples can be appreciated in the 0.1 Hz test for the respiratory signal obtained with the proposed method. This distortions in the signal can be due to forced breathing at slow frequencies, and produced when the subject can no longer inhale or exhale air, as the subject is waiting for the visual aid to start its ascend or descend, involuntary movements of the thoracic-abdominal region are produced, which originate the small notches in the respiratory signal. This same notches can be appreciated in a smaller scale in the reference method, but as the proposed method measures the changes in both the intercostal and abdominal muscles, this effect is magnified.

Regarding the impact of the aforementioned distortions, as they produce high frequency harmonics in the respiratory signal, this harmonics are also present in the IF signal. By using the synchrosqueezing transform instead of other methods to compute the IF, and due to the property of the synchrosqueezing transform to concentrate the spectral power into narrow bands, the impact of this type of distortions is greatly minimized as it can be seen in Figure 6.5b.

In Table 6.3 very similar results in the median IF can be appreciated between the reference method and the proposed method for both the 0.1 Hz and 0.3 Hz test, being the latter the one that presents a higher IQR range from the two tests, this could be caused by small errors in the synchrosqueezing estimation. Regarding the SD IF for both methods and tests, all present similar results between each other with comparable median and IQR.

Given the BA plots in Figure 6.6, even though no apparent trend can be appreciated in the four BA plots, as the samples do not present a normal distribution, it cannot be discarded that any of the BA plots do not contain any bias. In all the plots it can be appreciated that the zero value is inside the 95 % reference interval, with low median values and narrow 95 % reference interval (Table 6.3), which indicates a good agreement between methods. Given the SDE results shown in Table 6.3, both methods yield SDE results within the margin of the SD IF in Table 6.4 which indicate a low error between methods. The SDE results in Table 6.3 also show low median SDE and narrow IQR for both tests, indicating a good performance of the proposed method in terms of respiratory signal extraction at constant frequencies.

6.4.2 Free Breathing Test

The results of the free breathing test shown in Table 6.6, regarding the accuracy of the cycle detection, show a median sensitivity of 100 % with a very narrow IQR and a mean PPV of 94.7

% with low SD values. The median SEN value can be explained as the majority of the respiratory cycles were detected correctly in both methods, thus very low false negatives were detected. The PPV result on the other hand, indicate that there is a small number of false positives as well. One possible source of the false positives could be due to small movements of the subject, or even small differences between the thoracic and abdominal respiratory movements [32, 33], which if it is taken into account the measuring point of the proposed method and the measuring point of the reference method, the latter only measures the thoracic movement while the proposed method measures both thoracic and abdominal regions. Overall, the results in Table 6.6 for the SEN and PPV show a high successful rate in the respiratory cycle detection.

Regarding the MAE and MAPE results in Table 6.6, small values can be seen for the MAE indicator either for the mean and SD, indicating a low error between methods in the RC series. The MAPE results also yield values consistent with the MAE results with a mean of 4.87 %, which also indicates a high agreement between methods for the RC series. Given the SDE results in Table 6.7, the RC series SDE shows low values with mean of 0.34 s and SD of 0.16 s indicating a low error between methods. The correlation results for the RC series show a mean value of 0.94 with a small SD thus indicating a high agreement between method for the RC series. Overall all the indicators show a high agreement between the RC series obtained with the proposed method and the reference method.

The results for the instantaneous frequency in Table 6.5 for the mean IF show an approximate median value for both methods around 0.24 Hz, which as previously indicated, is within the range of 14 to 20 breaths per minute present in adult population at rest [38]. The results for the SDE IF present a median of 12.95 mHz with an increase IQR if compared to the constant breathing tests, which is caused by the greater variability between subjects and the increased length of the test. Given the correlation results for the IF, a mean of 0.9 and an SD of 0.09 is reported, which are in agreement with the SDE results. Either the SDE and the correlation indicate a high agreement between both methods in the case of IF analysis.

Regarding the BA analysis of the free breathing test, the cycles BA plot in Figure 6.8a and its numerical results in Table 6.8, show a high agreement between methods with a negligible bias and a very narrow limits of agreement, with no bias or trend present in the plot. The mean IF BA plot in Figure 6.8b, even though no apparent trend can be appreciated, as the samples do not present a normal distribution it cannot be assured that no trend is present in the BA plot. Moreover, a median bias of 2 mHz between the differences is reported in Table 6.8 with a narrow 95 % reference interval. It has to be noted that this bias if compared to the mean IF in Table 6.5, yields less than 1 % of the median IF. The source of this small bias can be due to small discrepancies between the theoretical sampling frequency of the camera and the real sampling frequency, which in return can produce misalignments between methods and errors in the IF estimation. For the SD IF BA plot in Figure 6.8c, no apparent trend can be appreciated either. A near zero median of the differences for the SD IF BA plot is reported in Table 6.5, with a narrow 96 % reference interval. Both IF BA indicate a high agreement between the IF obtained from the proposed method and the one obtained from the reference method.

6.4.3 Quality Index

Taking into account the last plot in Figure 6.9, the effects of the window length can be directly appreciated in the response time of the transitions between "bad quality" signals to "good quality" ones. As it can be seen, even when the respiratory signal has become quasi-periodic after the reading period, the quality index still yields high values for a certain period of time, this is due to the sliding window still containing a certain amount of samples of the "bad quality" signal, and until those samples are removed, the quality index does not decrease. An example of this behaviour is depicted in Figure 6.2, where the effect of the "bad quality" samples can be seen in the histogram. Even though this could be seen as a downside of the method, this provides an hysteresis which ensures that when the quality index values have decreased, the respiratory signal will have good quality, thus preventing false positives.

Regarding the statistical results shown in Table 6.9, a clear dependence between the length of the window and the quality of the signal is reported. As an example, whenever the window length decreases, the statistical results show lower p-values. This behaviour is true until the sliding window is sufficiently small, that the contained samples from the modulus of the optical flow only correspond to a few respiratory periods, in which case there is not enough information inside the window to accurately discern between good and bad quality. This behaviour can be seen in Table 6.9 if a 1 s window length or less is chosen, as the p-value results increase and the AUC results decrease. Given the previous constraints, the chosen window length has to take into account both the specificity and response time as a trade-off. As an example, if the 10 s window is chosen, it reports a good trade-off between specificity and response time, with an AUC of 0.9533 and very significant differences ($p < 0.01$) and significant differences ($p < 0.05$) between free breathing vs reading and free breathing vs activity respectively. Furthermore, and as it can be seen in Figure 6.9, the response time is sufficiently small which does not delay in excess the transition between "bad quality" signal to "good quality" one.

Given the plots in Figure 6.10 for a 10 s window length, the median and IQR of the Free pieces is clearly below the Reading and Activity pieces, which is in concordance with the statistical results shown in Table 6.9. Regarding the other two pieces, both report similar median and IQR results, with an increased IQR can for the Activity piece if compared to the Reading piece. This could be due to the difference in length between the two pieces, but more importantly as the Activity piece involved the subject moving freely, this increased IQR could be caused by the variability in the movements from the subjects.

Overall, both the statistical results and the box plot for the quality index show a good performance in terms of detecting good and bad quality signals from the modulus of the optical flow.

6.4.4 Limitations of the Study

Some limitations were present in this study, being the first the location of the reference method. While the reference method performed the measurement on the chest, the proposed method measured the intercostal and abdominal region, which can produce different results depending on how the subject is breathing. Even though the results do not show a significant error, either in the constant breathing test or the free breathing test, a new validation using two RIP sensors, one in the chest and the other in the abdominal region, has to be performed to discard possible errors due to abdominal respiration. The second limitation of this study was the narrow age gap of the subjects, as the respiratory mechanics in adult population can change, specially for elder people. To make the study extensible to a wider range of adult population, measurements with a wider age interval should be performed as well.

The last limitation of the study was that the test were performed in a very controlled environment, were no changing lighting conditions or external vibrations were taken into account for the study. While most of the lighting issues could be solved by using an IR camera (as the algorithm would still be valid), a validation with external vibrations should still be performed, as they will affect the obtained optical flow, hence, the respiratory signal.

6.5 Conclusions

In this chapter a new method to obtain the respiratory signal at the thoracic-abdominal region from a lateral perspective is presented. The proposed method uses the phase of the optical flow to obtain the respiratory signal of the subject while using the modulus to obtain the quality of the respiratory signal. To validate the algorithm a thoracic RIP sensor has been used as a reference system.

The constant breathing tests results show a high agreement between the proposed method and the reference method, indicating the viability of the proposed method to retrieve the respiratory signal at a constant frequency, and more importantly indicating the precision in the IF estimation.

The results for the free breathing test indicate a high sensitivity in the respiratory cycle detection, yielding low error results and showing a high agreement in the BA differences. Regarding the IF for the free breathing test, the results yield a good agreement between methods with small SDE and small median differences in the BA plots.

In this chapter, the hypothesis that a quality index can be simultaneously obtained from the modulus of the optical flow has been validated, with results indicating a good specificity in terms of discerning between good and bad signal. Finally, the trade-off between window length and specificity has also been assessed in this chapter.

In general, and thanks to the simplicity of the method, the proposed algorithm can extract the respiratory signal with real-time constraints while simultaneously obtaining the quality of the respiratory signal, by using the phase and the modulus of the optical flow respectively.

Chapter 7

Synchrosqueezing transform implementation for real-time systems

7.1 Introduction

As introduced in the state of the art (chapter 2, section 2.4), there are no specific signal processing techniques, methods or methodologies for respiratory signal analysis. As the respiration is a biological signal, most of the current techniques applied to respiratory signals are based on the electrocardiogram (ECG) signal processing.

Nowadays, most of the algorithms and methods used to obtain metrics and insights from the respiratory signal, rely on time domain analysis. This is related to the traditional approach in medical practice on signals obtained with pneumotachography [39], where the main objective was to obtain the time between breaths (respiratory cycle) and the amplitude of the signal (tidal volume). More recently, another approach based on frequency-domain analysis can also be found in the literature. Frequency-domain analysis allows to extract information regarding the different harmonics inside the signal, which is a good indicator of the variability of the signal.

Nevertheless, there are some limitations to time-domain and frequency-domain analysis. For instance, time-domain analysis relies on complete respiratory periods to compute the breath to breath distance (respiratory cycles), hence, time-domain analysis would not yield any information until a complete respiratory cycle is obtained, making this type of signal processing methods not viable if a sample by sample analysis is required. Frequency-domain analysis on the other hand, require for the signals to be stationary for a correct spectral estimation of the signal. As the respiratory signal is non-stationary, the error produced depending on the used window and its length when estimating the mean respiratory frequency and its variability, would have a great impact in the obtained results. Moreover, only the mean variability and mean respiratory frequency can be obtained from this type of analysis.

There are some methods that use time-frequency representation instead of purely time-domain or frequency-domain analysis, that have been proven reliable in respiratory signal processing [179].

This methods rely on obtaining the instantaneous frequency (IF) of the signal, where the mean IF holds the information of the breathing frequency, and the standard deviation the variability of the respiratory signal.

The Hilbert transform is a classical method to obtain a time-frequency representation of a signal, moreover, some studies can be found in the literature that use such approach in ECG signal processing [166, 162]. This methodology has also been used in chapter 4 and chapter 5 of this thesis to obtain the IF of the respiratory signal. The use of the Hilbert transform poses a series of limitations that have to be taken into account whenever is used in biological signals. The first limitation is that it can only be applied if the signal is quasi-stationary, meaning that if the signal contains transitory periods or non-periodic pieces, the resultant IF will certainly contain errors [184]. Another limitation of this method is that it is computationally expensive, and as such, time-frequency representation using this methodology is limited to post-processing techniques, were there is no limitation in the time expended in processing a sample.

A new methodology to obtain the time-frequency representation was presented in [177] by Daubechies *et al.* in 2010 as a tool to obtain the IF of an arbitrary waveform. The Synchrosqueezing transform, as defined by Daubechies *et al.* in [177], is a relocation method based on the CWT, where the CWT is "sharpened" by relocating the spectral power density from nearby frequencies to a local maximum. To summarize, the Synchrosqueezing transform allows a precise time-frequency representation of a signal, independently if it is stationary or not or even if it contains more than one harmonic.

This type of time-frequency analysis can be found in the literature applied to ECG [178], EEG [185] and even for respiratory signal analysis [179, 186], where the instantaneous variability of these signals is studied. Moreover, in this thesis this type of approach has also been used in chapter 6 to obtain the IF of the respiratory signal. The main drawback of this method is that it is computationally expensive, as it requires to compute the CWT (create one wavelet for each one of the frequencies to study, and convolute the result with the signal being analysed), and after that, perform the frequency relocation. As another drawback, if the signal being analysed contains a large amount of samples, or there are to many frequencies in the CWT, the memory overhead would make this method not viable in current off-shelf computers, where the amount of memory and CPU/GPU computing capability are limited.

In this chapter a novel implementation of the synchrosquezing transform is proposed, the novelty lies in using matrix operations that can be performed in a parallel fashion using the GPU. The framework used to implement the algorithm was pytorch, which is a GPU accelerated framework for deep learning. The proposed implementation also uses an advanced strategy to only compute all the needed wavelets once, and in the frequency domain. This approach is important in order to prevent memory overheads, as there is no need to instantiate all the wavelets at each iteration, moreover, as the wavelets are already computed in the frequency domain, the computational cost required to initiate the algorithm is greatly reduced.

The proposed implementation has been optimised to perform in real-time, specially for systems that use non-contact methods to obtain physiological variables. Moreover, the proposed approach is to use a sliding window of known length (to asses the computational cost) in order to obtain the IF evolution of an arbitrary signal in a sample by sample basis, thus making the algorithm viable on low-end systems and in real-time applications.

7.2 Materials And Methods

In this section an overview of the synchrosqueezing transform method [177] presented by Daubechies *et al.* is performed. The proposed implementation uses the mathematical definitions in [177] as a reference to obtain the Synchrosqueezing transform, as well as a series of mathematical definitions proposed by Stéphane *et. al* [187] regarding the wavelet generation and CWT computation.

7.2.1 Formal Synchrosqueezing Definition

In this subsection the formal definition of the synchrosqueezing transform is studied, as well as an overview of the different implications of each step from a computational point of view.

Wavelet Definition

As the synchrosqueezing transform relies on wavelets to obtain the time-frequency representation, the first step to obtain such transformation lies on how to obtain a wavelet. For this implementation, the complex Morlet wavelet (Gabor wavelet) [187, 188] has been used. In Equation 7.1 the temporal representation of the Morlet wavelet can be appreciated, whereas in Equation 7.2, the frequency domain Morlet wavelet can be found.

The Morlet wavelet is defined as a pure tone modulated by a Gaussian [187] as it can be seen in Equation 7.1, whereas $e^{i\sigma t}$ is the complex tone and $e^{-\frac{1}{2}t^2}$ is the Gaussian modulation.

$$\psi(t) = \pi^{-\frac{1}{4}} e^{i\sigma t} e^{-\frac{1}{2}t^2} \tag{7.1}$$

$$\psi(w) = \pi^{-\frac{1}{4}} H(w) e^{-(w-\sigma)^2/2} \tag{7.2}$$

where $\psi(t)$ is the temporal wavelet, $\psi(w)$ is the wavelet in the frequency domain, σ is the fundamental frequency of the wavelet, and $H(w)$ is the Heaviside function as the Morlet wavelet has to comply with $if:\ w < 0\ then:\ \psi(w) = 0$ in order to be considered an analytic function [187].

The definition of the Morlet wavelet in the frequency domain (Eq. 7.2), has been obtained from [188]. It has to be noted that while the temporal definition of the Morlet wavelet is complex, the frequency domain definition is only real. From a computational point of view, using the frequency definition offers a serious advantage, as the memory required to allocate a complex vector (wavelet) is twice as much the memory to allocate a real one. To comply with the admissibility condition imposed on this type of wavelet [189, 188], a $\sigma = 6$ has been chosen.

Continuous Wavelet Transform

The first step to compute the synchrosqueezing transform is to obtain the CWT of a signal. To perform this transformation, one wavelet is needed for each one of the frequencies that are going to be analysed. To compute all the needed wavelets the definition of the mother wavelet is presented in Equation 7.3 [187] either in the time domain and in the frequency domain. By the means of the mother wavelet, the same wavelet definition (Equations 7.1 and 7.2) can be scaled and/or shifted to obtain a wavelet that corresponds to a certain frequency. For the next steps a generic analytical wavelet is assumed, as the mother wavelet can be applied to any wavelet, all the expressions are defined in function of the mother wavelet instead of a particular wavelet definition.

$$\begin{aligned}\psi_s(t) &= \frac{1}{\sqrt{a}}\psi(\frac{-t}{a}) \\ \hat{\psi}_s(w) &= \sqrt{a}\hat{\psi}(aw)\end{aligned} \tag{7.3}$$

where ψ_s is the mother wavelet, ψ is the original wavelet and a is the scale at which the wavelet is being generated.

To compute the CWT from the mother wavelet, the signal being analysed has to be convoluted with each one of the generated wavelets, so a time-frequency representation can be achieved. In Equation 7.4 the CWT transform is shown [177, 187].

$$W_s(a,b) = \int s(t)\frac{1}{\sqrt{a}}\overline{\psi(\frac{t-b}{a})}dt \tag{7.4}$$

where $W_s(a,b)$ is the CWT transform for a given scale a at given time b and $s(t)$ represents the signal being analysed.

From a computational point of view, as the convolution in the time domain requires shifting one of the signals (mother wavelet) in respect to the other, and although nowadays making this kind of operations has become much more efficient due to the consolidation of deep learning and convolutional neural networks, it still requires a fair amount of processing power to compute temporal convolutions. On the other hand, and taking into account the property of the Fourier transform where a temporal convolution is transformed into a multiplication in the frequency domain, computing the CWT in the frequency domain is much more computationally efficient. In Equation 7.5, the CWT in the frequency domain is shown [187].

$$\hat{f}_s(w) = \frac{1}{2}\hat{s}(w)\sqrt{a}\hat{\psi}^*(aw) \tag{7.5}$$

where $\hat{f}_s(w)$ is the CWT transform and $\sqrt{a}\hat{\psi}^*(aw)$ is the Fourier transform of the mother wavelet. This approximation is only valid if $\hat{\psi}(w) = 0$ for $w < 0$.

Synchrosqueezing Transform

Based on the mathematical definitions by Daubechies *et al.* described in [177], if the wavelet ψ is concentrated only in the positive frequency axis, and given that $W_s(a,b) \neq 0$, a candidate instantaneous frequency can be computed following Equation 7.6.

$$\omega(a,b) = -\jmath(W_s(a,b))^{-1}\frac{\partial}{\partial b}W_s(a,b) \tag{7.6}$$

where $\omega(a,b)$ is the candidate IF and $W_s(a,b)$ is the CWT transform for a certain scale a at a certain temporal displacement b.

The previous equation describes how to obtain a candidate IF from the wavelet transform in the time domain. It has to be noted that this is not the synchrosqueezing transform as $\omega(a,b)$ is still in the time-domain plane, and its purpose is to obtain which frequencies are present in the CWT. If the signal were to be of the form $s(t) = Acos(wt)$ its candidate IF would be $\omega(a,b) = w$.

Finally, the synchrosqueezing transform is defined in [177] by Equation 7.7. Taking into account that the signal has been computed only at discrete scales a_k (binned), the synchrosqueezing transform will also be binned in ω_l intervals.

$$T_s(\omega_l,b) = (\Delta\omega^{-1}) \sum_{a_k |\omega(a_k,b)-\omega_l| \leq \Delta\omega/2} W_s(a_k,b)a_k^{-\frac{3}{2}}(\Delta a)_k \tag{7.7}$$

where $T_s(\omega_l,b)$ is the obtained synchrosqueezing transform, $\Delta\omega$ and Δa are respectively the increments between frequencies and scales.

To summarize, $T_s(\omega_l,b)$ will contain a version of the CWT where the spectral power density has been relocated to its local maxima, hence the time-frequency representation will be concentrated into narrow frequency bands, which have a more precise representation than the one obtained with the CWT.

7.2.2 Proposed Implementation

The aim of this implementation is to achieve real-time performance when computing the IF on a sample by sample basis, by computing the SST transform using matrices instead of loops for the linear operations. Nowadays Matrix operations can be easily parallelized in the GPU, thus decreasing the computing time for each operation. To achieve real-time performances, the proposed implementation tries to minimize the memory transactions between GPU and CPU by pre-loading all the essential matrices into GPU memory at the initialization stage. Another improvement, is that by performing most of the operations in the frequency domain, the performance of the algorithm is also increased.

The proposed implementation has three fundamental code functions: the code initialisation, the synchrosqueezing transform (SST) computation and the ridge extraction. In the code initialisation all the global variables as well as all the wavelets are allocated in GPU memory, so all the objects are always ready when a new operation takes place in order to minimise the memory transaction between GPU and CPU. In the SST computation, whenever a new sample arrives, the SST is computed for the refreshed window. Finally, the ridge extraction allows for a 1-D representation of the instantaneous frequency.

The framework used was the pytorch (version 1.4) library with CUDA (version 10.2) support, and executed under Python3 (version 3.6.9). In the following subsections a more in-detail explanation of the principal operations involved in the SST computation as well as the ridge extraction is shown.

Code Initialisation

The code initialisation is the first function that is called in the life-cycle of the proposed implementation when the class is instantiated. This function serves two purposes, the first one is to allocate in GPU memory all the required data that will be later used at each iteration, and to compute all the necessary wavelets to perform the CWT and the SST computation. The amount of scales used is a parameter that can be set at the class instantiation, as well as the sliding window length and the sample frequency, thus allowing a customization of the amount of wavelets used and to pre-allocate all the necessary data into GPU memory. This function is only executed once in all the life-cycle of the proposed implementation, meaning that all this operations are only performed once.

In order to generate all the wavelets needed to compute the CWT, the equation 7.2 and 7.3 are used. As the wavelets are generated in the frequency domain, and to avoid errors due to the Fourier transform, the amount of samples that will be used in the FFT of the signal are computed. To do so, the following equation has been used: $N_{fft} = ceil(log_2(win\_len \cdot Fs))$ where win_len represents the sliding window length in seconds and Fs is the sample frequency of the signal. N_{fft} is the nearest superior power of 2 number given the amount of samples in the sliding window, for example, if the signal is sampled at 30 Hz and the window length is 60 s, the amount of samples inside the window will be 1800, being 2048 the nearest immediately superior power of 2 number. For this reason N_{fft} samples are used to compute the wavelets in the frequency domain.

Once the amount of samples needed are computed, the frequency and scale distribution must be computed as well. As the immediate application of this implementation is to obtain the IF in respiratory signals, and because such signals are usually in the low frequency range (0.1 Hz to 0.3 Hz), a exponential frequency/scale distribution has been chosen. An advantage of using this type of frequency/scale distribution instead of a linear one, is that by narrowing the distance between frequencies/scales in the low frequency range allows for better resolution at these frequencies. One downside of this approach, if the signal contains frequencies in the high frequency range, the estimation of the IF in those frequencies will certainly contain errors.

To obtain the desired frequencies and scales, the following equations have been used.

$$
\begin{aligned}
f_{aux} &= (0, \cdots, i/N_{sc}, \cdots, 1) \\
freqs &= 1/2^{f_{aux}/\tau} \qquad (7.8) \\
scales &= \sigma/(2\pi freqs) \qquad (7.9)
\end{aligned}
$$

where $f_{aux} \in [0, 1)$ is a vector with $1/N_{sc}$ steps, N_{sc} is the total amount of scales that will be used. $freqs$ and $scales$ are the frequency and scale distributions respectively, where $freqs \in [0, 1)$ represents the normalized frequency vector, and $scales$ its associated scales. τ represents the exponential decay which has been set to 0.1 as default. Finally σ is a constant of value 6 as stated in equation 7.2. To simplify all the operations in the frequency domain, the normalized frequency between 0 and 1 has been used. For this reason the equation 7.8 which contain all the frequencies that will be analysed, is numerically constrained in the margin between [0,1).

Once the frequency/scale distribution is computed, by the means of equation 7.2 and 7.3, all the wavelets necessary to compute the CWT are generated with N_{fft} samples in the normalized frequency domain. All the wavelets are stored in a matrix of N_{sc} x N_{fft} dimensions allocated in GPU memory, being each row a different wavelet with N_{fft} samples.

SST computation

Once the code has been initialized, for each new sample the SST computation is triggered. The only input parameter for this function is the sliding window containing all the samples that will be analysed. A copy of all the samples is immediately performed into the GPU memory so that all the operations can be executed in the GPU domain without memory transactions being performed.

The first step to compute the SST is to perform an FFT of the sliding window. As the sliding window does not contain a power of 2 number of samples, it has to be symmetrically padded with zeros until this amount is reached. Once the signal has been zero-padded, the complex FFT can be performed.

In order to compute the CWT of the signal in the frequency domain, the complex multiplication between the Wavelet matrix and the signal has to be performed (equation 7.5). In pytorch, when performing an element wise multiplication between a vector and a matrix, if the vector is defined as a 1 row vector, each element inside the matrix column is multiplied by the corresponding element in the same column of the vector, and assigned to the same position inside the resultant matrix. As a result, the obtained matrix has the same dimensions than the original matrix. An example of the previous operation is shown in equation 7.10.

$$\mathbf{C\hat{W}T} = \hat{\mathbf{s}} \odot \hat{\boldsymbol{\psi}} = \begin{pmatrix} \hat{s}_1\hat{\psi}_{1,1} & \hat{s}_2\hat{\psi}_{1,2} & \cdots & \hat{s}_n\hat{\psi}_{1,n} \\ \hat{s}_1\hat{\psi}_{2,1} & \hat{s}_2\hat{\psi}_{2,2} & \cdots & \hat{s}_n\hat{\psi}_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{s}_1\hat{\psi}_{m,1} & \hat{s}_2\hat{\psi}_{m,2} & \cdots & \hat{s}_n\hat{\psi}_{m,n} \end{pmatrix} \tag{7.10}$$

where $\hat{\mathbf{s}}$ is the 1D vector of the sliding window in the frequency domain, $\hat{\boldsymbol{\psi}}$ is the matrix containing all the wavelets, m is the number of scales and n is the length of the FFT.

Once this operation is completed, and before the inverse fast fourier transform (IFFT) is performed on the $\mathbf{C\hat{W}T}$, as the next step described in equation 7.6 involves performing a derivative of $\mathbf{C\hat{W}T}$ and because the derivative in the Fourier domain si defined as follows: $\mathcal{F}\{f'(t)\} = jw\hat{f}(w)$, this derivative is computed in the frequency domain. It has to be noted that this operation does not involve matrix algebra, as the operation in equation 7.6 is defined for a 1D signal, thus this operation does not involve computing the first derivative of a matrix in the frequency domain. Once the $\mathbf{C\hat{W}T}$ and $\mathbf{dC\hat{W}T}$ have been obtained, the complex IFFT can be computed to proceed with the SST computation.

The next step involves obtaining the candidate frequencies for all the scales, to perform this operation the equation 7.6 has been applied. Note that in this case $(W_s(a,b))^{-1}$ does not involve the computation of a pseudo-inverse of the **CWT** matrix, but to perform an element wise division between **dCWT** and **CWT**. As explained before, if not specified, pytorch by default performs an element wise multiplication and division of matrices, which in this case can be expressed in the following equation.

$$\boldsymbol{\Omega} = -\jmath(\mathbf{dCWT} \oslash \mathbf{CWT}) \tag{7.11}$$

where $\boldsymbol{\Omega}$ is a matrix containing all the candidate IF.

Once the candidate IF has been computed, the last step to obtain the SST is to relocate all the frequencies as described in equation 7.7. As this step is the only one that cannot be done using matrices as it requires iterating for all the frequencies used, a for loop has been chosen instead. In the next pseudo-code the implementation of this last part is shown.

Algorithm 2 SST computation

```
for k in range(0,$N_{sc} - 1$)) do
    for i in range(0,$N_{fft} - 1$ do
        for j in range(0,$N_{sc} - 1$)) do
            if $\Omega_{i,j}$ - $freqs_k$ < $\Delta freqs_k$ then
                $aux_{i,j} \leftarrow \Omega_{i,j}$
            else
                $aux_{i,j} \leftarrow 0$
        $SST_{i,k} = \sum_k aux_{i,k}$
```

where **SST** is the synchrosqueezing transform and $\Delta freq_k$ is the distance between $freqs_k$ and $freqs_{k-1}$. Although the algorithm 2 shows three nested loops, the two interior loops are completely parallelized as the function **where** from pytorch has been used, but for illustrative purposes, all the involved search operations have been shown.

With this last step, the SST for the input signal has been computed. This procedure is performed each time a new sample enters the sliding window, and as it needs to perform in real-time, the computational time of all this procedure must not exceed $1/Fs$.

Ridge extraction

This last step is used to extract a 1D signal from the SST. This step is crucial in obtaining the IF, as it removes all the harmonics and only the evolution of the fundamental frequency is shown. To perform this operation, first the maximum values in **SST** and its positions are extracted. To extract the ridge from the **SST**, the previous indices are used to obtain the values in **SST**. If for a certain scale there are two maximums, the one with the lowest frequency is used.

7.2.3 Performance characterisation

Computing performance

To test the computing performance in both GPU and CPU of the proposed implementation, the computing time and absolute error for each combination of window length and scale has to be assessed. In order to obtain those parameters, a test based on changing the amount of scales and the window length has been designed. To perform this test, a file containing a 0.1 Hz sinusoidal signal sampled at 30 Hz with 200 s of length was generated, to ensure that for each one of the tested lengths the sliding window contains exactly the same signal being the only variable the number of samples used. The same test has been used to asses the performance either in the GPU and the CPU, with the particularity that in the GPU only one CPU thread was used to control the program execution, while in the CPU case up to 4 concurrent threads using OpenMP were used. To characterise the absolute error between the estimated IF and the fundamental frequency of the signal, the average of the IF has been used.

The tests were performed on an Intel i7-4710HQ CPU, with an Nvidia GeForce TX 850M graphics card and 8 GB of RAM. The CUDA toolkit version 10.2 was used along python (version 3.6.9).

Error Assessment

As the proposed implementation is intended to be used in respiratory signals in order to obtain its real-time IF evolution, the error between the proposed implementation and IF approximation and a reference method has to be assessed.

Dataset

The dataset used to evaluate the performance of the algorithm was the same used in chapter 6, were fifteen healthy subjects, eight male and seven female, were measured in a controlled environment. In Table 7.1 the anthropometric data including age, height, weight and body mass index (BMI) can be found.

Table 7.1: Anthropometric data expressed as mean ± SD.

Age [years]	Height [cm]	Weight [kg]	BMI [kg/m^2]
26.6 ± 3.54	170.47 ± 9.26	64.6 ± 11.97	22.09 ± 2.72

Table reproduced from [31] (CC BY 4.0).

The subjects were instructed to perform four tests as described in chapter 6, but only three were used in this study. The three test included two constant frequency breathing tests and the free breathing test. For the constant frequency test, the subjects were instructed to breath at a certain frequency (0.1 Hz and 0.3 Hz) with the help of a custom visual aid. The two constant frequency test had a length of 3 minute each while the free breathing test had a duration of 10 minute. From the dataset, only the signal from the reference method was used for this study, obtained by the means of commercial RIP from BioSignalsPlux [94] (chapter 6).

Signal Processing

Once the raw signal has been obtained and before the extraction of the IF, and in order to avoid errors due to the respiratory signal containing spurious high frequency components or amplitude changes that could have a negative impact on the IF, a filtering and normalization of the signal has been performed. The following signal processing steps were applied to the respiratory signal to achieve this normalization as partially described in chapter 6.

1. A 2nd order bidirectional zero-phase digital Butterworth bandpass filter was applied to the signal with cut-off frequencies of 0.05 Hz and 0.6 Hz.
2. A non-linear (equation 7.12) function based on the arctangent [155] was applied to the signal in order to eliminate the possible AM components, as well as to compress the signal between -1 and 1.

$$\widehat{R}[k] = \arctan \frac{R[k]}{\sqrt{\frac{\sum_{i=1}^{N}(R[i]-\bar{R})^2}{N-1}} * \sqrt{2}} \tag{7.12}$$

where $R[k]$ is a sample of the raw respiratory signal, $\bar{R}$ represents the mean of the raw respiratory signal, and finally $\widehat{R}[k]$ represents a normalized respiratory sample.

Instantaneous Frequency Estimation

Once the respiratory signal has been normalized, and in order to characterise the error between the real IF and the obtained IF from the proposed implementation, two reference methods have been used. In one hand the Hilbert IF has been computed, while on the other hand the IF signal obtained from the Matlab functions *wsst* and *wsstridge* was used.

To obtain the Hilbert IF, the procedure outlined in chapter 4 was used. This procedure is defined in [167] as the first derivative of the instantaneous phase of the respiratory signal. The necessary steps to compute the Hilbert IF are the following:

1. The Hilbert transform is obtained from the respiratory signal.
2. The phase of the Hilbert transform has to be unwrapped and corrected with 2π increments to ensure continuity.
3. The IF is obtained by the difference of adjacent samples in the previously corrected phase series.

4. To ensure that no abrupt phase changes are present in the signal, a Hodrick-Prescott [168] filter is applied with smoothing factor of $2 \cdot 10^6$.

On the other hand, the IF obtained with Matlab was computed using 48 octaves (480 scales) and with a penalty of 20 in the ridge computation. This specific configuration was used as it was the one that yielded the best accuracy in terms of IF. All the computations to obtain either the IF from the Hilbert transform or via the tools provided in the Matlab framework, were performed in the Matlab environment (version 2019b).

Regarding the proposed implementation, as the aim of the method is to work in real-time using a sliding window, the signals from the dataset have been parsed emulating a real-time system with a FIFO architecture, where each new sample is added to the sliding window while the oldest is removed. To obtain the IF, the central sample from the obtained IF has been added to an array to conform the IF signal. A 100 scales and a 60 s sliding window have been chosen to obtain the IF from the respiratory signal based on previous results. Regarding the sliding window length, and because the first available sample from the proposed implementation begins at the 30th second, the first and the last 30 seconds of the IF from the Hilbert method and the Matlab method have been cropped in order to align the IF signals from the proposed implementation with the IF signal from both reference methods. As the sample frequency of the normalised respiratory signal is 40 Hz, a total of 2400 samples are used in the sliding window at each iteration.

To assess the error of the proposed implementation, a plot comparing all the obtained mean IF and SD IF from the proposed implementaion, the Matlab Synchrosqueezing and the Hilbert transform for each one of the subjects has been performed. To quantify the error between the proposed method and the reference methods, the standard deviation of the error (SDE) has been computed. As the Hilbert IF has much more SD than the IF obtained with Matlab, only the latter one has been used to asses the SDE on a sample by sample basis.

Finally, an ADT [173] has been performed on the mean IF and SD IF for all the methods to verify if the samples presented a normal distribution. From all the tests, the Hilbert IF was the only method that showed a $p < 0.05$ for either the mean and SD IF, which discards the null hypothesis that the samples present a normal distribution.

7.3 Results

Two type of results will be presented on this section, the results regarding the computational performance of the algorithm, where the computational performance is evaluated, and the results regarding the error assessment of the proposed implementation, where the obtained IF is compared with other two IF reference methods.

Figure 7.1 shows an example of the input and output signal as well as the intermediate signals that can be obtained from the proposed implementation. Figure 7.1 corresponds to an input sliding window of 60 s and 100 scales for a 0.1 Hz signal sampled at 30 Hz, and from right to left, the signal being processed, the CWT of the signal, the SST of the signal and finally the IF for that window length are shown.

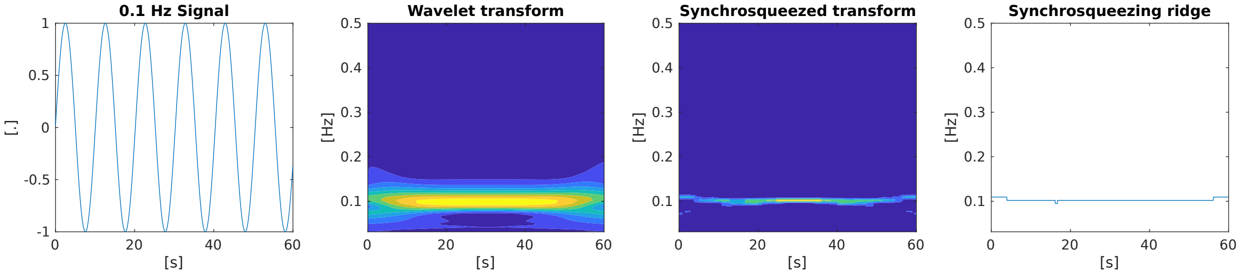
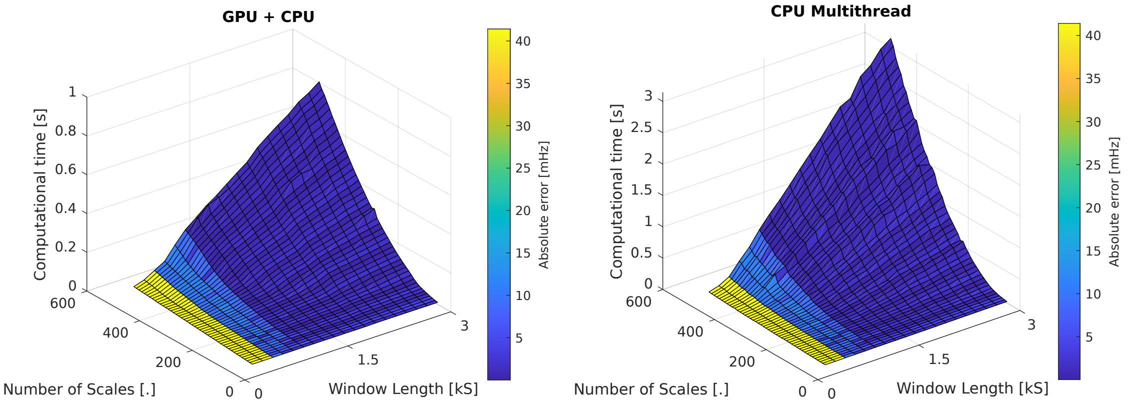


Figure 7.1: Input signal, CWT, SST and IF obtained with the proposed implementation at a 100 scales, for a 60 second window of 0.1 Hz signal sampled at 30 Hz.

As it can be seen in Figure 7.1, the CWT and SST steps depict the same 2D representation, but while in the CWT representation the spectral power density is concentrated around 0.1 Hz with a wide bandwidth, in the SST representation, this spectral density has been reduced ”squeezed” into thin line centred around 0.1 Hz. As previously stated, the last figure in Figure 7.1 is obtained by locating the maximum of the spectral power density of the SST step.

7.3.1 Computational Performance

As the aim of the proposed implementation is to compute the IF in real-time, the study of the performance when the algorithm is using a GPU or a multithread CPU is performed.



(a) (b)

Figure 7.2: a) Depicts the performance curve and absolute error for the GPU of the proposed implementation, while b) depicts the performance curve and absolute error for the CPU. The z-axis represents the execution time and the color-code the absolute error. The window length is reported in kilo-samples ”kS”

Figure 7.2 shows two figures which depict, respectively, the execution time (z-axis) and the absolute error (color-map) for the GPU (Figure 7.2a) and multithread CPU (Figure 7.2b). In Figure 7.2, for both GPU and CPU, a quadratic increase in computational time can be appreciated whenever the scales increase, while an almost linear increase is presented whenever the sliding window increases. Another characteristic of both figures, is that the error is almost constant when the sliding window is greater than 0.9 kS of length.

Table 7.2: GPU performance and absolute error.

	Scales		
W. Length [kS]	100	250	500
0.9	14.12 ms [2.02 mHz]	48.17 ms [8.69 mHz]	209.18 ms [2.84 mHz]
1.8	17.11 ms [2.02 mHz]	114.99 ms [0.77 mHz]	474.58 ms [0.77 mHz]
3	30.24 ms [2.02 mHz]	216.02 ms [0.77 mHz]	804.83 ms [0.77 mHz]

Where the data is presented as the execution time in ”ms” and [absolute error] in ”mHz” for each pair of sliding window length and scale. The window length is reported in kilo-samples ”kS”.

Table 7.2 shows a small subset of the computational time and absolute error for the GPU case. It can also be appreciated a quadratic increase of the computational time whenever the scale increases, as an example, for 100 scales the algorithm takes 14.12 ms while for 500 scales the algorithm takes 209.18 ms, both for a 0.9 kS window length. It can also be seen, that the error decreases whenever the scales increase, being the highest error for the 100 scale and 0.9 kS window, and the lowest for the 500 scale and 3 kS window.

Table 7.3: CPU performance and absolute error.

	Scales		
W. Length [kS]	100	250	500
0.9	17.61 ms [2.02 mHz]	68.61 ms [8.69 mHz]	585.5 ms [2.84 mHz]
1.8	29.48 ms [2.02 mHz]	309.4 ms [0.77 mHz]	1739 ms [0.77 mHz]
3	72.4 ms [2.02 mHz]	706.8 ms [0.77 mHz]	3142 ms [0.77 mHz]

Where the data is presented as the execution time in ”ms” and [absolute error] in ”mHz” for each pair of sliding window length and scale. The window length is reported in kilo-samples ”kS”.

Table 7.3 shows the same subset of data than Table 7.2 but for the CPU case. It can also be appreciated the same quadratic increase of the computational time whenever the number of used scales increases, as an example, for 100 scales the algorithm takes 17.61 ms while for 500 scales the algorithm takes 585.5 ms, both using the same 0.9 kS window. Regarding the absolute error, as the only change between the GPU and the CPU test does not imply a change in the proposed implementation as far as the device performing the operations, the absolute error for any given scale and window length pair is exactly the same in both Table 7.3 and Table 7.2.

7.3.2 Error Assessment

Signals

Figure 7.3 depicts an example of the obtained IF for the proposed implementation, the IF obtained with Matlab and the IF from the Hilbert transform, as well as the normalized respiratory signal from the Free test. As it can be seen, all three IFs follow the same temporal evolution being the proposed implementation and the Matlab IF on top of each other, indicating a high concordance between the proposed implementation and the reference methods. It can also be seen in Figure 7.3, that the IF from Matlab has more frequency resolution (more steps) than the proposed implementation, as it has been computed using 480 scales instead of the 100 of the proposed method.

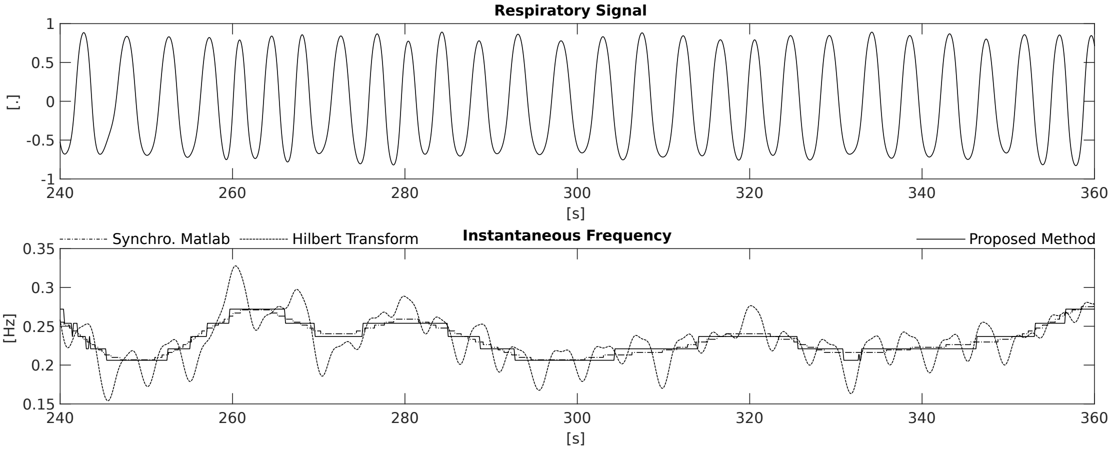


Figure 7.3: The first plot depicts the respiratory signal from the reference method, while the second plot depicts a comparison of the IF signals obtained from the proposed implementation, Matlab synchrosqueezing and Hilbert transform.

Performance

Figure 7.4 contains for each test and subject, the mean and SD IF for all the studied methods. As it can be seen for the 0.1 Hz and 0.3 Hz test, for all the subjects and methods, the mean IF is centred around the corresponding frequency. On the contrary, on the Free test each subject has a different mean IF. On this last test, it can also be seen that all the methods for the same subject report near identical results. Regarding the SD, both the proposed implementation and the Matlab results show very low SD while the Hilbert transform shows more dispersion. The SD results for the free test show more dispersion than the previous tests as each subject had its own variability, with the proposed implementation and Matlab results being closer to each other while the Hilbert transform presents more SD as in the previous cases.

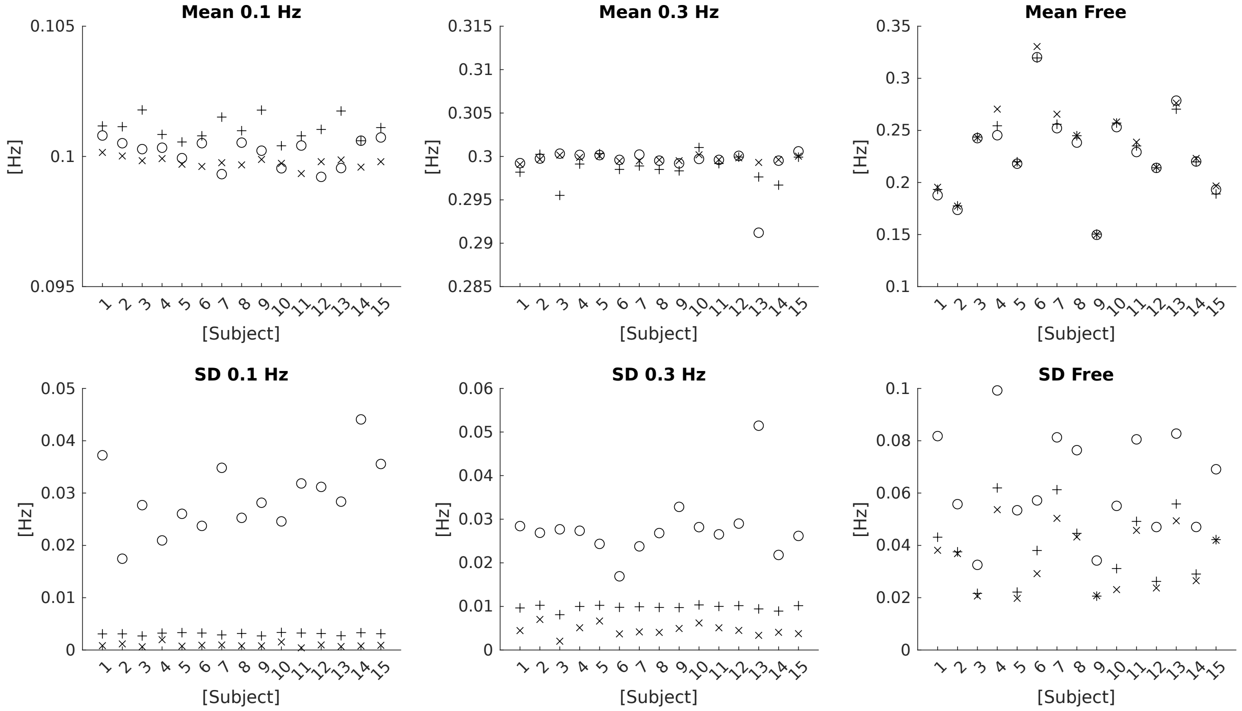


Figure 7.4: Comparison between Instantaneous frequency obtained using the Proposed implementation, the Hilbert transform and the Matlab Synchrosqueezing method. "+" represents the proposed method, "x" represents the symchrosqueezing transform obtained with Matlab and "o" represents the Hilbert transform

Table 7.4 contains the median and IQR [25 ; 75] of the mean IF, while Table 7.5 contains the median and IQR [25 ; 75] of the SD IF for all the subjects in Figure 7.4. The median and IQR have been chosen as the Hilbert IF samples do not present a normal distribution. Finally, Table 7.6 contains the SDE between the proposed implementation and the Matlab Synchrosqueezing on a sample by sample basis. The results are shown as mean ± sd as both methods presented a normal distribution.

Table 7.4: Median and IQR [25 ; 75] of the mean IF for all the subjects and methods.

Test	PI [mHz]	SM [mHz]	HT [mHz]
0.1 Hz	101.04 [100.78 ; 101.43]	99.8 [99.68 ; 99.88]	100.34 [99.65 ; 100.52]
0.3 Hz	298.9 [298.23 ; 299.92]	299.63 [299.49 ; 299.99]	299.71 [299.51 ; 300.17]
Free	235.07 [198.52 ; 255.52]	238.7 [201.15 ; 263.69]	229.23 [198.36 ; 250.44]

Where PI stands for Proposed Implementation, SM stands for Synchrosqueezing Matlab and HT stands for Hilbert Transform.

Table 7.5: Median and IQR [25 ; 75] of the SD IF for all the subjects and methods.

Test	PI [mHz]	SM [mHz]	HT [mHz]
0.1 Hz	3.15 [2.94 ; 3.24]	0.83 [0.76 ; 0.97]	28.15 [24.75 ; 34.08]
0.3 Hz	9.92 [9.67 ; 10.17]	4.46 [3.83 ; 5.11]	26.89 [24.8 ; 28.37]
Free	38.04 [26.94 ; 48.04]	36.85 [23.25 ; 45.1]	57.18 [48.63 ; 81.11]

Where PI stands for Proposed Implementation, SM stands for Synchrosqueezing Matlab and HT stands for Hilbert Transform.

Table 7.6: Mean ± SD of the SDE between the proposed implementation and the Matlab Synchrosqueezing for all the test.

	0.1 Hz [mHz]	0.3 Hz [mHz]	Free [mHz]
SDE	2.54 ± 0.35	7.19 ± 0.45	26.48 ± 15.66

7.4 Discussion

Regarding the results in Figure 7.2 and Tables 7.2 and 7.3, the first thing to be noted is the difference between the maximum computing time between GPU and CPU, being the maximum CPU computing time 3.9 times greater than the maximum GPU computing time. Moreover, when comparing the lowest computing time, only a difference of approximate 3.5 ms is found between GPU and CPU. The main reason why the CPU performance rapidly decreases when the scales and window length increases is due to the 4 cores rapidly maxed out at 100 % of its capacity and not being able to rapidly process all the data. On the other hand the GPU as it has more cores (hence more threads), and although is eventually maxed out, it can ingest and process more data per operation than the 4 cores of the CPU. Another reason for this difference in computation time is due to the GPU being optimized to perform matrix operations, and for this reason is more efficient in parallelizing all the operations involved in computing the SST.

Another important detail that can be appreciated in Figure 7.2 is the fact that the computing time increases quadratically whenever the scales increase, but it increases almost linearly when the window length increases. This can be due to two factors, the first one being the increase in number of operations that need to be performed whenever the scales increase, and the second reason can be due to when performing the "reallocation" of frequencies the proposed algorithm cannot parallelize the first loop described in Algorithm 2, which is translated into an increased computing time as more iterations of the loop are being performed. As for the linear increase whenever the window length is augmented, this is solely due to the increase in the number of operations and hence linear.

Given the results of the absolute error in Tables 7.2 and 7.3, it can be appreciated in Figure 7.2 that for a low number of scales the error can be up to 40 mHz. This is due to the ”tailing” error of the CWT as it can be seen in Figure 7.1, the beginning and ending samples as they pose a discontinuity an error on the CWT computation is produced, and if the window length is below a certain number of samples, the estimation of the mean IF for that window will certainly contain errors. As for the other scales and window lengths, whenever the window length is higher than 1.6 kS, the error due to this discontinuities does not contribute as much to the error as the difference between the discrete scale/frequency to the actual frequency. For this reason, whenever the window length is greater than 1.6 kS, and the number of scales is greater than 100, the absolute error becomes smaller than a 5 %.

For the results in the comparison between methods in Table 7.4, small differences can be seen between the proposed implementation and the reference methods, for the constant breathing frequency tests. Regarding the Free breathing test, all the methods show similar results as well, being the Hilbert IF the one with the lowest median. For the results in Table 7.5 in both the constant frequency tests, the proposed implementation has a median SD IF below 5 % of the median mean IF showing good results in the IF estimation. If compared with the other two reference methods, the proposed implementation has more SD than the Matlab Synchrosqueezing but way below the IF from the Hilbert transform. This last results indicate that the proposed implementation, and taking into account the different in scales between the Matlab results and the proposed implementation, can be used to obtain the IF of a constant frequency signal. Looking at the results for the Free breathing test, the results from the proposed implementation and the Matlab Syncrhosqueezing are very similar with only a few milihertz of difference either in median and interquartile range. As for the comparison with the results obtained from the Hilbert IF, the proposed implementation has lower median and lower interquartile range. Overall, the proposed implementation yields very good results when compared with the Matlab implementation, taking into account the difference in the number of used scales.

Finally, the results in Table 7.6, show a low SDE for both the constant frequency tests, which indicates a good agreement between the IF obtained from the proposed implementation and the Matlab Synchrosqueezing, being all the results below a 5% error for each test respectively. Regarding the Free test an increased SDE can be found. As each subject was breathing at a different frequency with different frequency modulations within the respiratory signal, and as the signals from the proposed implementation and the Matlab Synchrosqueezing have been obtained using a different number of scales, whenever a small frequency change occurs, the Matlab IF is able to shift to the next scale while the proposed implementation due to the limitation in the number of used scales, cannot jump two the next discrete frequency. Because of that difference, the SDE shows higher results for this test than the constant frequency breathing tests. The effect described before can be seen in Figure 7.3 within the seconds 260 and 280.

7.5 Conclusions

In this chapter a new implementation of the synchrosqueezing transform (SST) has been presented, as well as an approximation of the IF estimation of the respiratory signal. The proposed implementation is based on performing all the operations involved in the SST as matrix operations so its easy to parallelize and to obtain performances that otherwise could not be achieved. The proposed implementation has been build based on the pytorch library, which enables GPU parallelization as well as CPU parallelization, thus achieving real-time performance when computing the IF. The results regarding the computational performance show an that the proposed implementation can perform with low error at real-time speeds. The results also show the relationship between the absolute error and the number of scales, whenever the scale number increases the error decreases and the performance decreases as well. Finding a good trade-off between the number of scales, the maximum error and the computational performance is crucial for real-life applications.

Regarding the error assessment of the approximated version of the IF, the results show a high agreement between the Matlab IF and the IF for the proposed implementation on the constant breathing tests, with an increased SDE for the Free breathing test due to the difference in the number of scales used in both methods. The proposed implementation also showed much better results than the Hilbert IF in terms of SD.

In general, the proposed implementation can be used in real-time implementations to obtain the variability of the instantaneous frequency.

Chapter 8

Conclusions & Future work

In this chapter the general conclusions and main findings of this thesis are reported, as well as a last section detailing the proposed future work.

8.1 Conclusions

Non-contact unobtrusive methods for respiratory measurement have been recently established as an alternative to traditional methods. Moreover, those based on video analysis have experimented an increase in popularity in the recent years. In this thesis, different methods based on video processing have been presented and analysed, with the aim to improve the performance of the current methods as well as to validate the feasibility of these methods to measure respiratory signals, from which the respiratory rhythm can be obtained.

This section has been organized focusing on the different objectives of this thesis summarized in three different subsections, where all the different achievements and main findings are reported.

8.1.1 Video-based methods

In this thesis, Three different video-based methods have been studied, being two of them based on commercial consumer-grade RGB camera and one based on a stereo depth camera.

The first method described in chapter 3, is based on the detection inside the frame of a custom pattern placed on the thorax of the subject. The respiratory signal of the subject is obtained by tracking the evolution of a pattern inside the image along the successive frames, by the means of sparse optical flow techniques. The proposed method is comprised by two differentiated stages, the detection stage and the tracking stage. The detection stage is only performed whenever the pattern needs to be located inside the frame, while the tracking stage is performed every frame as long as there are enough features of the detected pattern inside the frame. Thanks to the asymmetry between the detection and tracking stages, the proposed method can achieve performances up to 80 Hz in the tracking stage, making the method feasible to extract the respiratory signal in real-time. The method has been validated with 21 healthy subjects and compared with a commercial inductive plethysmography system. The tests have been performed to validate the method against the reference system, with results for the intra-class Fisher correlation ranging from 0.945 for the best result (0.1 Hz test) to 0.85 for the worst (Reading test). Regarding the objectives proposed in this thesis, this method fulfils the first sub-objective of proposing and validating a method that uses a consumer-grade camera to obtain the respiratory signal of a subject.

The second method, presented in chapter 5, involves the use of an stereo depth camera (Intel Realsense$^{\text{TM}}$ ZR300) to obtain the respiratory signal through the evolution of the depth map. The novelty of this method is that by the means of the IR image obtained trough the same camera, the face of the subject is located and a region of interest (ROI) on the thoracic cage is obtained. Whenever the subject moves or changes its position, a new ROI is computed. This ROI is later used in conjunction with the depth map to obtain the changes in the distance between the subject and the camera, that are proportional to the respiration hence obtaining the respiratory signal of the subject. The method was designed to perform in real-time with simultaneous acquisition of the IR feed and the depth map from the camera. The method was evaluated with 20 healthy subjects using an inductive plethysmography system as a reference method. The measurements where taken in a car simulator were the subject could act without any restriction of movement within the capabilities of the simulator. Each subject was asked to drive normally within the car simulator parameters. The proposed method was designed taking into account the limitations of the first method regarding the illumination changes and the use of a pattern to extract the respiratory signal. By computing a ROI from the location of the face of the subject, and by using the depth map to extract the respiratory signal, both limitations are assessed and the third sub-objective is fulfilled.

The third method defined in chapter 6, uses a consumer-grade camera located at the side of the subject, to measure the displacement of the thoracic-abdominal region. The proposed method makes use of a dense optical-flow algorithm to measure the movement and track the evolution of the thorax. The main characteristic of the method, is that the respiratory signal is embedded into the phase of the optical flow, while the modulus can be used as a quality indicator. This method does not require any ROI or any previous calibration, and due to the simplicity of the algorithms used to extract the respiratory signal, the proposed method is suitable to be used in real-time.

To validate the method, 15 subjects were measured using and an inductive plethysmography system as a reference. Four tests where performed for each subject, two of them were designed to test the method in constant frequency breathing, one was designed to test the method in free breathing conditions, while the last test was designed to validate the quality indicator. This method fulfils the third and fourth sub-objective by improving the first two methods for respiratory signal extraction, while achieving a null user interaction as no ROI or previous calibration is needed.

8.1.2 Respiratory Rhythm Measurement

In this subsection, the performance in terms of respiratory rhythm measurement, for all the presented methods, is reported and the principal conclusions for each method are exposed.

For the *Pattern* method (chapter 3) in terms of respiratory rhythm analysis, the MAE, MAPE, standard deviation of the error (SDE) indicators and a Bland-Altman (BA) analysis of the cycles obtained from the respiratory signal have been performed. Moreover, an analysis of the accuracy of the respiratory cycle extraction has been obtained when compared with the reference method. Relative to the error results for all tests, the method yields low SDE and MAE results with MAPE results below 10 %, thus showing a good agreement between the proposed method and the reference method. Regarding the accuracy in the cycle detection, the lowest reported SEN value is 94.02 % with a corresponding PPV of 89.59 % for the reading test, implying a very accurate cycle detection for the proposed method. Given the aforementioned results, the proposed method yields very good results in terms of respiratory rhythm measurement for all tests.

In chapter 4 a comparison between three different methods, the one presented in chapter 3 (pattern), a method based on a Microsoft Kinect$^{\text{TM}}$ camera (depth) and a method based on a thermal camera (thermal), has been performed to assess the feasibility of each method on respiratory rhythm measurement. An inductive plethysmograph was used as a reference method for all the tests. To compare the obtained signals, the Hilbert IF was obtained for each method. The results show a high agreement between the pattern and depth method with the reference method in terms of SDE for all tests. Regarding the BA results for the two aforementioned methods, the constant breathing tests show narrow limits of agreement and small bias values in the order of mHz. For the free breathing and reading tests, both methods present good results being the depth method slightly better in terms of error than the pattern method. Regarding the thermal method, an increased SDE results can be seen, moreover, the BA analysis shows a gain bias that could be caused by slight differences between the real sampling frequency of the camera and the one used to process all the data.

This chapter fulfils the second sub-objective of performing an evaluation of different video-based methods in terms of respiratory rhythm analysis. The depth and pattern methods show results that indicate the feasibility of using these two methods for respiratory rhythm measurement in constant and free breathing conditions. Regarding the thermal method, it cannot be directly discarded due to the limitations of the study.

For the *Depth* method (chapter 5), an opportunistic approach in the respiratory signal analysis has been proposed, where each piece has been analysed independently, even if they belong to the same subject. Each piece had a minimum length of 60 seconds and it was obtained via visual inspection of the respiratory signal. An error and accuracy analysis of the respiratory cycles obtained from each piece has been performed, by using the MAE, MAPE and SDE indicators for the error assessment and the SEN and PPV indicators for the accuracy of the cycle detection. Moreover, an analysis of the Hilbert instantaneous frequency for each piece has also been performed, as well as a BA analysis of all the obtained cycles and the instantaneous frequency. Furthermore, a statistical study has been performed between the dependence of the length of each piece and the standard deviation of the error, for both the cycles and the IF. Regarding the cycle results, a global sensitivity of 77.21 % can be appreciated with a global PPV of 80.69 %. The MAE, MAPE and SDE results show low values that indicate a good agreement between the proposed method and the reference system. The BA results for the detected cycles show no bias error with narrow limits of agreement. As for the IF results, the SDE reports low values as well indicating a good agreement between methods. The BA plots for the IF show no observable bias, but as the samples did not present a normal distribution, this can not be assured. Given the aforementioned results, the proposed method has proven its reliability in measuring the respiration of the subject, as well as measuring the respiratory rhythm.

The last part of the study of the *Depth* method (chapter 5), was focused on the interaction between the length of each piece and the standard deviation of the error, either for the cycle detection and the IF estimation. The results showed non-significant differences between the length of each piece and the error between methods. These results confirm the hypothesis that using variable length pieces obtained in an opportunistic manner, does not have an impact on the error of the obtained signal. Moreover, the presented results also confirm the feasibility of using variable-length pieces in real-life systems based on the quality of the signal, thus fulfilling the fourth objective of this thesis.

Regarding the *Lateral* method (chapter 6), four tests have been performed, being three of them with the sole purpose of assessing the viability of the proposed method to measure respiratory rhythm, where the first two tests involved breathing at a constant frequency, while in the third test the subject was asked to breath freely. For the constant breathing frequency tests, the Synchrosqueezing instantaneous frequency was obtained, being the error assessed via a BA analysis and SDE results, for each breathing frequency. The BA results show narrow limits of agreement and no bias can be found, moreover, low SDE results are also reported thus indicating a good agreement between methods for both tests. For the free breathing test, the respiratory cycles and the synchrosqueezing IF have been extracted and analysed. For the respiratory cycle series, a sample by sample correlation has been obtained, the MAE, MAPE and SDE error indicators have been computed, as well as the SEN and PPV accuracy indicators and a BA analysis of all the obtained cycles. For the IF analysis, the correlation between IF signals and SDE has been obtained, and a BA analysis has been performed. The results for the cycle extraction yield a correlation of 0.94 for all the subjects, with low SDE, MAE and MAPE, indicating a high agreement between the proposed method and the reference method. Regarding the accuracy of the cycle detection, a median SEN of 100 % and a mean PPV of 94.7 % are obtained, thus validating the error results. For the BA analysis, the results show narrow limits of agreement with no bias present. The instantaneous frequency results show a correlation between IF signals of 0.9, with low SDE values indicating a high agreement in terms of IF between the proposed method and the reference method. The BA analysis for the IF show a narrow 95 % reference interval with no apparent bias bias or trend.

Given the results of the constant frequency breathing tests and the free breathing test, the proposed method can be used to extract the respiratory signal of a subject from a lateral perspective, while it can also be used to assess the respiratory rhythm. Also, thanks to the simplicity of the method, and that no interaction with the subject is required whatsoever, this fulfils the first and the third sub-objectives and partially fulfils the fifth sub-objective of this thesis.

Finally, the last test performed for the *Lateral* method in chapter 6 Section 6.3.3, was designed to validate the hypothesis that a quality index could be obtained from the modulus of the optical flow. To assess the previous hypothesis, the relationship between the length of the sliding window, used to obtain the quality index, and each activity performed in the test, was assessed using the Kruskal-Wallis statistical test as well as an Area Under the Curve analysis. The results show a clear dependence between the length of the window and the estimation of the quality index, while showing significant differences between the free breathing regions (good signal) and the reading/activity regions (bad signal). The area under the curve also shows a 0.9533 for a 10 s window thus reporting a good trade-off between length, specificity and statistical performance. For all the results reported, it can be asserted that the modulus of the optical-flow, for this method, can be used as a quality index, thus verifying the initial hypothesis and fulfilling the second part of the fifth sub-objective.

Cross-comparison of the proposed methods

As all the methods presented in this thesis have at least one test in free breathing conditions, where the respiratory cycles are obtained and analysed, hence a cross-comparison of the three methods can be performed. In Table 8.1 the results of the MAE, MAPE and SDE of the respective free breathing test for each method are reported.

Table 8.1: Comparison of the error results for the three presented methods, expressed as mean ± SD of the respective free breathing test.

Method	MAE [s]	MAPE [%]	SDE [s]
Pattern	0.27 ± 0.28	4.64 ± 2.69	0.36 ± 0.37
Depth	0.46 ± 0.25	8.49 ± 2.99	0.62 ± 0.32
Lateral	0.22 ± 0.1	4.87 ± 1.57	0.34 ± 0.16

For all the error results of all the methods, it can be seen that the *Lateral* method is the one with the lowest error values, while the *Depth* method presents the highest values. Regarding the *Depth* method, the test was performed without any constraint imposed to the subject, hence the reported error results are higher than the other two methods where the tests where performed in a more controlled environment. Regarding the *Lateral* method, the difference in the duration of the test has to be taken into consideration when comparing the error results with the other two methods.

In conclusion, all the proposed methods present a good agreement with the reference method, being the *Lateral* method the one with lower error results and higher agreement, thus fulfilling the main objective of obtaining quality respiratory signals from the proposed methods. Furthermore, all the presented methods have been designed to perform in real-time, with real-life conditions, being the last two methods specifically designed to perform the measurements without any subject interaction. This fulfils the secondary objective of this thesis.

8.1.3 Respiratory Signal Processing

In chapter 7 a new implementation of the Synchrosqueezing transform (SST) algorithm is presented. The proposed implementation has as an objective to improve the computational performance of the state of the art methodologies, with the aim to use the SST to obtain the instantaneous frequency of the respiratory signals in real-time. The proposed implementation shows a high agreement between the instantaneous frequency obtained through the proposed method and the reference method, and it is able to achieve real-time performances when used with GPU parallelization. This fulfils the last sub-objective of this thesis, and enables the computation in real-time of the instantaneous frequency of the respiratory signal.

8.2 Future Work

In reference to future lines of work, overcoming the limitations of the proposed methods in terms of lighting changing conditions and external vibrations, would be the first point to investigate. Improving the robustness of these methods against the aforementioned parameters would be a major step in the investigation for respiratory signal extraction by the means of methods based on video analysis.

As future lines of research to improve the current proposed methods in terms of respiratory rhythm analysis would be:

- A future study involving the quality index and the opportunistic piece selection could be performed, to assess the viability of using the proposed quality index to effectively obtain a piece of the respiratory signal.
- An study involving the variability of the instantaneous frequency obtained in real-time with the proposed synchrosqueezing implementation, could also be performed to assess if the variability is correlated with an increase in stress or drowsiness.
- A study involving the opportunistic approach and the real-time instantaneous frequency could be performed as well, to verify that by selecting a piece with sufficient length, the variability of the instantaneous frequency can still be obtained and it is statistically relevant.

Bibliography

[1] Majd AlGhatrif and Joseph Lindsay. A brief review: history to understand fundamentals of electrocardiography. *Journal of Community Hospital Internal Medicine Perspectives*, 2(1): 14383, 2012. ISSN 2000-9666. doi: 10.3402/jchimp.v2i1.14383.

[2] W. Einthoven, G. Fahr, and A. de Waart. On the direction and manifest size of the variations of potential in the human heart and on the influence of the position of the heart on the form of the electrocardiogram. *American Heart Journal*, 1950. ISSN 10976744. doi: 10.1016/0002-8703(50)90165-7.

[3] S. Serge Barold. Willem Einthoven and the birth of clinical electrocardiography a hundred years ago. *Cardiac Electrophysiology Review*, 2003. ISSN 13852264. doi: 10.1023/A:1023667812925.

[4] Julian F. Thayer, Shelby S. Yamamoto, and Jos F. Brosschot. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. *International Journal of Cardiology*, 141(2):122–131, may 2010. ISSN 01675273. doi: 10.1016/j.ijcard.2009.09.543. URL `http://dx.doi.org/10.1016/j.ijcard.2009.09.543`.

[5] Julian F. Thayer, Fredrik Åhs, Mats Fredrikson, John J. Sollers, and Tor D. Wager. A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. *Neuroscience and Biobehavioral Reviews*, 36(2): 747–756, 2012. ISSN 01497634. doi: 10.1016/j.neubiorev.2011.11.009. URL `http://dx.doi.org/10.1016/j.neubiorev.2011.11.009`.

[6] Yuan Ting Zhang and Carmen C.Y. Poon. Health informatics: Unobtrusive physiological measurement technologies. *IEEE Journal of Biomedical and Health Informatics*, 17(5):893, 2013. ISSN 21682194. doi: 10.1109/JBHI.2013.2279187.

[7] Duarte Dias and João Paulo Silva Cunha. Wearable health devices—vital sign monitoring, systems and technologies. *Sensors (Switzerland)*, 18(8), 2018. ISSN 14248220. doi: 10.3390/s18082414.

[8] Ju Wang, Joana M. Warnecke, Mostafa Haghi, and Thomas M. Deserno. Unobtrusive health monitoring in private spaces: The smart vehicle. *Sensors (Switzerland)*, 20(9), 2020. ISSN 14248220. doi: 10.3390/s20092442.

[9] European Comission. Green Paper on mobile Health (mHealth). Technical report, Brussels, 2014. URL `https://ec.europa.eu/digital-single-market/en/news/green-paper-mobile-health-mhealth`.

[10] World Health Organization. FROM INNOVATION TO IMPLEMENTATION eHealth in the WHO European Region, 2016. ISSN 978 92 890 5137 8. URL `http://www.euro.who.int/__data/assets/pdf_file/0012/302331/From-Innovation-to-Implementation-eHealth-Report-EU.pdf`.

[11] Yuan Ting Zhang, Ya Li Zheng, Wan Hua Lin, He Ye Zhang, and Xiao Lin Zhou. Challenges and opportunities in cardiovascular health informatics. *IEEE Transactions on Biomedical Engineering*, 60(3):633–642, 2013. ISSN 00189294. doi: 10.1109/TBME.2013.2244892.

[12] John Enderle, Susan M. Blanchard, and Joseph Bronzino. Introduction to Biomedical Engineering. In *Introduction to Biomedical Engineering*, pages 54–58. Academic Press, 1999. ISBN 9780122386602.

[13] Christoph Brüser, Christoph Hoog Antink, Tobias Wartzek, Marian Walter, and Steffen Leonhardt. Ambient and unobtrusive cardiorespiratory monitoring techniques. *IEEE Reviews in Biomedical Engineering*, 8:30–43, 2015. ISSN 19411189. doi: 10.1109/RBME.2015.2414661.

[14] Steffen Leonhardt, Lennart Leicht, and Daniel Teichmann. Unobtrusive vital sign monitoring in automotive environments—A review. *Sensors (Switzerland)*, 18(9):1–38, 2018. ISSN 14248220. doi: 10.3390/s18093080.

[15] Alexandros Pantelopoulos and Nikolaos G. Bourbakis. A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, 40(1):1–12, 2010. ISSN 10946977. doi: 10.1109/TSMCC.2009.2032660.

[16] Chenglu Sun, Wei Li, Chen Chen, Zeyu Wang, and Wei Chen. An unobtrusive and non-contact method for respiratory measurement with respiratory region detecting algorithm based on depth images. *IEEE Access*, 7:8300–8315, 2019. ISSN 21693536. doi: 10.1109/ACCESS.2018.2890082.

[17] Xiaobai Li, Jie Chen, Guoying Zhao, and Matti Pietikäinen. Remote heart rate measurement from face videos under realistic situations. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 4264–4271, 2014. ISSN 10636919. doi: 10.1109/CVPR.2014.543.

[18] Jelena Nikolic-Popovic and Rafik Goubran. Impact of motion artifacts on video-based non-intrusive heart rate measurement. *2016 IEEE International Symposium on Medical Measurements and Applications, MeMeA 2016 - Proceedings*, 2016. doi: 10.1109/MeMeA.2016.7533740.

[19] Changzhi Li and Jenshan Lin. Complex signal demodulation and random body movement cancellation techniques for non-contact vital sign detection. *IEEE MTT-S International Microwave Symposium Digest*, pages 567–570, 2008. ISSN 0149645X. doi: 10.1109/MWSYM.2008.4633229.

[20] Jianxuan Tu, Taesong Hwang, and Jenshan Lin. Respiration Rate Measurement under 1-D Body Motion Using Single Continuous-Wave Doppler Radar Vital Sign Detection System. *IEEE Transactions on Microwave Theory and Techniques*, 64(6):1937–1946, 2016. ISSN 00189480. doi: 10.1109/TMTT.2016.2560159.

[21] J G Webster. *Medical Instrumentation Application and Design, 4th Edition.* Wiley Global Education, 2009. ISBN 9781118312858.

[22] Michelle A. Cretikos, Rinaldo Bellomo, Ken Hillman, Jack Chen, Simon Finfer, and Arthas Flabouris. Respiratory rate: The neglected vital sign. *Medical Journal of Australia*, 188(11): 657–659, 2008. ISSN 0025729X. doi: 10.5694/j.1326-5377.2008.tb02165.x.

[23] Michael J. Tipton, Abbi Harper, Julian F.R. Paton, and Joseph T. Costello. The human ventilatory response to stress: rate or depth? *Journal of Physiology*, 595(17):5729–5752, 2017. ISSN 14697793. doi: 10.1113/JP274596.

[24] Andrea Nicolò, Carlo Massaroni, and Louis Passfield. Respiratory frequency during exercise: The neglected physiological measure. *Frontiers in Physiology*, 8(DEC):1–8, 2017. ISSN 1664042X. doi: 10.3389/fphys.2017.00922.

[25] Noelia Rodriguez Ibanez, MIREYA Fernandez Chimeno, Juan Jose Ramos Castro, Miguel Angel Garcia Gonzalez, Eduard Montseny Masip, and Daniel Bande Matinez. Method and System for Determining an Individual's State of Attention, 2014. URL `https://worldwide.espacenet.com/publicationDetails/biblio?FT=D&date=20110203&DB=&locale=en_EP&CC=US&NR=2011028857A1&KC=A1&ND=5`. US2011028857 (A1).

[26] Federico Guede-Fernandez, Mireya Fernandez-Chimeno, Juan Ramos-Castro, and Miguel A. Garcia-Gonzalez. Driver Drowsiness Detection Based on Respiratory Signal Analysis. *IEEE Access*, pages 1–1, 2019. ISSN 2169-3536. doi: 10.1109/ACCESS.2019.2924481. URL `https://ieeexplore.ieee.org/document/8744224/`.

[27] F Guede. *Una contribución a la evaluación de la adherencia a hábitos de vida saludables basado en aplicaciones móviles.* PhD thesis, Universitat Politècnica de Catalunya, 2018. URL `http://hdl.handle.net/2117/121196`.

[28] V Ferrer. *Estudio de la viabilidad de la estimación del grado de adherencia a estilos de vida saludables a partir de medidas cardíacas oportunistas.* PhD thesis, Universitat Politècnica de Catalunya, 2019. URL `http://hdl.handle.net/2117/170437`.

[29] 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. *Biomedical Signal Processing and Control*, 51:138–147, 2019. ISSN 17468094. doi: 10.1016/j.bspc.2019.02.004. URL `https://linkinghub.elsevier.com/retrieve/pii/S1746809419300400`.

[30] 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*, pages 1–1, 2019. ISSN 2169-3536. doi: 10.1109/ACCESS.2019.2947759. URL `https://ieeexplore.ieee.org/document/8871160/`.

[31] M. Mateu-Mateus, F. Guede-Fernández, M. A. García-González, J. J. Ramos-Castro, and M. Fernández-Chimeno. Camera-based method for respiratory rhythm extraction from a lateral perspective. *IEEE Access*, 8:154924–154939, 2020. ISSN 2169-3536. doi: 10.1109/ACCESS.2020.3018616.

[32] K. Konno and J. Mead. Measurement of the separate volume changes of rib cage and abdomen during breathing. *Journal of applied physiology*, 22(3):407–422, 1967. ISSN 00218987. doi: 10.1152/jappl.1967.22.3.407.

[33] Masafumi Nozoe, Kyoshi Mase, Sachie Takashima, Kazuhiro Matsushita, Yusuke Kouyama, Hiromi Hashizume, Yurina Kawasaki, Yuki Uchiyama, Noriyasu Yamamoto, Yoshihiro Fukuda, and Kazuhisa Domen. Measurements of chest wall volume variation during tidal breathing in the supine and lateral positions in healthy subjects. *Respiratory Physiology and Neurobiology*, 193(1):38–42, 2014. ISSN 15699048. doi: 10.1016/j.resp.2013.12.016. URL `http://dx.doi.org/10.1016/j.resp.2013.12.016`.

[34] A. Aliverti, S. J. Cala, R. Duranti, G. Ferrigno, C. M. Kenyon, A. Pedotti, G. Scano, P. Sliwinski, Peter T. Macklem, and S. Yan. Human respiratory muscle actions and control during exercise. *Journal of Applied Physiology*, 83(4):1256–1269, 1997. ISSN 87507587. doi: 10.1152/jappl.1997.83.4.1256.

[35] B. H. Brown, R. H. Smallwood, D. C. Barber, P. V. Lawford, and D. R. Hose. *Medical Physics and Biomedical Engineering.* IOP Publishing Ltd, 1999. ISBN 0-7503-0368-9.

[36] John E. Hall and Arthur C. Guyton. *Textbook of Medical Physiology.* Saunders 2011, 12th edition, 2010. ISBN 9780808924005. URL `https://www.elsevier.com/books/guyton-and-hall-textbook-of-medical-physiology/hall/978-0-8089-2400-5`.

[37] Christian F. Clarenbach, Oliver Senn, Thomas Brack, Malcolm Kohler, and Konrad E. Bloch. Monitoring of ventilation during exercise by a portable respiratory inductive plethysmograph. *Chest*, 128(3):1282–1290, 2005. ISSN 00123692. doi: 10.1378/chest.128.3.1282. URL `http://dx.doi.org/10.1378/chest.128.3.1282`.

[38] UK) Hilton, Penny (University of Sheffield. *Fundamental nursing skills.* John Wiley & Sons, Inc., 2006. ISBN 978-0-470-03239-8.

[39] Gila Benchetrit. Breathing pattern in humans: diversity and individuality. *Respiration Physiology*, 122(2-3):123–129, sep 2000. ISSN 00345687. doi: 10.1016/S0034-5687(00)00154-7. URL `http://linkinghub.elsevier.com/retrieve/pii/S0034568700001547`.

[40] Jacek Jaworski and Jason H.T. Bates. Sources of breathing pattern variability in the respiratory feedback control loop. *Journal of Theoretical Biology*, 469:148–162, 2019. ISSN 10958541. doi: 10.1016/j.jtbi.2019.03.001. URL `https://doi.org/10.1016/j.jtbi.2019.03.001`.

[41] C Gordon Douglas and J S Haldane. The regulation of normal breathing. *The Journal of Physiology*, 38(5):420–440, jun 1909. ISSN 14697793. doi: 10.1113/jphysiol.1909.sp001315. URL `http://doi.wiley.com/10.1113/jphysiol.1909.sp001315`.

[42] Jacopo P Mortola. How to breathe? Respiratory mechanics and breathing pattern, 2019. ISSN 18781519. URL `https://doi.org/10.1016/j.resp.2018.12.005`.

[43] Yushiro Kuratomi, Nobuo Okazaki, Teruo Ishihara, Tatsuo Arai, and Shiro Kira. Variability of Breath-By-Breath Tidal Volume and Its Characteristics in Normal and Diseased Subjects Ventilatory Monitoring with Electrical Impedance Pneumography. *Japanese Journal of Medicine*, 24(2):141–149, 1985. ISSN 00215120. doi: 10.2169/internalmedicine1962.24.141.

[44] Mudiaga Sowho, Jason Amatoury, Jason P Kirkness, and Susheel P. Patil. Sleep and Respiratory Physiology in Adults. *Clinics in Chest Medicine*, 35(3):469–481, sep 2014. ISSN 02725231. doi: 10.1016/j.ccm.2014.06.002. URL `https://linkinghub.elsevier.com/retrieve/pii/S0272523114000409`.

[45] Guillermo Gutierrez, Jeffrey Williams, Ghadah A. Alrehaili, Anna McLean, Ramin Pirouz, Richard Amdur, Vivek Jain, Jalil Ahari, Amandeep Bawa, and Shawn Kimbro. Respiratory rate variability in sleeping adults without obstructive sleep apnea. *Physiological Reports*, 4 (17):1–9, 2016. ISSN 2051817X. doi: 10.14814/phy2.12949.

[46] Young Jae Park and Young Bae Park. Clinical utility of paced breathing as a concentration meditation practice. *Complementary Therapies in Medicine*, 20(6):393–399, 2012. ISSN 09652299. doi: 10.1016/j.ctim.2012.07.008. URL `http://dx.doi.org/10.1016/j.ctim.2012.07.008`.

[47] Alfred Fleisch. Der Pneumotachograph; ein Apparat zur Geschwindigkeitsregistrierung der Atemluft. *Pflügers Archiv für die Gesamte Physiologie des Menschen und der Tiere*, 209(1): 713–722, 1925. ISSN 00316768. doi: 10.1007/BF01730956.

[48] JOHN C Lilly. Flow meter for recording respiratory flow of human subjects. *Methods in medical research*, 2:113–121, 1950.

[49] Ismael Perez-Suarez, Marcos Martin-Rincon, Juan José Gonzalez-Henriquez, Chiara Fezzardi, Sergio Perez-Regalado, Victor Galvan-Alvarez, Julian W. Juan-Habib, David Morales-Alamo, and Jose A. L. Calbet. Accuracy and Precision of the COSMED K5 Portable Analyser. *Frontiers in Physiology*, 9(December):1–12, 2018. ISSN 1664-042X. doi: 10.3389/fphys.2018.01764.

[50] Carlo Massaroni, Andrea Nicolò, Daniela Lo Presti, Massimo Sacchetti, Sergio Silvestri, and Emiliano Schena. Contact-based methods for measuring respiratory rate. *Sensors (Switzerland)*, 19(4):1–47, 2019. ISSN 14248220. doi: 10.3390/s19040908.

[51] Mustafa I. Beyaz, S. Habibiabad, H. Yildiz, U. Goreke, and K. Azgin. A Turbine-Based MEMS Sensor for Spirometry With Wearable Devices. *IEEE Sensors Journal*, 19(19):8612–8617, 2019. ISSN 1530-437X. doi: 10.1109/jsen.2019.2919599.

[52] P. Plakk, P. Liik, and P. H. Kingisepp. Hot-wire anemometer for spirography. *Medical & Biological Engineering & Computing*, 36(1):17–21, jan 1998. ISSN 0140-0118. doi: 10.1007/BF02522852. URL `http://link.springer.com/10.1007/BF02522852`.

[53] I Yoshiya, Y Shimada, and K Tanaka. Evaluation for clinical of a hot-wire respiratory flowmeter. *Journal of Applied Physiology*, 47(5):1131–1135, 1979.

[54] Magdalena Zimová-Herknerová and Richard Plavka. Expired tidal volumes measured by hot-wire anemometer during high-frequency oscillation in preterm infants. *Pediatric Pulmonology*, 41(5):428–433, 2006. ISSN 87556863. doi: 10.1002/ppul.20367.

[55] I. Yoshiya, T. Nakajima, I. Nagai, and S. Jitsukawa. A bidirectional respiratory flowmeter using the hot wire principle. *Journal of Applied Physiology*, 38(2):360–365, 1975. ISSN 87507587. doi: 10.1152/jappl.1975.38.2.360.

[56] W. Perez and M. J. Tobin. Separation of factors responsible for change in breathing pattern induced by instrumentation. *Journal of Applied Physiology*, 59(5):1515–1520, 1985. ISSN 87507587. doi: 10.1152/jappl.1985.59.5.1515.

[57] Ramon Pallas-Areny, Josep Colominas-Balague, and F. Javier Rosell. Effect of respiration-induced heart movements on the ECG. *IEEE Transactions on Biomedical Engineering*, v(n):585–590, 1992. ISSN 00189294. doi: 10.1109/10.29452.

[58] J. A. Hirsch and B. Bishop. Respiratory sinus arrhythmia in humans: How breathing pattern modulates heart rate. *American Journal of Physiology - Heart and Circulatory Physiology*, 10(4):620–629, 1981. ISSN 03636135.

[59] Paul Grossman and Edwin W. Taylor. Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*, 74(2):263–285, 2007. ISSN 03010511. doi: 10.1016/j.biopsycho.2005.11.014.

[60] Peter H. Charlton, Mauricio Villarroel, and Francisco Salguiero. Waveform Analysis to Estimate Respiratory Rate. In *Secondary Analysis of Electronic Health Records*, number i, pages 377–390, Cham, 2016. Springer International Publishing. ISBN 978-3-319-43740-8. doi: 10.1007/978-3-319-43742-2_26. URL `http://link.springer.com/10.1007/978-3-319-43742-2_26`.

[61] Peter H. Charlton, Drew A. Birrenkott, Timothy Bonnici, Marco A.F. Pimentel, Alistair E.W. Johnson, Jordi Alastruey, Lionel Tarassenko, Peter J. Watkinson, Richard Beale, and David A. Clifton. Breathing Rate Estimation from the Electrocardiogram and Photoplethysmogram: A Review. *IEEE Reviews in Biomedical Engineering*, 11:2–20, 2018. ISSN 19411189. doi: 10.1109/RBME.2017.2763681.

[62] Ciara O'Brien and Conor Heneghan. A comparison of algorithms for estimation of a respiratory signal from the surface electrocardiogram. *Computers in Biology and Medicine*, 37(3): 305–314, 2007. ISSN 00104825. doi: 10.1016/j.compbiomed.2006.02.002.

[63] Drew A. Birrenkott, Marco A.F. Pimentel, Peter J. Watkinson, and David A. Clifton. A robust fusion model for estimating respiratory rate from photoplethysmography and electrocardiography. *IEEE Transactions on Biomedical Engineering*, 65(9):2033–2041, 2018. ISSN 15582531. doi: 10.1109/TBME.2017.2778265.

[64] Mojtaba Nazari and Sayed Mahmoud Sakhaei. Variational Mode Extraction: A New Efficient Method to Derive Respiratory Signals from ECG. *IEEE Journal of Biomedical and Health Informatics*, 22(4):1059–1067, 2018. ISSN 21682194. doi: 10.1109/JBHI.2017.2734074.

[65] Justin Boyle, Niranjan Bidargaddi, Antti Sarela, and Mohan Karunanithi. Automatic detection of respiration rate from ambulatory single-lead ECG. *IEEE Transactions on Information Technology in Biomedicine*, 13(6):890–896, 2009. ISSN 10897771. doi: 10.1109/TITB.2009.2031239.

[66] Raquel Bailón, Leif Sörnmo, and Pablo Laguna. A robust method for ECG-based estimation of the respiratory frequency during stress testing. *IEEE Transactions on Biomedical Engineering*, 53(7):1273–1285, 2006. ISSN 00189294. doi: 10.1109/TBME.2006.871888.

[67] J. Lindemann, R. LEIACKER, G. Rettinger, and T. KECK. Nasal mucosal temperature during respiration. *Clinical Otolaryngology and Allied Sciences*, 27(3):135–139, jun 2002. ISSN 0307-7772. doi: 10.1046/j.1365-2273.2002.00544.x. URL `http://doi.wiley.com/10.1046/j.1365-2273.2002.00544.x`.

[68] Karl Storck, Matts Karlsson, Per Ask, and Dan Loyd. Heat transfer evaluation of the nasal thermistor technique. *IEEE Transactions on Biomedical Engineering*, 43(12):1187–1191, 1996. ISSN 00189294. doi: 10.1109/10.544342.

[69] R. Farré, J. M. Montserrat, M. Rotger, E. Ballester, and D. Navajas. Accuracy of thermistors and thermocouples as flow-measuring devices for detecting hypopnoeas. *European Respiratory Journal*, 11(1):179–182, 1998. ISSN 09031936. doi: 10.1183/09031936.98.11010179.

[70] Michael K. Marks, Michael South, and Bradley G. Carter. Measurement of respiratory rate and timing using a nasal thermocouple. *Journal of Clinical Monitoring*, 11(3):159–164, 1995. ISSN 07481977. doi: 10.1007/BF01617716.

[71] Yoshiaki Itasaka, Soichiro Miyazaki, Toshihiko Tanaka, Yutaka Shibata, and Kazuo Ishikawa. Detection of Respiratory Events during Polysomnography—Nasal-Oral Pressure Sensor Versus Thermocouple Airflow Sensor—. *Practica oto-rhino-laryngologica. Suppl.*, 129:60–63, 2010. doi: 10.5631/jibirinsuppl.129.60.

[72] J. M. Gehring, J. G. Cho, J. R. Wheatley, and T. C. Amis. Response characteristics for thermal and pressure devices commonly used for monitoring nasal and oral airflow during sleep studies. *Physiological Measurement*, 35(3):455–470, 2014. ISSN 13616579. doi: 10.1088/0967-3334/35/3/455.

[73] Jin Fei and Ioannis Pavlidis. Thermistor at a distance: Unobtrusive measurement of breathing. *IEEE Transactions on Biomedical Engineering*, 57(4):988–998, 2010. ISSN 00189294. doi: 10.1109/TBME.2009.2032415.

[74] Kirk H. Shelley. Photoplethysmography: Beyond the calculation of arterial oxygen saturation and heart rate. *Anesthesia and Analgesia*, 105(SUPPL. 6):13–18, 2007. ISSN 00032999. doi: 10.1213/01.ane.0000269512.82836.c9.

[75] M. Cejnar, H. Kobler, and S. N. Hunyor. Quantitative photoplethysmography: Lambert-Beer law or inverse function incorporating light scatter. *Journal of Biomedical Engineering*, 15(2): 151–154, 1993. ISSN 01415425. doi: 10.1016/0141-5425(93)90047-3.

[76] Mohamed Elgendi. On the analysis of fingertip photoplethysmogram signals. *Current cardiology reviews*, 8(1):14–25, 2012. ISSN 1875-6557. doi: 10.2174/157340312801215782.

[77] John Allen and Alan Murray. Age-related changes in the characteristics of the photoplethysmographic pulse shape at various body sites. *Physiological Measurement*, 24(2):297–307, 2003. doi: 10.1088/0967-3334/24/2/306. URL `https://doi.org/10.1088%2F0967-3334%2F24%2F2%2F306`.

[78] A. A.R. Kamal, J. B. Harness, G. Irving, and A. J. Mearns. Skin photoplethysmography - a review. *Computer Methods and Programs in Biomedicine*, 28(4):257–269, 1989. ISSN 01692607. doi: 10.1016/0169-2607(89)90159-4.

[79] D. J. Meredith, D. Clifton, P. Charlton, J. Brooks, C. W. Pugh, and L. Tarassenko. Photoplethysmographic derivation of respiratory rate: A review of relevant physiology. *Journal of Medical Engineering and Technology*, 36(1):1–7, 2012. ISSN 03091902. doi: 10.3109/03091902.2011.638965.

[80] Walter Karlen, Srinivas Raman, J. Mark Ansermino, and Guy A. Dumont. Multiparameter respiratory rate estimation from the photoplethysmogram. *IEEE Transactions on Biomedical Engineering*, 60(7):1946–1953, 2013. ISSN 00189294. doi: 10.1109/TBME.2013.2246160.

[81] Paul S. Addison, James N. Watson, Michael L. Mestek, and Roger S. Mecca. Developing an algorithm for pulse oximetry derived respiratory rate (RRoxi): A healthy volunteer study. *Journal of Clinical Monitoring and Computing*, 26(1):45–51, 2012. ISSN 13871307. doi: 10.1007/s10877-011-9332-y.

[82] Shishir Dash, Kirk H. Shelley, David G. Silverman, and Ki H. Chon. Estimation of respiratory rate from ECG, photoplethysmogram, and piezoelectric pulse transducer signals: A comparative study of time-frequency methods. *IEEE Transactions on Biomedical Engineering*, 57 (5):1099–1107, 2010. ISSN 00189294. doi: 10.1109/TBME.2009.2038226.

[83] G. John. Gibson. *Clinical Tests of Respiratory Function*. Chapman & Hall Medical, 2nd edition, 1998. ISBN 9780412568909.

[84] Jayant Sirohi and Inderjit Chopra. Fundamental understanding of piezoelectric strain sensors. *Journal of Intelligent Material Systems and Structures*, 11(4):246–257, 2000. ISSN 1045389X. doi: 10.1106/8BFB-GC8P-XQ47-YCQ0.

[85] SS5LB Biopac Systems, 2019. URL `https://www.biopac.com/product/respiratory-effort-transducer-bsl/`. Accessed: 30-11-2019.

[86] Piezo Crystal Effort Sensor Kit, SleepSense, 2019. URL `https://sleepsense.com/shop/sleepsense-respiratory-effort-sensors/piezo-crystal-effort-sensor-kit-double-buckle-safety-din-connectors/`. Accessed: 30-11-2019.

[87] Zhengbo Zhang, Jiewen Zheng, Hao Wu, Weidong Wang, Buqing Wang, and Hongyun Liu. Development of a Respiratory Inductive Plethysmography Module Supporting Multiple Sensors for Wearable Systems. *Sensors*, 12(12):13167–13184, sep 2012. ISSN 1424-8220. doi: 10.3390/s121013167. URL `http://www.mdpi.com/1424-8220/12/10/13167`.

[88] S. Chadha, H. Watson, s. Birch, G. A. Jenouri, A. W. Schneider, M. A. Cohn, and M. A. Sackner. Validation of respiratory inductive plethysmography using different calibration procedures. *American Review of Respiratory Disease*, 125(6):644–649, 1982. ISSN 00030805. doi: 10.1164/arrd.1982.125.6.644.

[89] Jean Paul Cantineau, Pierre Escourrou, Richard Sartene, Claude Gaultier, and Michael Goldman. Accuracy of Respiratory Inductive Plethysmography during Wakefulness and Sleep in Patients with Obstructive Sleep Apnea. *Chest*, 102(4):1145–1151, oct 1992. ISSN 00123692. doi: 10.1378/chest.102.4.1145. URL `https://linkinghub.elsevier.com/retrieve/pii/S0012369216342969`.

[90] J. D. Sackner, A. J. Nixon, B. Davis, N. Atkins, and M. A. Sackner. Non-invasive measurement of ventilation during exercise using a respiratory inductive plethysmograph. I. *American Review of Respiratory Disease*, 122(6):867–871, 1980. ISSN 00030805. doi: 10.1164/arrd.1980.122.6.867.

[91] Pierre Yves Carry, Pierre Baconnier, Andre Eberhard, Pierre Cotte, and Gila Benchetrit. Evaluation of respiratory inductive plethysmography: Accuracy for analysis of respiratory waveforms. *Chest*, 111(4):910–915, 1997. ISSN 00123692. doi: 10.1378/chest.111.4.910. URL `http://dx.doi.org/10.1378/chest.111.4.910`.

[92] Marie-Noëlle Fiamma, Ziyad Samara, Pierre Baconnier, Thomas Similowski, and Christian Straus. Respiratory inductive plethysmography to assess respiratory variability and complexity in humans. *Respiratory Physiology & Neurobiology*, 156(2):234–239, may 2007. ISSN 15699048. doi: 10.1016/j.resp.2006.12.001. URL `https://linkinghub.elsevier.com/retrieve/pii/S156990480600334X`.

[93] Gregor Brüllmann, Karsten Fritsch, Robert Thurnheer, and Konrad E. Bloch. Respiratory monitoring by inductive plethysmography in unrestrained subjects using position sensor-adjusted calibration. *Respiration*, 79(2):112–120, 2009. ISSN 00257931. doi: 10.1159/000212117. URL `https://www.karger.com/Article/FullText/212117`.

[94] BioSignalsPlux. RIP Sensor. URL `https://www.biosignalsplux.com/index.php/rsp-respiration`.

[95] David Girbau, Antonio Lázaro, Ángel Ramos, and Ramón Villarino. Remote sensing of vital signs using a doppler radar and diversity to overcome null detection. *IEEE Sensors Journal*, 12(3):512–518, 2012. ISSN 1530437X. doi: 10.1109/JSEN.2011.2107736.

[96] Amy Droitcour, Victor Lubecke, Jenshan Lin, and Olga Boric-Lubecke. A microwave radio for Doppler radar sensing of vital signs. *IEEE MTT-S International Microwave Symposium Digest*, 3:175–178, 2001. ISSN 0149645X. doi: 10.1109/MWSYM.2001.966866.

[97] Mehrdad Nosrati and Negar Tavassolian. Accurate Doppler Radar-Based Cardiopulmonary Sensing Using Chest-Wall Acceleration. *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, 3(1):41–47, 2019. ISSN 24697249. doi: 10.1109/JERM.2018.2879452.

[98] Byung-Kwon Park, Shuhei Yamada, Olga Boric-Lubecke, and Victor Lubecke. Single-channel receiver limitations in doppler radar measurements of periodic motion. In *2006 IEEE Radio and Wireless Symposium*, volume 2006, pages 99–102. IEEE, 2006. ISBN 0-7803-9412-7. doi: 10.1109/RWS.2006.1615104. URL `http://ieeexplore.ieee.org/document/1615104/`.

[99] Ming-chun Huang, Jason J Liu, Wenyao Xu, Changzhan Gu, Changzhi Li, and Majid Sarrafzadeh. A Self-Calibrating Radar Sensor System for Measuring Vital Signs. 10(2):352–363, 2016.

[100] Seyed Mohammad Ali Tayaranian Hosseini and Hamidreza Amindavar. A New Ka-Band Doppler Radar in Robust and Precise Cardiopulmonary Remote Sensing. *IEEE Transactions on Instrumentation and Measurement*, 66(11):3012–3022, 2017. ISSN 00189456. doi: 10.1109/TIM.2017.2714480.

[101] Changzhan Gu, Yuchu He, and Jiang Zhu. Noncontact Vital Sensing With a Miniaturized 2.4 GHz Circularly Polarized Doppler Radar. *IEEE Sensors Letters*, 3(7):1–4, 2019. ISSN 2475-1472. doi: 10.1109/lsens.2019.2924695.

[102] Yuchu He, Changzhan Gu, Haijiang Ma, Jiang Zhu, and George V. Eleftheriades. Miniaturized Circularly Polarized Doppler Radar for Human Vital Sign Detection. *IEEE Transactions on Antennas and Propagation*, 67(11):7022–7030, 2019. ISSN 15582221. doi: 10.1109/TAP.2019.2927777.

[103] Yuri Feldman, Alexander Puzenko, Paul Ben Ishai, Andreas Caduff, Issak Davidovich, Fadi Sakran, and Aharon J. Agranat. The electromagnetic response of human skin in the millimetre and submillimetre wave range. *Physics in Medicine and Biology*, 54(11):3341–3363, 2009. ISSN 00319155. doi: 10.1088/0031-9155/54/11/005.

[104] Lydia Chioukh, Halim Boutayeb, Ke Wu, and Dominic Deslandes. f / nf Harmonic Radar System with Optimal Detection of Vital Signs. pages 25–28, 2012.

[105] Yanming Xiao, Changzhi Li, and Jenshan Lin. A portable noncontact heartbeat and respiration monitoring system using 5-GHz radar. *IEEE Sensors Journal*, 7(7):1042–1043, 2007. ISSN 15581748. doi: 10.1109/JSEN.2007.895979.

[106] Yee Siong Lee, Pubudu N Pathirana, Robin J Evans, and Christopher L Steinfort. Noncontact Detection and Analysis of Respiratory Function Using Microwave Doppler Radar. *Journal of Sensors*, 2015(February):1–13, 2015. ISSN 1687-725X. doi: 10.1155/2015/548136. URL `http://www.hindawi.com/journals/js/2015/548136/`.

[107] Amy D. Droitcour, Olga Boric-Lubecke, Victor M. Lubecke, Jenshan Lin, and Gregory T.A. Kovacs. Range correlation and I/Q performance benefits in single-chip silicon Doppler radars for noncontact cardiopulmonary monitoring. *IEEE Transactions on Microwave Theory and Techniques*, 52(3):838–848, 2004. ISSN 00189480. doi: 10.1109/TMTT.2004.823552.

[108] Ashikur Rahman, Victor M. Lubecke, Olga Boric-Lubecke, Jan H. Prins, and Takuya Sakamoto. Doppler Radar Techniques for Accurate Respiration Characterization and Subject Identification. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 8(2): 350–359, 2018. ISSN 21563357. doi: 10.1109/JETCAS.2018.2818181.

[109] Marco Mercuri, Yao Hong Liu, Ilde Lorato, Tom Torfs, Fokko Wieringa, Andre Bourdoux, and Chris Van Hoof. A Direct Phase-Tracking Doppler Radar Using Wavelet Independent

Component Analysis for Non-Contact Respiratory and Heart Rate Monitoring. *IEEE Transactions on Biomedical Circuits and Systems*, 12(3):632–643, 2018. ISSN 19324545. doi: 10.1109/TBCAS.2018.2813013.

[110] Giovanni Cerasuolo, Orsola Petrella, Luigi Marciano, Francesco Soldovieri, and Gianluca Gennarelli. Metrological characterization for vital sign detection by a bioradar. *Remote Sensing*, 9(10), 2017. ISSN 20724292. doi: 10.3390/rs9100996.

[111] Yao-hong Liu, Sunil Sheelavant, Marco Mercuri, Paul Mateman, and Masoud Babaie. An Ultralow Power Burst-Chirp UWB Radar Transceiver for Indoor Vital Signs and Occupancy Sensing in 40-nm CMOS. *IEEE Solid-State Circuits Letters*, 2(11):256–259, nov 2019. ISSN 2573-9603. doi: 10.1109/LSSC.2019.2951423. URL `https://ieeexplore.ieee.org/document/8891756/`.

[112] IEEE Standard for Safety Levels with Respect to Human Exposure to Electric, Magnetic, and Electromagnetic Fields, 0 Hz to 300 GHz. *IEEE Std C95.1-2019 (Revision of IEEE Std C95.1-2005/ Incorporates IEEE Std C95.1-2019/Cor 1-2019)*, pages 1–312, 2019. ISSN null. doi: 10.1109/IEEESTD.2019.8859679.

[113] Guidelines for limiting exposure to time-varying electric, magnetic and electromagnetic fields (100 kHz TO 300 GHz). Technical report, ICNIRP, 2018. URL `https://www.icnirp.org/cms/upload/consultation_upload/ICNIRP_RF_Guidelines_PCD_2018_07_11.pdf`.

[114] Rikke Gade and Thomas B. Moeslund. Thermal cameras and applications: A survey. *Machine Vision and Applications*, 25(1):245–262, 2014. ISSN 09328092. doi: 10.1007/s00138-013-0570-5.

[115] Abbas K Abbas, Konrad Heimann, Katrin Jergus, Thorsten Orlikowsky, and Steffen Leonhardt. Neonatal non-contact respiratory monitoring based on real-time infrared thermography. *BioMedical Engineering OnLine*, 10(1):93, 2011. ISSN 1475-925X. doi: 10.1186/1475-925X-10-93.

[116] Abdulkadir H. Alkali, Reza Saatchi, Heather Elphick, and Derek Burke. Thermal image processing for real-time non-contact respiration rate monitoring. *IET Circuits, Devices and Systems*, 11(2):142–148, 2017. ISSN 1751858X. doi: 10.1049/iet-cds.2016.0143.

[117] Carina Barbosa Pereira, Xinchi Yu, Tom Goos, Irwin Reiss, Thorsten Orlikowsky, Konrad Heimann, Boudewijn Venema, Vladimir Blazek, Steffen Leonhardt, and Daniel Teichmann. Noncontact Monitoring of Respiratory Rate in Newborn Infants Using Thermal Imaging. *IEEE Transactions on Biomedical Engineering*, 66(4):1105–1114, 2019. ISSN 15582531. doi: 10.1109/TBME.2018.2866878.

[118] H. Budzier and G. Gerlach. Calibration of uncooled thermal infrared cameras. *Journal of Sensors and Sensor Systems*, 2015. ISSN 2194-878X. doi: 10.5194/jsss-4-187-2015.

[119] Othon Gonzalez-Chavez, Daniel Cardenas-Garcia, Sertac Karaman, Mariano Lizarraga, and Joaquin Salas. Radiometric Calibration of Digital Counts of Infrared Thermal Cameras. *IEEE Transactions on Instrumentation and Measurement*, 68(11):4387–4399, 2019. ISSN 15579662. doi: 10.1109/TIM.2018.2887070.

[120] Shourjya Sanyal and Koushik Kumar Nundy. Algorithms for Monitoring Heart Rate and Respiratory Rate From the Video of a User's Face. *IEEE Journal of Translational Engineering in Health and Medicine*, 6(May 2017):1–11, 2018. ISSN 21682372. doi: 10.1109/JTEHM.2018.2818687.

[121] Carlo Massaroni, Daniel Simões Lopes, Daniela Lo Presti, Emiliano Schena, and Sergio Silvestri. Contactless Monitoring of Breathing Patterns and Respiratory Rate at the Pit of the Neck: A Single Camera Approach. *Journal of Sensors*, 2018:1–13, 2018. ISSN 1687-725X. doi: 10.1155/2018/4567213.

[122] Flavia Benetazzo, Alessandro Freddi, Andrea Monteriù, and Sauro Longhi. Respiratory rate detection algorithm based on RGB-D camera: theoretical background and experimental results. *Healthcare Technology Letters*, 1(3):81–86, sep 2014. ISSN 2053-3713. doi: 10.1049/htl.2014.0063. URL `http://digital-library.theiet.org/content/journals/10.1049/htl.2014.0063`.

[123] Marek Bartula, Timo Tigges, and Jens Muehlsteff. Camera-based system for contactless monitoring of respiration. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pages 2672–2675, 2013. ISSN 1557170X. doi: 10.1109/EMBC.2013.6610090.

[124] Kazuki Nakajima, Yoshiaki Matsumoto, and Toshiyo Tamura. A monitor for posture changes and respiration in bed using real time image sequence analysis. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, 1:51–54, 2000. ISSN 05891019. doi: 10.1109/IEMBS.2000.900665.

[125] Jean-Yves Bouguet. Pyramidal implementation of the affine lucas kanade feature tracker—description of the algorithm. Technical report, 2001. URL `http://robots.stanford.edu/cs223b04/algo_tracking.pdf`.

[126] Gunnar Farnebäck. Two-Frame Motion Estimation Based on Polynomial Expansion. In *Lecture Notes in Computer Science*, volume 2749, pages 363–370. 2003. ISBN 978-3-540-40601-3. doi: 10.1007/3-540-45103-X_50. URL `http://link.springer.com/10.1007/3-540-45103-X_50`.

[127] Till Kroeger, Radu Timofte, Dengxin Dai, and Luc Van Gool. Fast optical flow using dense inverse search. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9908 LNCS:471–488, 2016. ISSN 16113349. doi: 10.1007/978-3-319-46493-0_29.

[128] D. Lucas; Takeo Kanade Bruce. An Iterative Image Registration Technique with an Application to Stereo Vision. In *International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.

[129] Rik Janssen, Wenjin Wang, Andreia Moço, and Gerard De Haan. Video-based respiration monitoring with automatic region of interest detection. *Physiological Measurement*, 37(1): 100–114, 2015. ISSN 13616579. doi: 10.1088/0967-3334/37/1/100.

[130] A. P. Prathosh, Pragathi Praveena, Lalit K. Mestha, and Sanjay Bharadwaj. Estimation of Respiratory Pattern From Video Using Selective Ensemble Aggregation. *IEEE Transactions on Signal Processing*, 65(11):2902–2916, 2017. ISSN 1053587X. doi: 10.1109/TSP.2017.2664048.

[131] Michael H. Li, Azadeh Yadollahi, and Babak Taati. Noncontact Vision-Based Cardiopulmonary Monitoring in Different Sleeping Positions. *IEEE Journal of Biomedical and Health Informatics*, 21(5):1367–1375, sep 2017. ISSN 2168-2194. doi: 10.1109/JBHI.2016.2567298. URL `http://ieeexplore.ieee.org/document/7468522/`.

[132] Intel. Intel ® RealSense TM Camera SR300 Embedded Coded Light 3D Imaging System with Full High Definition Color Camera, 2016. URL `https://www.intelrealsense.com/wp-content/uploads/2019/07/RealSense_SR30x_Product_Datasheet_Rev_002.pdf`.

[133] Intel. Intel ® RealSense TM 3D Camera ZR300, 2017. URL `https://click.intel.com/media/ZR300-Product-Datasheet-Public-002.pdf`.

[134] Intel. D400 Series Product Family - Datasheet. (January), 2019. URL `https://www.intelrealsense.com/wp-content/uploads/2019/10/Intel-RealSense-D400-Series-Datasheet-Oct-2019.pdf`.

[135] Aleš Procházka, Martin Schätz, Oldřich Vyšata, and Martin Vališ. Microsoft Kinect Visual and Depth Sensors for Breathing and Heart Rate Analysis. *Sensors*, 16(12):996, jun 2016. ISSN 1424-8220. doi: 10.3390/s16070996. URL `http://www.mdpi.com/1424-8220/16/7/996`.

[136] Jochen Kempfle and Kristof Van Laerhoven. Respiration Rate Estimation with Depth Cameras. pages 1–10, 2018. doi: 10.1145/3266157.3266208.

[137] Shinobu Kumagai, Ryohei Uemura, Toru Ishibashi, Susumu Nakabayashi, Norikazu Arai, Takenori Kobayashi, and Jun'ichi Kotoku. Markerless Respiratory Motion Tracking Using Single Depth Camera. *Open Journal of Medical Imaging*, 06(01):20–31, 2016. ISSN 2164-2788. doi: 10.4236/ojmi.2016.61003. URL `http://www.scirp.org/journal/doi.aspx?DOI=10.4236/ojmi.2016.61003`.

[138] Evan Silverstein and Michael Snyder. Comparative analysis of respiratory motion tracking using Microsoft Kinect v2 sensor. *Journal of Applied Clinical Medical Physics*, 19(3):193–204, 2018. ISSN 15269914. doi: 10.1002/acm2.12318.

[139] Edgar A. Bernal, Lalit K. Mestha, and Eribaweimon Shilla. Non contact monitoring of respiratory function via depth sensing. *2014 IEEE-EMBS International Conference on Biomedical and Health Informatics, BHI 2014*, pages 101–104, 2014. doi: 10.1109/BHI.2014.6864314.

[140] Manuel Martinez and Rainer Stiefelhagen. Breathing rate monitoring during sleep from a depth camera under real-life conditions. *Proceedings - 2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017*, pages 1168–1176, 2017. doi: 10.1109/WACV.2017.135.

[141] Mireya Fernandez Chimeno, Juan Ramos Castro, Miguel Angel García Gonzalez, Federico Guede Fernandez, Marc Mateu Mateus, Noelia Rodriguez Ibañez, Bernat Bas Pujols, and Jose Maria Alvarez Gomez. Respiratory Signal Extraction, 2018. URL `https://patentscope.wipo.int/search/en/detail.jsf?docId=WO2018121861`.

[142] Christoph Bruser, Christoph Hoog Antink, Tobias Wartzek, Marian Walter, and Steffen Leonhardt. Ambient and Unobtrusive Cardiorespiratory Monitoring Techniques. *IEEE Reviews in Biomedical Engineering*, 8:30–43, 2015. ISSN 1937-3333. doi: 10.1109/RBME.2015.2414661. URL `http://ieeexplore.ieee.org/document/7063243/`.

[143] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. ORB: An efficient alternative to SIFT or SURF. In *2011 International Conference on Computer Vision*, pages 2564–2571. IEEE, nov 2011. ISBN 978-1-4577-1102-2. doi: 10.1109/ICCV.2011.6126544. URL `http://ieeexplore.ieee.org/document/6126544/`.

[144] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004. ISSN 09205691. doi: 10.1023/B:VISI.0000029664.99615.94.

[145] Marius Muja and David G Lowe. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. *International Conference on Computer Vision Theory and Applications (VISAPP '09)*, pages 1–10, 2009. ISSN 00301299. doi: 10.1.1.160.1721.

[146] Marius Muja and David G Lowe. Scalable nearest neighbor algorithms for high dimensional data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(11):2227–2240, 2014. ISSN 01628828. doi: 10.1109/TPAMI.2014.2321376.

[147] Martin a Fischler and Robert C Bolles. Random Sample Paradigm for Model Consensus: A Apphcatlons to Image Fitting with Analysis and Automated Cartography. *Communications of the ACM*, 24(6):381–395, 1981. ISSN 00010782. doi: 10.1145/358669.358692.

[148] Stuart P. Lloyd. Least Squares Quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982. ISSN 15579654. doi: 10.1109/TIT.1982.1056489.

[149] Tapas Kanungo, D.M. Mount, N.S. Netanyahu, C.D. Piatko, Ruth Silverman, and A.Y. Wu. An efficient k-means clustering algorithm: analysis and implementation. *IEEE Transactions*

on Pattern Analysis and Machine Intelligence, 24(7):881–892, 2002. ISSN 0162-8828. doi: 10.1109/TPAMI.2002.1017616.

[150] E. Vincent, R. Laganiere, T Vincent, and R. Laganiere. Detecting planar homographies in an image pair. *2nd International Symposium on Image and Signal Processing and Analysis*, 0(2):182–187, 2001. doi: 10.1109/ISPA.2001.938625. URL `http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=938625`.

[151] Anders Hast, Johan Nysjö, and Andrea Marchetti. Optimal RANSAC - Towards a repeatable algorithm for finding the optimal set. *Journal of WSCG*, 21(1):21–30, 2013. ISSN 12136972. doi: 10.1.1.699.9947.

[152] Alexander; Alekhin, Steven; Puttemans, and Joao Cartucho. *lk_track.py*. OpenCV, 2016. URL `https://github.com/opencv/opencv/blob/master/samples/python/lk_track.py`. [Source Code].

[153] World Medical Association. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *Bulletin of the World Health Organization*, 79(4):373–4, 2001. ISSN 0042-9686. URL `http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2566407/`.

[154] Gonzalo R. Arce. *Nonlinear Signal Processing: A Statistical Approach*. Wiley, 2004. ISBN 978-0-471-67624-9. URL `https://www.wiley.com/en-es/Nonlinear+Signal+Processing:+A+Statistical+Approach-p-9780471676249`.

[155] Arturo Martinez, Raul Alcaraz, and Jose Joaquin Rieta. Application of the phasor transform for automatic delineation of single-lead ECG fiducial points. *Physiological Measurement*, 31 (11):1467–1485, 2010. ISSN 09673334. doi: 10.1088/0967-3334/31/11/005.

[156] Ronald Aylmer Fisher. *Statistical methods for research workers*, volume 5. Oliver and Boyd, 6 edition, 1936.

[157] Davide Giavarina. Understanding Bland Altman analysis. *Biochemia Medica*, 25(2):141–151, 2015. ISSN 13300962. doi: 10.11613/BM.2015.015.

[158] Microsoft Research. Kinect for Windows SDK beta, 2011. URL `https://www.microsoft.com/en-us/research/project/kinect-for-windows-sdk-beta/`.

[159] Paul Viola and Michael J. Jones. Robust Real-Time Face Detection. *International Journal of Computer Vision*, 57(2):137–154, may 2004. ISSN 0920-5691. doi: 10.1023/B:VISI.0000013087.49260.fb. URL `http://link.springer.com/article/10.1023/B:VISI.0000013087.49260.fb`.

[160] Rainer Lienhart, Alexander Kuranov, and Vadim Pisarevsky. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. *Proceedings of the 25th*

DAGM Pattern Recognition Symposium, pages 297–304, 2003. ISSN 0021-9541. doi: 10.1007/978-3-540-45243-0{_}39.

[161] Jianbo Shi and Tomasi. Good features to track. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94*, pages 593–600. IEEE Comput. Soc. Press, 1994. ISBN 0-8186-5825-8. doi: 10.1109/CVPR.1994.323794. URL `http://ieeexplore.ieee.org/document/323794/`.

[162] Christina Orphanidou. Derivation of respiration rate from ambulatory ECG and PPG using Ensemble Empirical Mode Decomposition: Comparison and fusion. *Computers in Biology and Medicine*, 81(September 2016):45–54, 2017. ISSN 18790534. doi: 10.1016/j.compbiomed.2016.12.005. URL `http://dx.doi.org/10.1016/j.compbiomed.2016.12.005`.

[163] R. Balocchi, D. Menicucci, E. Santarcangelo, L. Sebastiani, A. Gemignani, B. Ghelarducci, and M. Varanini. Deriving the respiratory sinus arrhythmia from the heartbeat time series using empirical mode decomposition. *Chaos, Solitons and Fractals*, 20(1):171–177, 2004. ISSN 09600779. doi: 10.1016/S0960-0779(03)00441-7.

[164] Sung Bin Park, Yeon Sik Noh, Sung Jun Park, and Hyoung Ro Yoon. An improved algorithm for respiration signal extraction from electrocardiogram measured by conductive textile electrodes using instantaneous frequency estimation. *Medical and Biological Engineering and Computing*, 46(2):147–158, 2008. ISSN 01400118. doi: 10.1007/s11517-007-0302-y.

[165] Ds S Benitez, Pa a Gaydecki, a Zaidi, and a P Fitzpatrick. A new QRS detection algorithm based on the Hilbert transform. *Computing in Cardiology*, 27:379–382, 2000. ISSN 0276-6547. doi: 10.1109/CIC.2000.898536.

[166] Santanu Sahoo, Prativa Biswal, Tejaswini Das, and Sukanta Sabut. De-noising of ECG Signal and QRS Detection Using Hilbert Transform and Adaptive Thresholding. *Procedia Technology*, 25(Raerest):68–75, 2016. ISSN 22120173. doi: 10.1016/j.protcy.2016.08.082. URL `http://linkinghub.elsevier.com/retrieve/pii/S2212017316304297`.

[167] Boualem Boashash. Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals. *Proceedings of the IEEE*, 80(4):520–538, apr 1992. ISSN 00189219. doi: 10.1109/5.135376. URL `http://ieeexplore.ieee.org/document/135376/`.

[168] Robert J. Hodrick and Edward C. Prescott. Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking*, 29(1):1, feb 1997. ISSN 00222879. doi: 10.2307/2953682. URL `http://www.jstor.org/stable/2953682?origin=crossref`.

[169] J. P. Royston. An Extension of Shapiro and Wilk's W Test for Normality to Large Samples. *Applied Statistics*, 31(2):115, 1982. ISSN 00359254. doi: 10.2307/2347973. URL `http://www.jstor.org/stable/10.2307/2347973?origin=crossref`.

[170] Myles Hollander, Douglas A. Wolfe, and Eric Chicken. *The Two-Way Layout*, pages 289–392. John Wiley & Sons, Inc., Hoboken, NJ, USA, jul 2015. doi: 10.1002/9781119196037.ch7. URL `http://doi.wiley.com/10.1002/9781119196037.ch7`.

[171] Davis E. King. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10:1755–1758, 2009.

[172] Morgan Quigley, Ken Conley, Brian P. Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y. Ng. Ros: an open-source robot operating system. In *ICRA Workshop on Open Source Software*, 2009. URL `http://www.willowgarage.com/sites/default/files/icraoss09-ROS.pdf`.

[173] T W Anderson and D A Darling. Asymptotic Theory of Certain. *Ann. Math. Statist.*, (2): 193–212. doi: 10.1214/aoms/1177729437.

[174] Robert V Hogg and Johannes Ledolter. *Engineering statistics.* New York : Macmillan ; London : Collier Macmillan, 1987. ISBN 0023557907.

[175] William H Kruskal and W Allen Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621, 1952.

[176] Danielle C. Gomes, Ingrid G. Azevedo, Ana G. Figueiredo Araújo, Lenice D. Costa Lopes, Danilo A. P. Nagem, Fabrício A. Magalhães, and Silvana A. Pereira. Thoracoabdominal motion in newborns: Reliability between two interactive computing environments. *Pediatric Pulmonology*, (May 2019):1–6, 2020. ISSN 8755-6863. doi: 10.1002/ppul.24709.

[177] Ingrid Daubechies, Jianfeng Lu, and Hau Tieng Wu. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. *Applied and Computational Harmonic Analysis*, 30(2):243–261, 2011. ISSN 10635203. doi: 10.1016/j.acha.2010.08.002. URL `http://dx.doi.org/10.1016/j.acha.2010.08.002`.

[178] Hau Tieng Wu, Yi Hsin Chan, Yu Ting Lin, and Yung Hsin Yeh. Using synchrosqueezing transform to discover breathing dynamics from ECG signals. *Applied and Computational Harmonic Analysis*, 36(2):354–359, 2014. ISSN 10635203. doi: 10.1016/j.acha.2013.07.003. URL `http://dx.doi.org/10.1016/j.acha.2013.07.003`.

[179] N. Rodríguez-Ibáñez, M. A. García-González, M. Fernández-Chimeno, H. De Rosario, and J. Ramos-Castro. Synchrosqueezing Index for Detecting Drowsiness Based on the Respiratory Effort Signal. In *IFMBE Proceedings*, volume 41, pages 965–968, 2014. doi: 10.1007/978-3-319-00846-2_239. URL `http://link.springer.com/10.1007/978-3-319-00846-2_239`.

[180] Paul S. Addison. Modular continuous wavelet processing of biosignals: Extracting heart rate and oxygen saturation from a video signal. *Healthcare Technology Letters*, 3(2):111–115, 2016. ISSN 20533713. doi: 10.1049/htl.2015.0052.

[181] Lin Li, Haiyan Cai, and Qingtang Jiang. Adaptive synchrosqueezing transform with a time-varying parameter for non-stationary signal separation. *Applied and Computational Harmonic Analysis*, 1(61201287):1–32, 2019. ISSN 1096603X. doi: 10.1016/j.acha.2019.06.002. URL `https://doi.org/10.1016/j.acha.2019.06.002`.

[182] William H Kruskal and W Allen Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621, 1952.

[183] Thorsten Pohlert. *The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR)*, 2014. URL `https://CRAN.R-project.org/package=PMCMR`. R package.

[184] Norden E. Huang, Zheng Shen, Steven R. Long, Manli C. Wu, Hsing H. Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, mar 1998. ISSN 1364-5021. doi: 10.1098/rspa.1998.0193. URL `https://doi.org/10.1098/rspa.1998.0193`.

[185] T. Nimmy John, Puthankattil Subha Dharmapalan, and N. Ramshekhar Menon. Exploration of time-frequency reassignment and homologous inter-hemispheric asymmetry analysis of MCI-AD brain activity. *BMC Neuroscience*, 20(1):1–14, 2019. ISSN 14712202. doi: 10.1186/s12868-019-0519-3. URL `https://doi.org/10.1186/s12868-019-0519-3`.

[186] Muammar M. Kabir, Reza Tafreshi, Diane B. Boivin, and Naim Haddad. Enhanced automated sleep spindle detection algorithm based on synchrosqueezing. *Medical and Biological Engineering and Computing*, 53(7):635–644, 2015. ISSN 17410444. doi: 10.1007/s11517-015-1265-z. URL `http://dx.doi.org/10.1007/s11517-015-1265-z`.

[187] Mallat Stéphane. CHAPTER 4 - Time Meets Frequency. In Mallat Stéphane, editor, *A Wavelet Tour of Signal Processing (Third Edition)*, pages 89–153, Boston, 2009. Academic Press. ISBN 978-0-12-374370-1. doi: https://doi.org/10.1016/B978-0-12-374370-1.00008-2. URL `http://www.sciencedirect.com/science/article/pii/B9780123743701000082`.

[188] Christopher Torrence and Gilbert P. Compo. A Practical Guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79(1):61–78, jan 1998. ISSN 0003-0007. doi: 10.1175/1520-0477(1998)079⟨0061:APGTWA⟩2.0.CO;2. URL `http://journals.ametsoc.org/doi/abs/10.1175/1520-0477%281998%29079%3C0061%3AAPGTWA%3E2.0.CO%3B2`.

[189] Marie Farge. Wavelet Transforms And Their Applications To Turbulence. *Annual Review of Fluid Mechanics*, 24(1):395–458, 1992. doi: 10.1146/annurev.fl.24.010192.002143. URL `https://doi.org/10.1146/annurev.fl.24.010192.002143`.