· U C ·

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Carlos Manuel Silva Lúcio

Voluntary Cough Detection By Internal Sound Analysis

Dissertação apresentada à Universidade de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Biomédica

Orientador(es):
Prof. Dr. Rui Pedro Paiva (Orientador, DEIUC)
Prof. Dr. Paulo Carvalho (Co-orientador, DEIUC)

Coimbra, 2014

Este trabalho foi desenvolvido em colaboração com:

Departamento de Engenharia Informática
da Faculdade de Ciências e Tecnologia da
Universidade de Coimbra

WELCOME Project

CISUC – Centre for Informatics and Systems of
University of Coimbra

Esta cópia da tese é fornecida na condição de que quem a consulta reconhece que os direitos de autor são pertença do autor da tese e que nenhuma citação ou informação obtida a partir dela pode ser publicada sem a referência apropriada.

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognize that its copyright rests with its author and that no quotation from the thesis and no information derived from it may be published without proper acknowledgement.

Acknowledgements

This work was conducted under the guidance of Professors Rui Pedro Paiva and Paulo Carvalho, to whom I must express my sincere appreciation for the support, which the present work reflects. I extend to them my grateful admiration.

A special recognition I would like to address, in Portuguese, to my family and friends. *Pelo vosso perseverante e carinhoso apoio, compreensão e disponibilidade, o meu mais verdadeiro agradecimento. Além do presente trabalho, tenho o coração grato pelo agradável e incomparável amparo que sempre foram.*

Abstract

Cough can be defined as a forced expulsive onrush, normally against a closed glottis, producing a characteristic three-phase sound, and as a symptom, it can be an indicator of many respiratory diseases. An objective measure of cough would be of use in clinical practice, clinical research and the assessment of novel therapies and pharmaceuticals.

In the present work, a method to automatically identify, count and (partly) qualify cough sounds, based on internal sound signals, is proposed. This approach relies on explosive phase detection, because of its acoustic and spectral distinctive characteristics, and its potential for accurate onset detection of cough sounds. The features analyzed, related with tonality, pitch, timbre and frequency, prove to be very relevant in our explosive phase detection approach. Our results show an accurate detection, for a wide testing population with and without respiratory perturbations, which demonstrates the ruggedness of this approach. The internal sound analysis reveals advantageous in external noise reduction, therefore internal sounds are highlighted and better characterized. The explosive phase detection approach demonstrates usefulness in detecting the onset of cough sounds.

Resumo

A tosse pode ser definida como um evento explosivo forçado, normalmente contra a glote fechada, produzindo um som característico com 3 fases, e como sintoma, pode ser um indicador de muitas doenças respiratórias. Uma medição objetiva da tosse seria útil na prática e pesquisa clínica, e na avaliação de terapias e produtos farmacêuticos inovadores.

O presente trabalho propõe um método baseado no som interno para automaticamente identificar, contar e parcialmente qualificar sons de tosse. Esta abordagem basea-se na deteção de fase explosiva, devido à sua acústica e às características distintivas no espetro, e ao seu potencial para a deteção precisa do início dos sons de tosse. As características analisadas, relacionadas com tonalidade, entoação, timbre e frequência, revelam-se muito relevantes na abordagem de deteção da fase explosiva. Os nossos resultados evidenciam uma boa deteção, para uma vasta população de teste, com e sem perturbações respiratórias, o que demonstra a robustez desta abordagem. A análise do som interno revela-se vantajosa na redução de ruído externo, portanto, os sons internos são realçados e melhor caracterizados. A abordagem de deteção da fase explosiva demonstra utilidade na deteção do início dos sons de tosse.

List of Acronyms

1^{st}	First
2^{nd}	Second
BMEI	International Conference on BioMedical Engineering and Informatics
COPD	Chronic obstructive pulmonary disease
EIT	Electrical impedance tomography
ERS	European Respiratory Society
FFT	Fast Fourier transform
HACC	Hull Automatic Cough Counter
HMM	Hidden Markov Models
KNN	K-nearest neighbors
LCM	Leicester Cough Monitor
LS	LifeShirt®
MFCC	Mel frequency cepstral coefficient
MOBIHEALTH	International Conference on Wireless Mobile Communication and Healthcare
PNN	Probabilistic neural network
RPG	With respiratory perturbations group
SpO_2	Saturation of peripheral oxygen
STD	Standard deviation
WELCOME	Wearable Sensing and Smart Cloud Computing for Integrated Care to COPD Patients with Co-morbidities
WPG	Without respiratory perturbations group

List of Figures

List of Tables

Contents

Chapter 1

Introduction

1.1 Scope

The present master thesis is part of the European project Wearable Sensing and Smart Cloud Computing for Integrated Care to COPD Patients with Co-morbidities (WELCOME), which aims to bring about a change in the management of chronic diseases and in particular the Chronic Obstructive Pulmonary Disease (COPD). The project is intended to produce a patient centered approach to COPD management, by the design, development and evaluation of a platform that will integrate the fundamental elements of care into a unified system targeting COPD for early detection of complications. The combination of continuous monitoring, information and communication technologies, shared decision support systems, and personalized guidance will provide a shift from reactive to predictive, preventive, personalized, and participatory medicine. The project includes the development of a vest with a large number of non-invasive chest sensors and devices dedicated to the treatment of diabetes for measuring and monitoring various parameters, like high spatial resolution electrocardiogram, chest sounds, Electrical Impedance

Tomography (EIT) and saturation of peripheral oxygen (SpO_2). The WELCOME solution will then integrate and exploit the monitoring data, in order to discover predictive patterns and organize the healthcare pathway. Here, it will be included signal processing and detection algorithms for cough, dyspnea and chest sounds like crackles, rhonchi or wheeze. A schematic view of WELCOME project is presented in Figure 1.1.

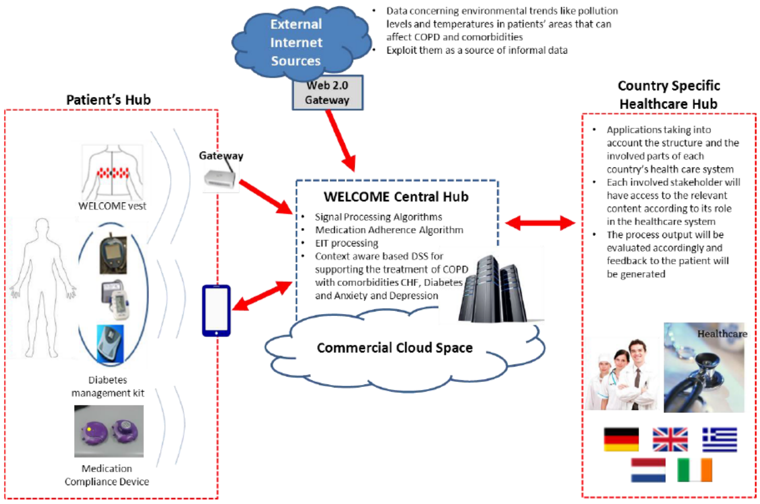

Figure 1.1: Schematic view of WELCOME. From Project Proposal Document

1.2 Motivation and Objectives

The original goal of the present master thesis was the detection of acute dyspnea by the non-intrusive parameters contemplated in the project. Dyspnea can be defined as the patient's subjective perception of shortness of breath, and the current clinical gold standard for detection and measurement is by X-ray and spirometry, both highly intrusive and, hence, not appropriated for continuous monitoring. It was intended to develop a solution for paroxysmal dyspnea and dyspnea under exertion detection, based on information provided by multi-sensor parameters able to producing surrogates for respiration and chemoreflex mechanism, as well as technology already developed by the University of Coimbra that enables cardiac output and contractility index assessment from systolic time intervals. The lack of a suitable dataset led to the acquisition of data on hospital environment, but due to difficulties and delays in this procedure, the scope of the work had to be changed. In late April of 2014 the decision to abandon the dyspnea detection study was made, and our attention turned to detection and quantification of cough.

Counting and classifying cough automatically for ambulatory monitoring has proven to be an important issue, with several challenges to address. As a common and clinic descriptive symptom of many respiratory diseases [Chung et al.,1996; Irwin et al.,1990; Irwin et al.,1998; Chang et al.,2003], cough has been vastly explored by the scientific community as a diagnosis marker. In some conditions like pulmonary fibrosis, lung cancer and COPD, the daily life quality depends on the assessment, monitoring and control of this symptom. With the potential of reducing hospital admissions and the prevention and mitigation of co-morbidities, like chronic heart failure, diabetes, anxiety and depression, a continuous monitoring of cough could contribute for improvement of financial profitability and efficiency.

The cough sound had consistently been cloven in three main phases: explosive phase, intermediate phase and voicing (or voiced) phase. The first explosive phase is characterized by an initial burst of high frequency sound that emerges in the moment of glottal opening, and so, it seems to assume importance not only for accurate onset detection of the cough sound but also for counting cough, since this phase is always present in each cough sound produced. In fact, much difficulties in counting cough rely on a variety of patterns, molded by the causing pathology, presence of sputum, among other factors. The first phase endows the signal of a more explosive and louder characteristic, therefore less variable between subjects, being an ideal candidate for identifying cough.

Besides the identification and counting of cough, three other characteristics of the cough sound proved to provide important clinical information: the pattern of coughing; the intensity of the sound; and the general acoustic properties. The analysis of those may identify the presence of sputum, wheeze or mechanical blockage, providing information about the causing mechanism of the symptom.

To sum up, the main objectives of this work were:

- Elaborate the data acquisition protocol
- Data acquisition for obtaining a relevant dataset
- Summarize the state-of-the-art methods for detecting and counting cough
- Development of a cough detection approach
- Develop and test the proposed algorithms
- Dissertation writing

1.3 Approaches

In this work, a method for automatic cough detection based on internal sounds analysis is proposed, aiming to not only count cough sounds but also to characterize the event by intensity and pattern. To this end, we use an explosive phase detection approach, analyzing 50 individuals, which produced 411 voluntary cough epochs (continuous coughing sounds without a 2-s pause), 383.4 seconds of speech, 24.55 seconds of laughing and 26 throat clear events, captured by a digital stethoscope. It was also desired in this work a broader analysis of the detected cough sound. The performance of the method to discriminate the number of cough sounds detected in a multiple cough sounds (fits of cough) was measured and it was evaluated if internal chest sound energy could be a surrogate for the intensity analysis.

The results achieved for quantification of cough showed that the explosive detection approach is a reliable method for identifying cough sounds. The detection approach shows robustness across subjects with different respiratory perturbations and demonstrates advantages in one of the main challenges of audio signal analysis, the external-to-subject noise. Moreover, internal events seem to be better characterized, with some features related with tonality, pitch, timbre and spectral analysis revealing themselves very descriptive. The ability of the approach to discriminate the number of cough sound by fits of cough indicate a first step in the analysis of the pattern, and the energy metrics exhibit potential to become a surrogate for intensity assessing.

1.4 Main Contributions

It was obtained significant dataset for a wide population with and without respiratory perturbations, recording voluntary events such as cough, speech, laugher and throat clears. The results achieved for quantification of cough showed that the explosive phase detection approach is a reliable method for identifying cough, demonstrating robustness across subjects with different respiratory perturbations and mitigating the inherent difficulty of the variety of patterns in cough sounds.

Two papers about this results were written in September 2014. "Voluntary Cough Detection by Internal Sound Analysis" was submitted and accepted at 7th International Conference on BioMedical Engineering and Informatics (BMEI 2014). "Combining Pervasive Technologies and Cloud Computing for COPD and Comorbidities Management" was submitted at 4th International Conference on Wireless Mobile Communication and Healthcare (MOBIHEALTH 2014).

1.5 Outline of the Dissertation

The master thesis document is structured into six chapters. The Chapter 2 presents the state of the art of cough pathophysiology, methods for cough identification and the guidelines for cough counting and assessing. In Chapter 3 it presented the applied methods and approaches used in this work. Chapter 4 presents the results of our work, and discussion. In Chapter 5 it is summarized the conclusions and contributions of this work.

Chapter 2

State of the Art

In this chapter we investigate the state of the art of cough pathophysiology and the former and recent methods for cough identification, as well as the scientific consensus guidelines for cough counting and assessing. Pathophysiology of cough is described in Section 2.1, it is given to the COPD a special attention, within the framework of the project, in Section 2.2, it is reviewed the specifications of the definition of cough in Section 2.3, and it is analyzed the methods and systems for cough detection in Section 2.4.

2.1 Pathophysiology of Cough

Cough is a protective reflex, a component of normal respiratory physiology that enhances the mucociliary function and clears excessive secretions and airway debris from the respiratory tract. Although cough in healthy individuals is physiologically important, it is normally a very uncommon event [Loudon et al.,1966; Sumner et al.,2013]. Mostly, it represents a symptom of a respiratory (or not) disease. Cough also assumes great importance as a factor in the spread of infections and as a patient-initiated tactic to provide cardiopulmonary resuscitation to maintain consciousness during a potentially lethal arrhythmia or convert arrhythmias to a normal rhythm.

Because cough is an easily described and recognizable physical act, patients know what is being referred to as cough, thereby lending credibility to findings from patient surveys on prevalence of cough. This had enabled the development of patient reported outcome tools, by which physicians assessed the impact of cough on patients, and still do.

Research on chronic cough has been revealed difficult over the years, because unlike bronchoconstriction, the cough reflex is blunted in anesthetized animals [Lalloo et al.,1996]. Consequently, a better understanding of the human cough reflex was delayed until human trials became possible and secure.

The pathophysiology of the cough reflex began to be better characterized by experiments with the use of inhalational challenge tests in human subjects with chemicals such as capsaicin, chloride-deficient solutions, citric acid, and prostaglandins [Lalloo et al.,1996]. Those substances with a variety of chemical characteristics can securely stimulate the cough reflex in human beings.

In neurophysiological terms, cough arises following activation of a complex sensorimotor reflex arc [Irwin et al.,2014]. The cough reflex has vagal afferent input, namely two different classes of afferent nerves - the myelinated rapidly adapting receptors, and non-myelinated C-fibers with endings in the lungs [Goldsobel et al.,2010] - and also brain stem centralization with cortical modulation and motor efferent activity involving respiratory muscles. Cough receptors are located in the respiratory tract from the larynx to the segmental bronchi [Chang et al.,1999]. Cough reflex sensitivity can be modulated either by disease or pharmacologically. Because it can be initiated at numerous anatomic sites, and it is therefore not surprising that chronic cough may have a variety of causes. Two or three different conditions may occur together in the same patient, thus complicating the clinical picture [Irwin et al.,1991; Stone,1993].

There are a variety of respiratory diseases that can be related with the symptom of cough. The most common cause of an acute cough is a viral respiratory tract infection, which can be a common cold, pneumonia, pertussis, or tuberculosis. After a viral infection has cleared, the subject may be left with a post-infectious cough. This typically is a dry, non-productive cough that produces no phlegm. Symptoms may include a tightness in the chest, and a tickle in the lungs.

When the symptoms last longer than 8 weeks, it can be designated as chronic cough, and most of the cases are due to asthma, bronchitis, post-nasal drip (excessive mucus produced by the nasal mucosa) and gastro-esophageal reflux disease [Goldsobel et al.,2010]. Asthma is a chronic inflammatory disease of the airways, normally related with allergic factors, that results from chronic inflammation of the airways which increase contractility of the surrounding smooth muscles, as presented in Figure 2.1. Its symptoms are recurring and variable [Prevention et al.,2007].

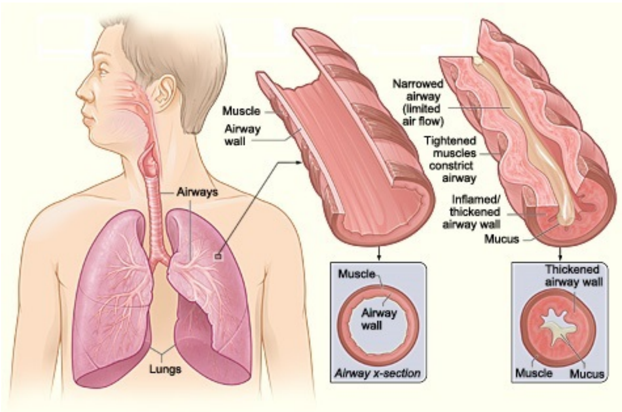

Figure 2.1: A cross-section of a normal airway and a cross-section of an airway during asthma symptoms. From [61].

Bronchitis is an inflammation of the mucous membranes of the bronchi and can be divided into acute and chronic [57]. Acute bronchitis is usually caused by viruses or bacteria, and most cases of chronic bronchitis are caused by smoking, which causes secretion of mucus into the airway, and difficulty clearing that mucus out of the airways, as Figure 2.2 shows.

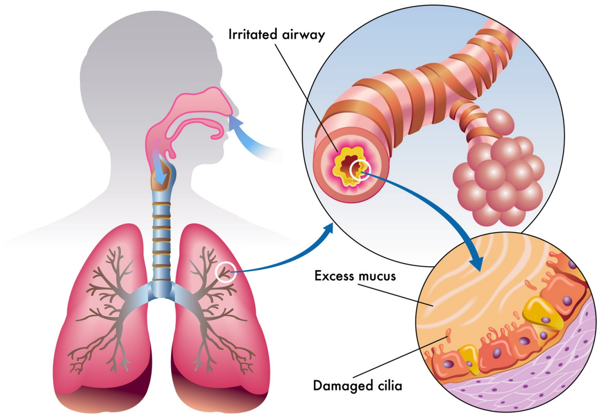

Figure 2.2: Bronchitis increases the amount of mucus in the bronchi, damaging cilia, the tiny hair-like organelles que reside on the surface of cells, and causing chronic cough. From [60].

2.2 Chronic Obstructive Pulmonary Disease

Chronic Obstructive Pulmonary Disease (COPD) is an umbrella term used to describe progressive lung diseases, most notably including emphysema and chronic bronchitis [Minkoff et al.,2005]. Emphysema is marked by progressive destruction of alveolar tissue and irreversible enlargement of the air spaces. Patients with COPD suffer symptoms of dyspnea, mucus production and chronic cough, with impairment in ability to carry out daily activities and progressive decline in quality of life.

Complex diseases such as COPD are most often the result of geneenvironment interactions that determine the clinical presentation of the disease [Agusti et al.,2012]. The diagnosis of COPD combines symptoms and a relevant exposure to risk factors as tobacco smoking and ambient pollutants, with the presence of persistent airflow limitation. For many years, the assessment of COPD, as well as the majority of respiratory diseases, has been based almost exclusively on the severity of airflow limitation. The most common of the pulmonary function tests is the spirometry, measuring the amount (volume) and/or speed (flow) of air that can be inhaled and exhaled. The Figure 2.3 presents a scheme of this exam.

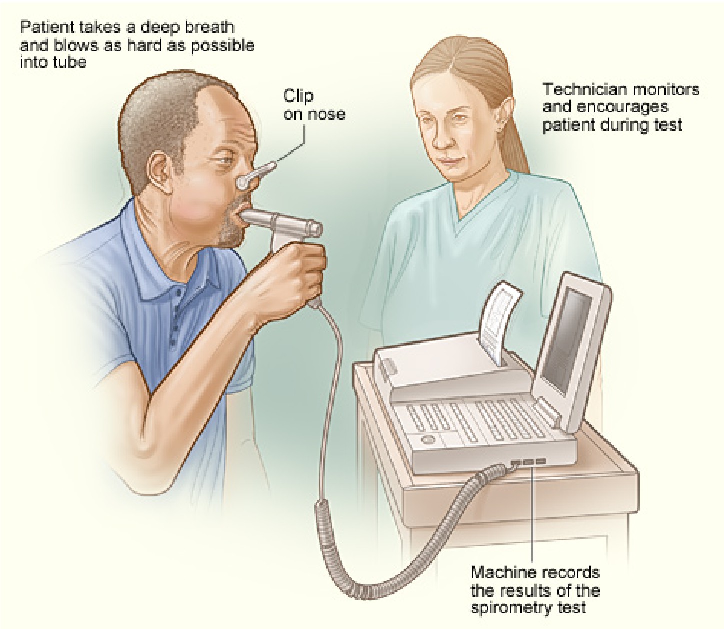

Figure 2.3: How spirometry is done. The patient takes a deep breath and blows into a tube connected to a spirometer. From [59].

Nowadays, it is known that the assessment and management of COPD patients requires a multidisciplinary approach, which should include genetic, biological, clinical and environmental levels of analysis [Agusti et al.,2012]. By the clinical point of view, COPD should be considered in any patient presenting with cough, sputum production or dyspnea, especially if the patient has been exposed to risk factors for the disease [Pauwels et al.,2004]. Cough may initially occur intermittently but it is usually the first symptom of COPD to develop [Georgopoulas et al.,1991]. COPD is also associated with several co-morbidities such as cardiovascular disease, metabolic syndrome (e.g. diabetes), osteoporosis, mental health diseases and lung cancer.

Pharmacological treatment of patients with COPD should be initiated using a short-acting or a long-acting bronchodilator. Treatment with inhaled corticosteroids is needed in patients with severe COPD. Other aspects of treatment include vaccinations, antibiotics and mucolytics. In the late-stage of severity long-term oxygen therapy, non-invasive ventilation and surgical treatment become necessary.

It was proven that chronic cough and sputum production are associated with COPD exacerbations, including severe exacerbations requiring hospitalizations, in pharmacological treated patients [Burgel et al.,2009].In the European Union, COPD severe exacerbations are the leading cause of lost work days, resulting approximately in [Loddenkemper et al.,2003]:

- 41,300 lost work days per 100,000 population, leading to productivity losses amount to a total of 28.5 billion annually.
- 4.7 billion for outpatient care.
- 2.9 billion for inpatient care.

The estimated costs of exacerbations vary widely across studies: $88 to $7,757 per exacerbation, the largest component of the total costs being typically hospitalisation [Toy et al.,2010]. Costs were highly correlated with exacerbation severity, although indirect costs have rarely been measured [Toy et al.,2010]. The important indicator is that every exacerbation event has a gradual increment phase preceeding the peak exacerbation time for several hours to several days [Rabe et al.,2007]. Therefore, an acute, objective and continuous monitoring of cough can be used for early detection of complications, and effective management of COPD may lead to improved patient outcomes and reduction in total healthcare costs for long-term management of COPD

2.3 Definition of Cough

A clear and consensual definition of cough is lacking in the majority of textbooks and scientific papers concerning cough. Two possible ways to define it are:

1. Cough is a forced expulsive maneuver, usually against a closed glottis and which is associated with a characteristic sound [Korpas et al.,1979].

2. Cough is a three-phase expulsive motor act, initialized by an inspiratory effort (inspiratory moment), followed by a forced expiratory against a closed glottis (compressive moment) and then by opening of the glottis and rapid expiratory airflow (expulsive moment) [Morice et al.,1991].

The major discrepancy between these two and between these and all other definitions lies in the several respiratory patterns associated with cough, one of its challenges. Moreover, neither these two definitions adequately deals with the common clinical scenario whereby an initial cough is followed by a series of cough efforts. For the patient, this is often described as a cough "attack". To the researcher, they may represent an extended single cough with different characteristics or peals of two or more single coughs in a short time. Clearly, this is of importance to those concerned with the accurate recording of cough frequency, therefore must be defined precisely.

For the purposes of acoustic recordings in clinical studies, cough should be defined as a forced expulsive onrush against a closed glottis that is associated with a characteristic sound, that literature have consistently cloven in three main phases associated with the mechanisms of cough-sound creation [Korpas et al.,1987; Thorpe et al.,1992]: explosive phase, intermediate phase and voiced phase.

First, the explosive phase, characterized by an initial burst of sound that emerges in the moment of glottal opening. It provides information about bronchus, inasmuch that the high frequency sound yields in the vibrations produced by the forced air flux in the airway and the bronchial narrowing places.

Then, the intermediate phase, steady-state flow with the glottis wide open. It reflects the status of trachea, the presence of sputum add a characteristic high frequency component to the sound and is directly related with the duration of this phase.

Finally, the voiced phase, where glottis narrows again, with the vocal cords approaching each other. This third phase may not take place, but the occurrence probability in voluntary cough is about 50% higher than in a spontaneous event [Hirtum et al.,2002]. In Figure 2.4 is represented a typical three-phase cough sound, acquired by a lapel microphone.

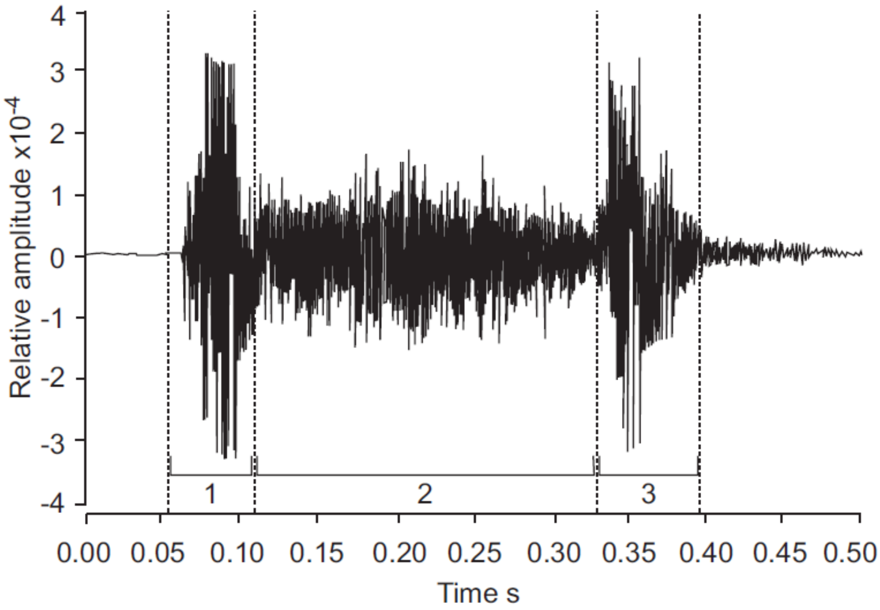

Figure 2.4: A typical three-phase cough sound (1: explosive phase; 2: intermediate phase; 3: voiced phase). From [Morice et al.,2007].

This definition clarifies the fits of cough as peals of two or more single coughs. Yet, a more careful description of cough events is needed for an accurate identification and quantification of cough, and there are several:

1. Counting the characteristic 3-phase cough sounds defined above is the most intuitive way of counting cough.

2. Nevertheless, systems that continuously monitor breathing usually quantifies cough as the number of breaths that contain at least one explosive cough sound.

3. Another cough quantification can be the time spent coughing, i.e. the number of seconds per hour containing at least one explosive cough sound.

4. It can also be done as cough epochs, continuous coughing sounds without a 2-s pause.

These four ways to counting cough can lead to differing quantification, as shown in Figure 2.5.

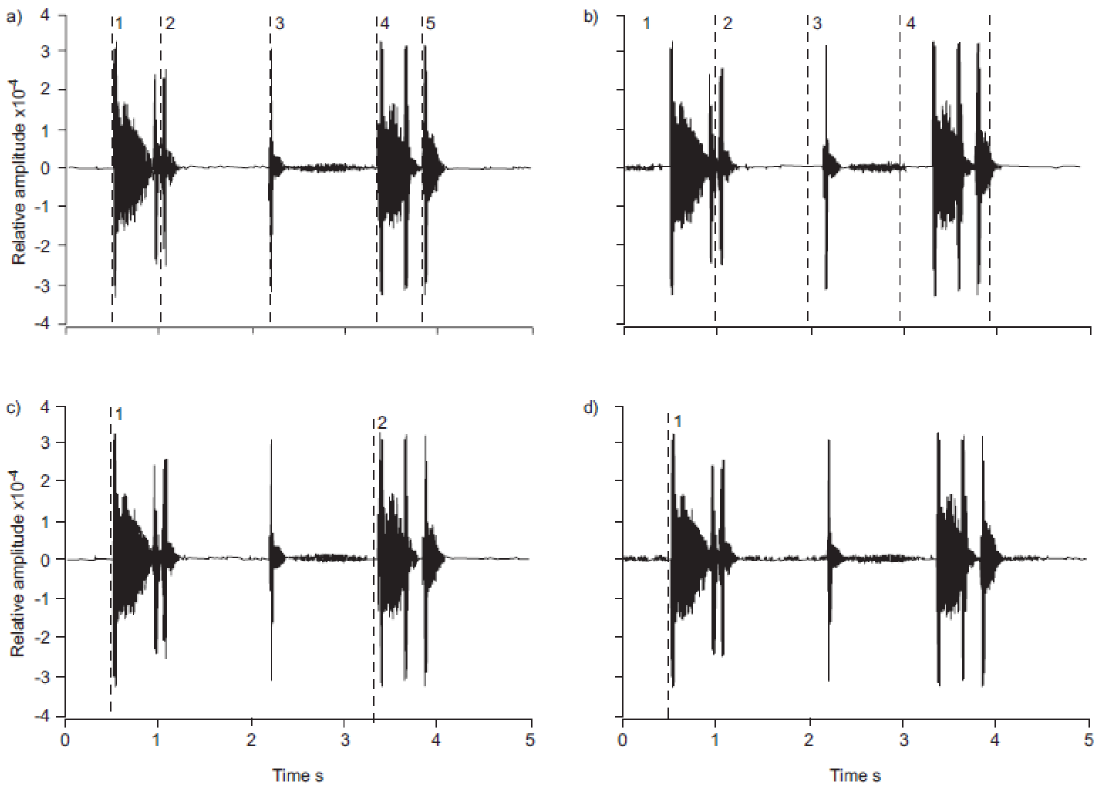

Figure 2.5: Methods for quantifying coughing: a) explosive cough sounds; b) cough seconds; c) cough breaths; d) cough epochs. Dashed lines divide units of cough and numbers represent cough count. From [Morice et al.,2007].

Counting the 3-phase cough sounds seems to be the more precise way to discriminate peal events, since every single coughs are counted, but the other three quantifications greatly simplify the process. Moreover, the European Respiratory Society (ERS) defend that there is a tight linear relationship between cough sounds and cough seconds in a variety of conditions [Morice et al.,2007], and that is not known whether any of these methods is more valid than any other in clinical terms. In [Kelsall et al.,2008] , it is found a strong correlation between counting explosive phases, cough seconds and cough epochs. What is consensual is the mandatory definition of the unit of cough used.

In this work, the definition of cough epoch is used as a time interval that contains cough sounds spaced no more than 2 seconds [Hamutcu et al.,2002; Munyard et al.,1996]. This defines peal events, and even nearby single coughs, as a single cough epoch. In physiological terms, since coughing serves the purpose of unblocking the airways, nearby 3-phase coughs aim to solve the same block, so this definition indirectly counts the occurrence of discomforts that lead to cough happenings. Moreover, excellent inter and intra-subject agreement has been found for quantification of cough seconds, explosive phases and cough epochs [Hamutcu et al.,2002; Munyard et al.,1994].

However, in clinical terms it is relevant the way that body deals with the blocks, a long peal event with numerous coughs contrasts with few strong single coughs although both can resolve the same. Therefore, other features of the cough signal apart from the number of coughs are potentially of use as clinical end-points for classifying the event. ERS defined three characteristics of the cough sound which provide important information [Morice et al.,2007]:

- The pattern of coughing is important, since either single coughs or fits of coughing serve different mechanical purposes and affect the patient's experience.

- The intensity of the sound is also relevant, which could be given by both peak intensity and overall energy released, is important in subjects that despite a small number of coughs may still find the symptom very distressing if associated with chest pains, retching or syncope.

- Finally, the acoustic properties of the cough sounds, which may identify the presence of sputum and wheeze.

2.4 Automate Counting of Cough

The evaluation of cough severity was for many years subjective, based on cough scores, diaries, visual analogue scales, and symptom questionnaires, which are completed either by the patient himself or a parent [Birring et al.,2003; Marsden et al.,2008]. However, it has been shown that subjective assessments correlate modestly with objective measures of cough frequency [Decalmer et al.,2007].

An objective measure of cough would be of use in clinical practice, clinical research and the assessment of novel therapies and pharmaceuticals. It would permit validation of the presence of cough, grading of severity and monitoring of responses to therapeutic trials. But identifying, quantifying and classifying cough has proven to be difficult, not only because of cough variety of phases, patterns and adjustments to pathologies with chronic cough, but also because the aim is to perform ambulatory long term monitoring, which, manually, can become a time-consuming and arduous task.

There have been attempts to achieve a consensual and reliable system for monitoring cough, with experiments based on both voluntary and pathologic events. Some approaches from the past used audio signals either alone or combined with others [Munyard et al.,1994; Chang et al.,1997; Hsu et al.,1994], but they only enabled to manually spot the cough sounds by visualizing the signals, which does not avoid the loss of time in counting.

Therefore, the ideal cough monitoring system needs to be able to detect and count cough automatically and with high accuracy. One way of classifying cough monitoring devices is by the degree of user input required [Smith,2008]. The ideal cough monitoring system would be small, robust and as less intrusive as possible for the subject.

Recently, other research works tried to automate the recognition and counting of cough sounds. Many methods use ambient audio signal only. The use of Hidden Markov Models (HMM), for instance, to detect cough signals as keyword spotting in continuous ambient audio recordings, exhibit promising results [Matos et al.,2007]. The Leicester Cough Monitor (LCM) use this approach to presegment possible cough events from 24-h ambulatory ambient audio recordings [Birring et al.,2008]. Some of these possible cough segments are then presented to a human expert in order to develop a statistical model tailored to the current recording. Finally, the full recordings are processed with the developed models. In [Birring et al.,2008], the system achieved an overall recall and specificity of 91% and 99%, respectively, for tests in 6-h recordings from 9 respiratory patients. A scheme of this system can be found in Figure 2.6.

Matos et al.[Matos et al.,2006] had also previously used HMM trained on ambient audio features, developed to characterize cough events, but also to represent the set of all other possible events (it also selects the event candidates from recordings, by energy thresholding). This two models compete to score new recordings and the most likely sequence of coughs and fillers is retained. This system achieved a recall of 71%, lower than the LCM, but here the process is fully automatic.

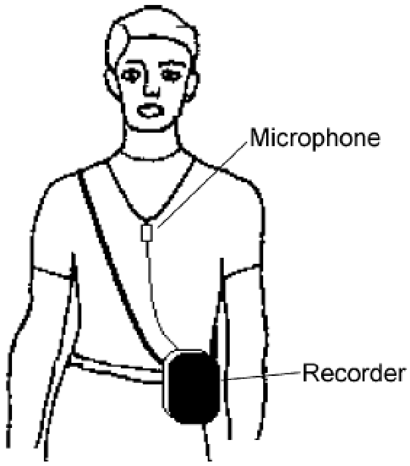

Figure 2.6: Leicester cough recording system scheme, with the lapel microphone for ambient sound recording. From [Matos et al.,2007].

Currently, this is one of the most promising approaches in cough counting, but some authors deem that coughs should not be treated as speech, since their acoustic differing characteristics need to be fully considered in the design of algorithms [Chunmei et al.,2013].

The Hull Automatic Cough Counter (HACC) system uses digital signal processing to calculate characteristic spectral coefficients of ambient sound events, which are then classified into cough and non-cough sounds by the use of a probabilistic neural network (PNN)[Barry et al.,2006]. It uses an event detection logic based on adaptable thresholding, which basically removes the predominant silence phases and allows focusing in probable cough sounds. This technique reduces the computation time of the analysis, by cutting a large percentage of data to be analyzed. The spectral coefficients are calculated for detected event candidates, which are then classified into cough and non-cough events by the use of PNNs. As the HACC system just identifies coughs and labels them, and does not automatically count them, a technician has to listen to and count the labelled coughs using a graphical user interface, presented in Figure 2.7. Tests performed only in smoking subjects achieved a specificity of 96% and a recall of 80%.

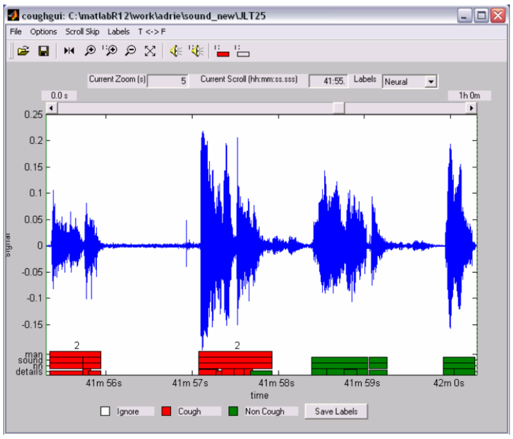

Figure 2.7: HACC system graphical user interface. From [Barry et al.,2006].

Drugman et al. investigated the use of contact microphone signal in complement to the ambient audio signal, with the use of PNNs too [Drugman et al.,2012]. The key idea was to focus only on the detection of the explosive phase of cough. Indeed, the intermediate phase had proven to be very similar to a forced expiration [Korpas et al.,1996], or in some healthy sputum-free subjects cases to a silence phase. As the voiced phase may not occur and resembles to a speech sound, the explosive phase can assume primacy in the cough sound analysis. As we can see in Figure 2.8, although this phase demonstrates irregularity, it is characteristic of the beginning of any cough sound and possibly, its irregularity can be more nonspecific among individuals.

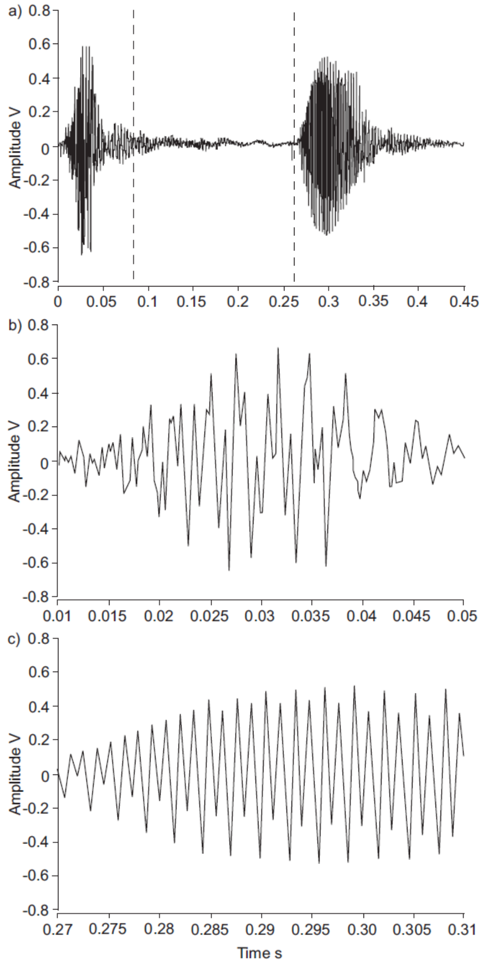

Figure 2.8: a) Typical cough ambient sound waveform divided into the three acoustic phases; b) The explosive phase on an expanded timescale, demonstrating the irregular, noise-like appearance; c) The voicing phase on an expanded timescale, showing its contrasting regular, periodic appearance. From [Tracey et al.,2008]

Drugman et al. approached the explosive phases of each cough sound by the first 60 milliseconds of the total sound, and selected a set of 50 features from the total 222 features calculated. This approach was experimented in voluntary cough from healthy subjects, achieving specificity and recall of 88%, for ambient audio signal analysis, and specificity and recall of 71% for the signal of a contact microphone over trachea and over thorax. The combination of those two signals was proved to convey little new relevant information compared to the audio signal modality alone.

There are, to date, three major cough-counting industrial devices: the Lifeshirt System, the PulmoTrack-CC system and the VitaloJAK system. Industrial devices as they are, little information was found about methods used in each. The LifeShirt®(VivoMetrics, Inc., Ventura, California, United States of America) system, incorporates respiratory inductance plethysmography for the non-invasive measurement of volume and timing ventilatory variables and also incorporates a unidirectional contact microphone, a single channel ECG, and a centrally located, 3-axis accelerometer. In [Coyle et al.,2005], the system was evaluated in eight patients with a documented history of COPD, 24-h ambulatory sound recording, and with a specialized software (VivoLogic®, VivoMetrics, Inc., Ventura, California, United States of America) used to view the data and a proprietary algorithm housed within the software to identify cough. A recall of 78.1% and a precision of 84.6% was achieved. The system sensors are presented in Figure 2.9.

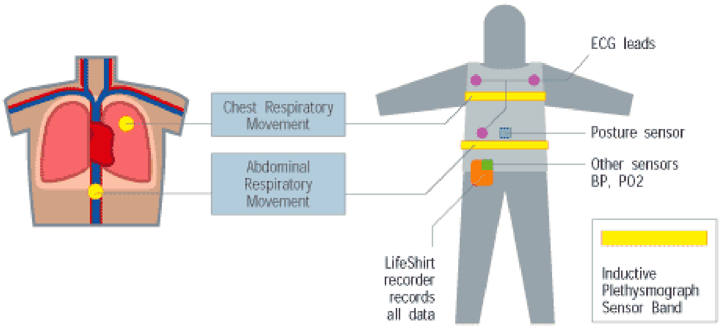

Figure 2.9: The LS system sensors. The inductive plethysmography sensors capture chest and abdominal respiratory movements. With ECG sensors, a pulse oximeter, and a posture accelerometer, all data are recorded in a small device attached to the waist. From [58].

The Karmelsonix®(KarmelSonix Limited, Baulkham Hills, New South Wales, Australia) company launched the PulmoTrack-CC, which includes a piezoelectric belt, one lapel microphone and two contact microphones placed on the trachea and the thorax. In [Vizel et al.,2010], the algorithm specifications are not clearly explained, but there is a first detection of cough candidates, and then a validation phase by detection of specific characteristics of cough in all signals data. The performance of this device reached a recall of 96% and a specificity of 94% on voluntary cough from 12 volunteers.

The VitaloJAK system uses a contact microphone placed on the chest wall and a custom-made digital recording device to detect cough from sound. In [McGuinness et al.,2007a], this system was adapted to a physiological approach tailoring. Subjects perform voluntary coughs, which are recorded, from set lung volumes. The same author concludes in [McGuinness et al.,2007b] that much of the variability in cough sounds within an individual can be explained by the lung volume from which the cough occurs. Acoustic parameters extracted from these voluntary coughs can be then used to interrogate a 24-h sound recording and pick out candidate events. The study, in 10 subjects (n=5 chronic cough, n=5 asthma), reaches a recall higher than 99% in this first selection, while compressing the amount of data to check manually in about 65%. The final cough detection achieved a recall of 97.5% and a specificity of 97.7%. In Figure 2.10 we can see an image of the VitaloJAK cough monitor.

Figure 2.10: VitaloJAK cough monitoring device, with both lapel and contact microphone attached. From [Smith,2008].

Despite all of those systems and approaches, the ERS Committee defends that there are at the moment no standardized methods for recording cough. Moreover, there are no adequately validated, commercially available, and clinically acceptable cough monitors [Morice et al.,2007].

Chapter 3

Methods

In this chapter we present the methods and approaches used in this work. In the Experimental Setup Section 3.1 it is explained the acquisition protocol for obtained dataset. The annotation phase is descripted in Annotation Algorithm Section 3.2. In the Algorithm Design Section 3.3 we present the specifications of the proposed algorithms. The feature extraction proceedings are described in Feature Extraction sub-Section 3.3.1, as well as feature selection proceedings are described in Feature Selection sub-Section 3.3.2. The Classification sub-Section 3.3.3 presents the classification algorithm and the Post-processing and Performance Analysis sub-Section 3.3.4 presents the proceedings after classification for cough counting and performance analysis.

3.1 Experimental setup

In order to evaluate the explosive phase approach for cough detection in chest sounds, a dataset was required. To the best of our knowledge, no free-access datasets were available for this kind of purpose. It was desired to have not only cough sounds, but also other respiratory and prosody-related occurrences, which can possibly be confounded with cough. Related works that use recordings of voluntary cough often include in their protocols events such as speech, laugh, throat clearings and forced expirations [Drugman et al.,2012; Drugman et al.,2013], and sneezes are also present in ambulatory recordings [Matos et al.,2007].

The employed recording system consisted in a 3M Littmann digital stethoscope, model 3200, St. Paul, Minnesota, USA, with a sampling frequency of 4000 Hz. The sound signal was acquired with individuals seated in a room and without any external sound cancellation. For the auscultation site, it was considered the posterior inferior lobe site of the left or right lung, and the posterior middle lobe site of the left or right lung. Related works have considered auscultations in the throat, trachea and thorax [Drugman et al.,2012; Drugman et al.,2013].

For each voluntary subject, 4 recordings of 15-s duration each were performed, and subjects were asked to produce in each recording: a single cough, a peal of two or more coughs, and around 5 seconds of one of the following events: speech, laughter and throat clears. In Figure 3.1 we can find the appearance of speech, laughter and throat clears in amplitude and their spectrogram. The subject initiated each event commanded by feedback of the acquisition technician, always keeping more than 2-seconds spacing between each event. The order of the events in the recording was also randomly varied for each of the four recordings. It was also requested that the subject perform the events with a minimum of breaks, i.e., to speak as much continuously as possible, to facilitate annotation.

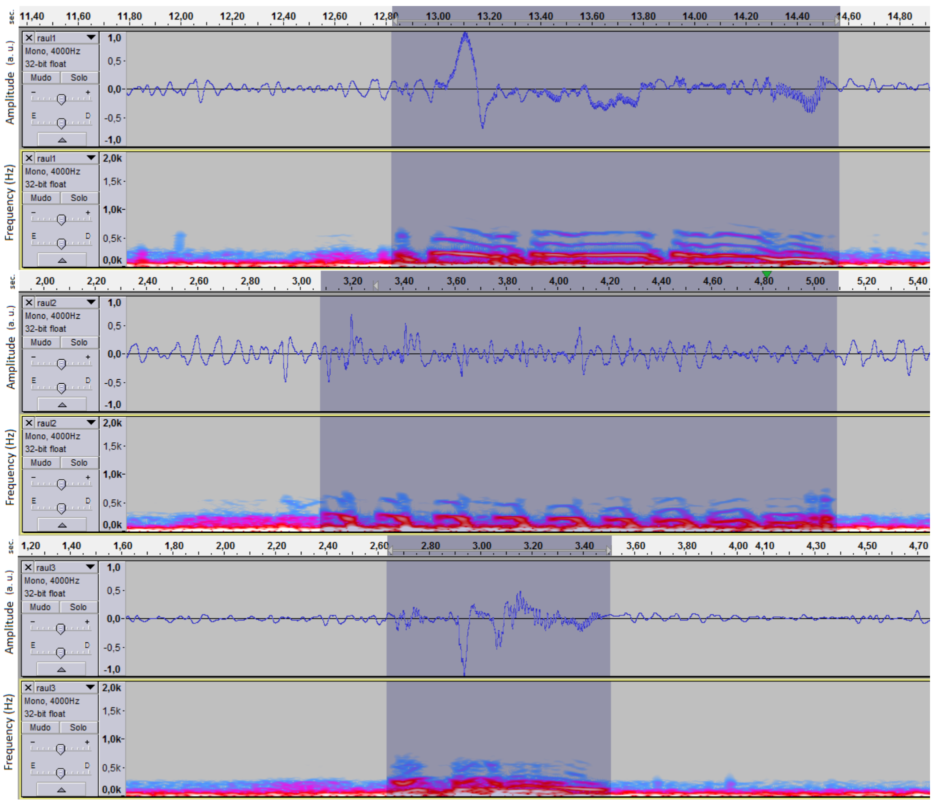

Figure 3.1: Examples of confusing events present in the acquisition protocol in amplitude and spectrogram: from top to bottom: first, a speech event; second, a laughter event; third, three throat clears

Recordings were performed on 36 healthy individuals without any known respiratory perturbation (without respiratory perturbations group - WPG) and from 14 individuals with respiratory perturbations (respiratory perturbations group - RPG). In the RPG group, smokers can be found (n=1), as well as cough-related pathologies like asthma (n=3), bronchitis (n=1), rhinitis (n=1) and simple colds (n=8).

The biometric characteristics of the testing groups are shown in Table 3.1. All the data related with the population can be found in the Attachments Section (ref).

	WPG+RPG	**WPG**	**RPG**
Age (years±STD)	33.26±14.73	36.21±23.25	29.89±9.98
BMI - Body mass index (kg/m^2±STD)	24.85±4.17	24.84±5.28	24.86±3.74
Gender (males/ females)	26/24	7/7	18/18

Table 3.1: Biometric characteristics of the testing groups

The final total audio data consisted of 50 recordings of 1 minute acquired from 50 individuals, containing 411 cough epochs (single cough or peal events), 383.4 seconds of speech, 24.55 seconds of laughter and 26 throat clears.

3.2 Annotation Process

All the audio data was analyzed by an observer, using Audacity audio software in order to obtain the annotation of explosive phases of cough sounds. Each audio file was loaded into Audacity and, by listening to the audio, the observer detected the cough sound. Also by observing the signal's spectrogram, the onset and offset were finely adjusted. In Figure 3.2 we can find the appearance of a cough sound in amplitude and their spectrogram.

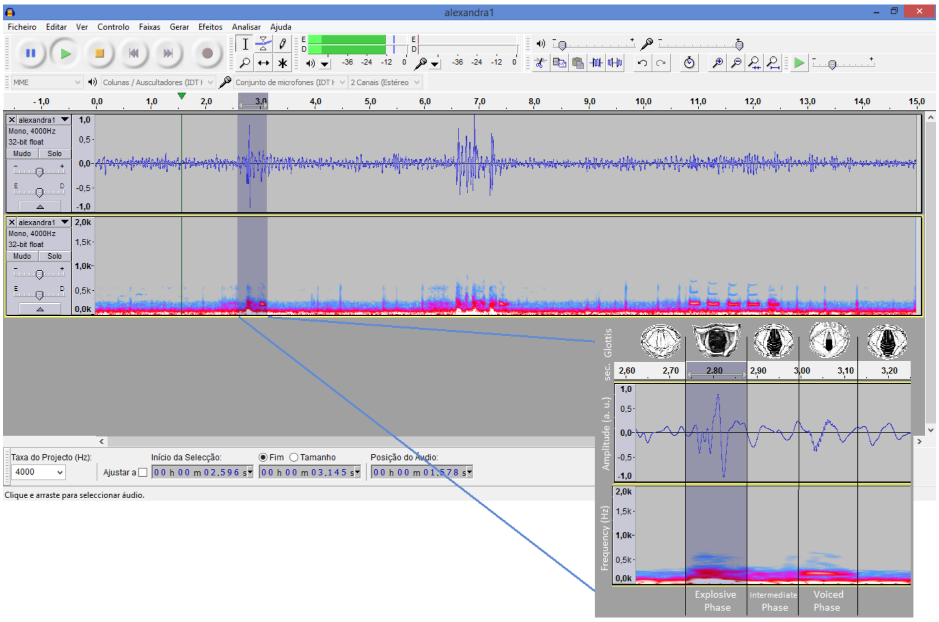

Figure 3.2: Audacity software interface with a signal loaded: top, waveform; bottom, spectrogram. A single cough sound is highlighted, evidencing the three phases and relating them with the status of the glottis.

The annotated onsets and offsets for each explosive phase of cough sounds were used to obtain the annotation vector of each recording, which consists in a vector containing the annotation of each frame-interval chosen a priori. In a later stage, the final voicing phases of each cough epoch were also annotated by listening to the audio and observing the spectrogram of the signal. If they did not occur, the offset of the final intermediate phases were annotated. In Figure 3.3 we can find one annotated voicing phase and offset of intermediate phase. The confusing events were also annotated by the same method.

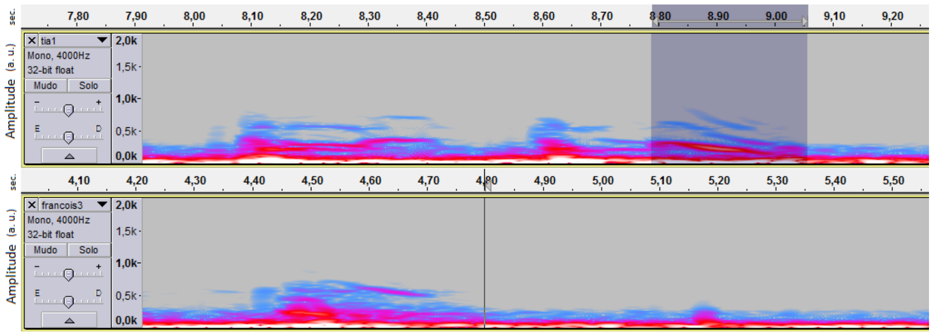

Figure 3.3: The annotation of voicing phases and offset of intermediate phases. In the top signal, the final voicing phase of a peal event with 2 cough sounds is highlighted. In the second signal, the offset of the intermediate phase of a single cough sound is pointed out.

3.3 Algorithm Design

3.3.1 Feature Extraction

A total of 79 features were calculated for each frame interval of the analyzed recordings. These features were extracted in 50-milliseconds frames, without overlapping. Most of these features were computed using the open-source MIR toolbox [Lartillot et al.] and VOICEBOX [Brookes et al.,2002] for Matlab, covering a broad range of sound dimensions including frequency, timbre, pitch, tonality and also speech-related analysis. All the 79 calculated features are presented in Table 3.2.

Feature	Description	Dimension	Functions (Toolboxes)
Mean_ener	Mean of the squared data signal	Energy	mean (Matlab)
Peak	Largest value of data signal	Basic Operator	max (Matlab)
Fft	Mean of the decomposition of the energy of the data signal along 128 frequencies using a Fast Fourier Transform	Frequency	mirspectrum + mirgetdata + mirstat (MIR toolbox)
Evo	Distance between the Fft of each successive frames	Frequency	mirflux + mirgetdata (MIR toolbox)
Ter	Modulates the energy by an attenuation in the lower and higher registers of the spectrum, and an emphasis around 25 KHz, where much of the speech information is carried	Frequency	mirspectrum + mirgetdata + mirstat (MIR toolbox)
Bark	Convert the Fft value in Hertz to the Bark frequency scale	Frequency	frq2bark (VOICEBOX)
Cent	Convert the Fft value in Hertz to cents scale	Frequency	frq2cent (VOICEBOX)
Erb	Convert the Fft value in Hertz to erb rate scale	Frequency	frq2erb (VOICEBOX)
Mel	Convert the Fft value in Hertz to mel scale	Frequency	frq2mel (VOICEBOX)
Rhar	Mean of the Hartley transform of data signal	Frequency	rhartley (VOICEBOX) + mirstat (MIR toolbox)
Rdct	Mean of the Discrete cosine transform of data signal	Frequency	rdct (VOICEBOX) + mirstat (MIR toolbox)
Zoomfft	Mean of the Discrete Fourier transform evaluated over a linear frequency range	Frequency	zoomfft (VOICEBOX) + mirstat (MIR toolbox)
Rsfft	Mean of the Fast Fourier Transform of real symmetric data	Frequency	rsfft (VOICEBOX) + mirstat (MIR toolbox)
Vu	Calculate volume unit level of data signal in linear units rather than dB	Frequency	v_ppmvu (VOICEBOX) + mirstat (MIR toolbox)
Zerocross	Calculate the number of times the data signal crosses the X-axis	Timbre	mirzerocross + mirgetdata (MIR toolbox)

Rolloff(1:2)	Calculate the frequency such that 85% (Rolloff(1)) and 95% (Rolloff(2)) of the total energy is contained below that frequency	Timbre	mirrolloff + mirgetdata (MIR toolbox)
Brightness	Fix the cut-off frequency of 1500 Hz and calculate the amount of energy above that frequency	Timbre	mirbrightness + mirgetdata (MIR toolbox)
Centroid	Calculate the spectral distribution centroid	Timbre	mircentroid + mirgetdata (MIR toolbox)
Spread	Calculate the spectral distribution spread	Timbre	mirspread + mirgetdata (MIR toolbox)
Skewness	Calculate the spectral distribution skewness	Timbre	mirskewness + mirgetdata (MIR toolbox)
Kurtosis	Calculate the spectral distribution kurtosis	Timbre	mirkurtosis + mirgetdata (MIR toolbox)
Flatness	Calculate the spectral distribution flatness	Timbre	mirflatness + mirgetdata (MIR toolbox)
Entropy	Calculate the spectral distribution entropy	Timbre	mirentropy + mirgetdata (MIR toolbox)
Regularity	Calculate the mean of the variation of the successive peaks of the spectrum	Timbre	mirregularity + mirgetdata (MIR toolbox)
Mfcc(1:14)	Calculate the mel-frequency cepstral coefficients of 13 ranks, plus the coefficient related to the average energy, that is by convention of rank 0 (Mfcc(1))	Timbre	mirmfcc + mirgetdata (MIR toolbox)
Mfccd(1:14)	Calculate the first derivative of mel-frequency cepstral coefficients of 13 ranks, plus the coefficient related to the first derivative of the average energy, that is by convention of rank 0 (Mfccd(1))	Timbre	mirmfcc + mirgetdata (MIR toolbox)
Mfccdd(1:14)	Calculate the second derivative of mel-frequency cepstral coefficients of 13 ranks, plus the coefficient related to the second derivative of the average energy, that is by convention of rank 0 (Mfccdd(1))	Timbre	mirmfcc + mirgetdata (MIR toolbox)

Roughness	Calculate the estimation of the sensory dissonance, or roughness, related to the beating phenomenon whenever pair of sinusoids are closed in frequency	Timbre	mirroughness + mirgetdata (MIR toolbox)
Midi	Convert the Fft value in Hertz to midi scale of semitones	Pitch	frq2midi (VOICEBOX)
Pitch(1:2,1:2)	Calculate the discretized note events of the signal data, for no filterbank configuration and Gammatone filterbank configuration (Pitch(1:2,:)) and for each of this, calculate the mean of the 2 best pitches (Pitch(:,1:2))	Pitch	mirpitch + mirgetdata (MIR toolbox)
Inharmonicity	Calculate the amount of partials of the signal data that are not multiples of the fundamental frequency	Pitch	mirinharmonicity + mirgetdata (MIR toolbox)
Key	Calculate an estimation of tonal center positions and their respective clarity of the signal data	Tonality	mirkey + mirgetdata (MIR toolbox)
Mode	Calculate an estimation of the modality of the signal data	Tonality	mirmode + mirgetdata (MIR toolbox)
Noisem	Calculate an estimation of the noise spectrum from noisy speech using minimum mean-square error method	Speech-related analysis	estnoisem (VOICEBOX)
Noiseg	Calculate an estimation of the noise spectrum from noisy speech using minimum statistics	Speech-related analysis	estnoiseg (VOICEBOX)
Teager	Calculate the mean of the Teager energy of the signal data	Speech-related analysis	teager (VOICEBOX)

Table 3.2: Description of the calculated features and the Matlab functions used for each one.

3.3.2 Feature Selection

By merging the total feature matrix with the annotation vector, the final Feature Matrix (FM matrix) is obtained. This FM will feed the classifier and will also be used for feature selection. This FM matrix was exported as an *.arff* format file for posterior loading in the WEKA data-mining software.

The feature selection was based on the Relief [Robnik-Sikonja et al.,2003] algorithm, which outputs a weight for each feature, based on which the feature ranking is determined.

3.3.3 Classification

After obtaining the FM matrix and the features ranking, the classification phase, based on the training-testing approach, was conducted to discriminate between sound frames with and without cough.

The chosen classifier was the K-Nearest Neighbors (KNN) pattern recognition algorithm. KNN is a simple and non-parametric approach that is known to be a good choice when data distribution is unknown or difficult to determine. The algorithm determines the class of a given pattern based on a distance metrics (defined here as Euclidean) and on the class of surrounding neighbor patterns. More precisely, the algorithm finds the closest k neighbors by using the chosen distance metrics and the class of a given pattern will be the class of the majority of its neighbors. In this way, the unique training parameter is the number of neighbors (k).

Classification results were validated with repeated stratified 10-fold cross validation (20 repetitions). Therefore, for the total analysis of the 50 subject's data, training was carried out using collected data from 45 subjects, while testing was performed with data from the remaining 5 subjects. In each fold, the 5 tested subjects were changed.

Before classification, the assembled training and testing data were preprocessed, and the set of selected features was rearranged by Principal Component Analysis, which uses an orthogonal transformation to convert the set with possibly correlated features into a set of values linearly uncorrelated called principal components. The reconstruction was performed using a 90% value of the covariance.

3.3.4 Post-processing and Performance Analysis

The results obtained in the classification stage were then post-processed by merging events spaced by less than 2 seconds, according to our definition of cough epoch. For the explosive phases classification, the accuracy, recall and precision metrics were obtained frame by frame by comparing the annotated vector of explosive phases for the testing data and the resulting classification vector, obtaining then the true positive, false positive and false negative rates. The recall and precision metrics, obtained for each explosive phase, were also obtained, defining a true detection as a classified event present in some part of an annotated one. The recall, also known as sensitivity, is the ratio of the number of true positive events to the total number of positive events. The precision, also known as positive predictive value, is the ratio of the number of true positive events to the sum of the number of true positive events with false positive events.

As for the evaluation of approximate cough epochs detection accuracy, the same procedure was applied, by merging events spaced less than 2 seconds. To those resulting events, just the intermediate and voicing phases of the last cough sound of each cough epoch were absent, comparing with true cough epochs. This resulting vector was compared with the post-processed classification vector, obtaining the true positive, false positive and false negative rates for the detected approximate cough epochs.

The mean and standard deviation (STD) of the onset and offset flaw of the successful detected events were computed as well. It was also measured how many times a classified event encompassed two, three or four annotated events, recording the mean and STD of the time lapse between those encompassed events.

Using only the annotation of the last voicing phases of each cough epoch, the recall and precision metrics were calculated by this second approximation of the cough epoch. To these resulting events, just the intermediate phases of the last cough sound of each cough epoch were absent, when the voicing phase was absent. Finally, to this last approximate events, the offset of the last intermediate phases of each cough epoch was also included, resulting in the real cough epochs recorded in the signal. The recall and precision metrics were also obtained.

At last, the recall and precision metrics for the explosive phases detected on peals of cough were obtained. For each peal, the true number of explosive phases was compared with the number of single hits detected by classification on the event interval. Moreover, the mean of the energy signal, was calculated in the detected explosive phases and approximate cough epochs. Also the mean of the maximum energy in each detected event interval was calculated.

The complete software for feature extraction and learning was run under Windows 8.1 on a 2.9 GHz I7 3520M PC with 8 GB of RAM, using Matlab R2012a 64-bits, Weka v3.6 and Audacity 2.0.5.

Chapter 4

Results and Discussion

In this chapter it is presented the results of our work, discussing their relevance, importance and comparison with other similar methods.

All data acquisition was accomplished without external noise cancelation to prove the impact of this in the chest sounds signal. The sampling frequency of the recording system used was appropriated given that Chunmei et al.[Chunmei et al.,2013] located the frequencies of cough with and without sputum below 2000 Hz, our Nyquist frequency range top.

The auscultation sites considered resulted from the predicted auscultation sites for the WELCOME project vest, excluding the anterior sites, more suitable for heart sounds auscultation. By consulting a pulmonologist physician from the research project group, it was considered the lower basal sites as the best for pulmonary auscultation. Between the right and left side it was not found any reference and it was chosen the right side to minimize the auscultation of heart beats.

The preliminary 15-seconds acquisitions revealed to be insufficient, and we preferred to repeat this short-time acquisition four times for each subject, rather than extend the time interval, to facilitate the script of the subjects. The commanded start of the events by feedback of the acquisition technician assured the spacing of more than 2 seconds between events. For the confusing events tested, sneezing was discarded by the inability of reproducing voluntarily, and forced expirations were found to be present in the foregoing part of the cough events recorded, so it would test the ruggedness of our approach by the deviation of the onset of detected events. Speech events were more requested to the subjects, resulting in an increased prevalence of those relative to others, because of the ethic importance of discarding sound segments containing speech. Furthermore, we advocate that noisy and powerful throat clears, which can be confounded by the classifier as a cough sound, are not so bad to be counted, as an important respiratory event too.

To minimize the inherent errors of parallax on the adjustments in the annotation phase, the observer tried to obtain values scaled visually with precision of around 0.05 seconds, and it was considered that a frame was part of a given event if more than half of the frame belonged to the noted event.

The established frame length of 50 milliseconds is in the order of the standard magnitude for microphone audio processing and the frame lengths of other similar studies. The results with no overlapping showed a good performance, but still, it was roughly tried the use of frames with overlap, achieving much longer computations and lower results. It may be a consequence of the already low performances frame by frame, which with the frame overlap should lower even more. Thus, it is was not followed up this study.

Our definition of cough event as an epoch that contains cough sounds spaced no more than 2 seconds, as valid as every other aforementioned, results in 411 cough events counted in the final total audio data. In those cough events, we found a total of 896 explosive phases, and consequently, the same number of 3-phase cough sounds. The merging of the annotated and the detected explosive phases obtained an approximation of cough events, since the final intermediate and voicing phases of each event are not accounted for in the result of that merging process. The assured condition of no spacing of less than 2 seconds between events confirms that no pair of fits of cough was merged after this procedure, and it was confirmed that no fit of cough of our data have a spacing of more than 2 seconds between subsequent explosive phases, ensuring that no fit of cough was splitted after this procedure.

KNN algorithm was chosen due to its simplicity and lower computation, and proved to be efficient for the desired classification. This allows its applicability in continuous monitoring systems, with limited battery.

We performed tests in the whole dataset (WPG+RPG), and for the WPG and RPG groups alone. For all of the testing groups, the Relief algorithm ranking is shown in Table 4.1.

For all of the testing groups, the most relevant feature was the key, obtained by the mirkey function of MIRtoolbox, which relates with tonality and gives a broad estimation of tonal center positions and their respective

	WPG+RPG	**WPG**	**RPG**
1st	Key	Key	Key
2nd	Evo	Evo	Evo
3rd	PitchNoFilterbank	Mfcc4	Mfcc4
4th	Midi	Midi	Midi
5th	Cent	Cent	Cent
6th	Mfcc4	PitchNoFilterbank	PitchNoFilterbank
7th	Rolloff95	Mfcc1	Mfcc1
8th	Mfcc1	Mfcc0	Mfcc0
9th	Mfcc0	Mfcc11	Mfcc11
10th	Mode	Rolloff95	Rolloff95

Table 4.1: Results obtained for Feature Selection. Rakings up to the 10 best features

clarity. Minimally, key consists of tonic plus the mode. Mode represents an estimation of the modality of the signal, i.e., if the signal frame corresponds to a major or a minor scale. Modes and scales may or may not have a tonic, e.g., the chromatic scale has no tonic, and the C major music scale has the tonic C. Mode feature, also present in the ranking, calculates an estimation of the modality of the signal, and also relates with tonality. This Key feature proved to be very relevant in our explosive phase detection approach. It shows that explosive phases of cough have a tonal center distinct from the tonal center of all other sounds tested, i.e., the tonic elements (which tend to assert their dominance over all others) in explosive phases seem to be distinctive.

For the remaining selected features, it can be found a maintenance of the occurrences in the three testing groups. This reveals the relevance of these features for the problem, and also the proximity between the groups. Evo is the second best feature for all groups, and gives an estimate distance between the Fast Fourier Transform (FFT) of the signal in each successive frames. This means that there is a leap in the FFT in the beginning of the explosive phase frames, which is expected and already verified in the annotation. Mfcc4 corresponds to the value of the fourth Mel Frequency Cepstral Coefficient (MFCC), as Mfcc1 is the first coefficient and Mfcc11 is the eleventh coefficient. The Mfcc0 corresponds to the coefficient related to the average energy. MFCC's are features widely used in automatic speech recognition. The selection of those features reveal therefore the importance of the spectral shape of the sound for identifying explosive phases. Midi is the mean by frame of the conversion of the FFT of the signal to midi scale of semitones. Cent is the mean by frame of the conversion of the FFT of the signal in Hertz to the logarithmic cents scale. Like a decibel's relation to intensity, a cent is a ratio between two close frequencies. Those scale changes give to the features an ability to detect variations in the FFT of the signal. PitchNoFilterbank calculate the mean of the best pitch of the discretized note events of the signal data, for no filterbank configuration. These two features demonstrates the importance of the pitch in explosive phase detection. Pitch is a perceptual property that allows the ordering of sounds on a frequency-related scale [Klapuri et al.,2006]. It depends on the frequency and is used in music to describe the extent at which a note is high or low. The explosive phase can therefore be interpreted as a high note event. Rolloff95 calculates the roll-off frequency, the frequency such that 95% of the total energy is contained below that value. This metric relates with timbre, which gives an idea of the quality of a sound. It is expected to obtain a higher roll-off frequency for the explosive phase frames, which are characterized by a high frequency content. PitchNoFilterbank is related with pitch analysis and calculate the mean of the best pitch of the discretized note events of the signal data, for no filterbank configuration. A filterbank is filter shape inspired by the auditory system.

The whole set of best features are related with all sound dimensions analyzed, not being found a pattern that allows to say that one is more relevant than another. This reveals flexibility in characterization of our sound signal, with several areas of sound analysis contributing for the identification of the explosive phase.

We proceeded to the fixation of best k value for the KNN and the optimal number of features, in order to maximize results, for all of the testing groups. The best k value was selected by considering, iteratively, odd numbers in the range of 3 to 15, for the set of the first 20 ranked features, representing 25% of the whole set of 79 features. This procedure was performed with 5 cycles of repeated stratified 10-fold cross validation method, with the final result obtained by the mean of the set of results. The fixation of the k value was done by maximization of the result of the F1 score, a measure of accuracy that considers both the recall and precision metrics, event by event.

After setting the best k, the optimal number of features was determined experimentally by using a simple forward feature selection approach that consists on adding one feature at a time based on the resulting ranking by the Relief algorithm. This procedure was also performed with 5 cycles of repeated stratified 10-fold cross validation method, obtaining the best four numbers of features by set. For those, a 20 cycle method was performed, to determine the best final number by maximization of the result of the F1 score event by event. One exception occurred in RPG testing group, where bigger values were achieved for the set length of 20 features. Here, it was tested set lengths until 30 features with 5 cycles of repeated stratified 10-fold cross validation, performing then a 20 cycles test for the best four numbers of features, confirming then the best result under 20 features. So, it was considered still valid to select the k value for the set of the first 20 ranked features in here.

The results for maximization of the k value are shown in Table 4.2, Table 4.3 and Table 4.4. The results for maximization of the number of features for WPG+RPG are shown in Table 4.5 and Table 4.6. The results for maximization of the number of features for WPG are shown in Table 4.7 and Table 4.8. The results for maximization of the number of features for RPG are shown in Table 4.9 and Table 4.10.

k value	**3**	**5**	**7**	**9**	**11**	**13**	**15**
Recall by 1^{st} approx. cough epochs	86.4%	86.1%	86.3%	86.7%	86.9%	86.7%	86.8%
Precision by 1^{st} approx. cough epochs	80.2%	81.7%	81.6%	83.6%	85.0%	84.9%	84.6%

Table 4.2: Results obtained for WPG+RPG analysis (5 cycles), with optimized k=11 (F1 score = 85.9%)

k value	**3**	**5**	**7**	**9**	**11**	**13**	**15**
Recall by 1^{st} approx. cough epochs	85.4%	85.4%	84.9%	84.9%	86.9%	85.2%	84.5%
Precision by 1^{st} approx. cough epochs	80.5%	81.9%	82.2%	83.0%	82.6%	81.9%	83.4%

Table 4.3: Results obtained for WPG analysis (5 cycles), with optimized k=11 (F1 score = 84.7%)

k value	**3**	**5**	**7**	**9**	**11**	**13**	**15**
Recall by 1^{st} approx. cough epochs	92.0%	88.1%	89.2%	88.5%	87.2%	87.8%	87.2%
Precision by 1^{st} approx. cough epochs	76.9%	74.2%	83.8%	78.7%	80.9%	83.8%	81.6%

Table 4.4: Results obtained for RPG analysis (5 cycles), with optimized k=7 (F1 score = 86.4%)

Number of features	**Recall by 1^{st} approx. cough epochs**	**Precision by 1^{st} approx. cough epochs**
1	0.0%	0%
2	74.0%	76.5%
3	80.1%	80.6%
4	84.1%	82.5%
5	84.6%	85.2%
6	82.6%	83.8%
7	86.2%	83.0%
8	86.3%	85.0%
9	85.5%	84.0%
10	86.8%	84.8%
11	85.5%	82.3%
12	87.1%	83.8%
13	86.4%	82.8%
14	87.9%	84.5%
15	86.4%	83.4%
16	86.3%	84.0%
17	87.3%	83.7%
18	86.3%	84.8%
19	87.3%	84.3%
20	87.0%	83.5%

Table 4.5: Results obtained for WPG+RPG analysis (5 cycles), with optimized numbers of [8 14 18 19] features (F1 score = [85.6% 86.1% 85.5% 85.8%])

Number of features	**Recall by 1st approx. cough epochs**	**Precision by 1st approx. cough epochs**
8	86.6%	84.6%
14	86.3%	84.5%
18	86.4%	83.8%
19	86.8%	83.8%

Table 4.6: Results obtained for WPG+RPG analysis (20 cycles), with optimized number of 8 features (F1 score = 85.6%)

Number of features	**Recall by 1^{st} approx. cough epochs**	**Precision by 1^{st} approx. cough epochs**
1	0.0%	0.0%
2	73.5%	76.80%
3	82.5%	80.37%
4	82.5%	81.52%
5	81.6%	83.10%
6	80.7%	82.20%
7	80.6%	82.25%
8	79.8%	81.74%
9	78.7%	82.70%
10	84.2%	81.65%
11	84.4%	82.68%
12	85.2%	81.69%
13	83.9%	82.39%
14	84.2%	85.16%
15	85.0%	82.09%
16	86.1%	84.12%
17	85.2%	81.49%
18	85.5%	83.55%
19	85.2%	83.34%
20	83.2%	83.54%

Table 4.7: Results obtained for WPG analysis (5 cycles), with optimized numbers of [12 14 16 18] features (F1 score = [83.4%, 84.7%, 85.1%, 84.5%])

Number of features	Recall by 1^{st} approx. cough epochs	Precision by 1^{st} approx. cough epochs
12	84.6%	84.0%
14	84.9%	83.7%
16	85.0%	83.0%
18	85.1%	83.4%

Table 4.8: Results obtained for WPG analysis (20 cycles), with optimized number of 12 features (F1 score = 84.3%)

Number of features	Recall by 1^{st} approx. cough epochs	Precision by 1^{st} approx. cough epochs
1	21.9%	34.0%
2	77.5%	81.8%
3	85.3%	74.3%
4	85.8%	72.9%
5	88.8%	75.1%
6	88.1%	82.3%
7	85.3%	76.4%
8	83.6%	76.9%
9	83.8%	74.8%
10	83.9%	72.9%
11	89.8%	82.6%
12	86.8%	79.8%
13	89.0%	79.9%
14	91.0%	86.3%

15	88.8%	79.6%
16	89.9%	82.8%
17	88.5%	83.6%
18	88.6%	83.6%
19	91.7%	82.5%
20	88.2%	82.7%
21	88.1%	78.8%
22	88.1%	83.8%
23	89.3%	79.8%
24	88.2%	80.7%
25	87.6%	79.4%
26	89.3%	79.6%
27	90.7%	85.0%
28	90.4%	83.6%
29	89.5%	81.5%
30	87.7%	84.0%

Table 4.9: Results obtained for RPG analysis (5 cycles), with optimized numbers of [14, 19, 26 27] features (F1 score = [88.6%, 86.9%, 87.8%, 86.9%])

Number of features	**Recall by 1^{st} approx. cough epochs**	**Precision by 1^{st} approx. cough epochs**
14	88.5%	81.6%
19	88.7%	81.6%
26	89.9%	80.1%
27	88.5%	81.2%

Table 4.10: Results obtained for RPG analysis (20 cycles), with optimized number of 19 features (F1 score = 85.0%)

With the fixed k value and number of ranked features for each testing group, the final results for the testing groups were obtained by redoing 20 cycles of repeated stratified 10-fold cross validation method. The overall results of recall and precision metrics for the cough sounds detection are present in Table 4.11.

	WPG+RPG	WPG	RPG
Recall frame by frame	42.4%	42.6%	40.5%
Precision frame by frame	65.1%	65.6%	63.2%
Recall by explosive phases	77.3%	75.2%	76.0%
Precision by explosive phases	66.4%	68.1%	61.7%
Recall by 1st approx. cough epochs	86.6%	84.6%	88.8%
Precision by 1st approx. cough epochs	84.3%	83.1%	81.9%
Recall by 2st approx. cough epochs	89.8%	88.2%	93.3%
Precision by 2st approx. cough epochs	87.3%	86.5%	85.6%
Precision by cough epochs	90.0%	88.6%	93.3%
Recall by cough epochs	87.6%	87.2%	85.6%

Table 4.11: Results obtained for all analysis (20 cycles).

The overall results are similar for the three testing groups, revealing the robustness and applicability of this approach. As more phases are being added to the final target events, better scores are achieved. This shows that some misclassified explosive phases are present in both the intermediate and voicing phases of cough sounds. This misclassified events are more present in voicing phases, because of the higher improvement in metrics from the 1^{st} approximate cough epochs to the 2^{nd} approximate cough epochs, than from the 2^{nd} approximate cough epochs to the real cough epochs. Actually, there was no improvement in this last, in the RPG group, which shows that no misclassified events occurred in the intermediate phases of cough sounds here. As it was verified during the annotation, in some subjects, voicing phases can be quite similar to the explosive phases, revealing the absence of prosody-related characteristics, such as harmonics in spectrogram, as shown in Figure 4.1.

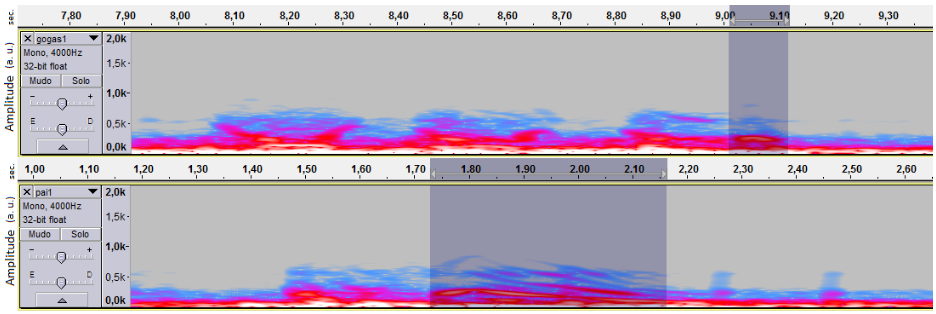

Figure 4.1: Different patterns of the voicing phase by subject: in the first signal it is highlighted the final voicing phase of one peal of three cough sounds, very similar with the initial explosive phases; in the second signal it is highlighted the final voicing phase of a single cough sound, with the presence of the harmonics related with prosody.

Since in fits of cough the chances of classifying at least one of the multiple explosive phases are higher, and only one positive hit here results in a successful classified event, it can be assumed that the increase in results in due to single cough sounds that are not classified in explosive phases, as intended, but are then detected, mostly, in the voicing phase. Although this was not the objective, it can be important to understand that some voicing phases are not as prosody-related as others, depending on the subject, and those others may show similarities with the explosive phases, in the internal sound.

The results by testing group are present in Table 4.12, Table 4.13 and Table 4.14, .

	Value
Accuracy frame by frame	95.7%
Recall frame by frame	42.4%
Precision frame by frame	65.1%
Recall by 1^{st} approx. cough epochs	86.6%
Precision by 1^{st} approx. cough epochs	84.3%
True Positive Events	6985
False Positive Events	1330
False Negatives Events	1074
Onset (mean)	17.3 ms
Onset (STD)	318.9 ms
Offset (mean)	83.1 ms
Offset (STD)	194.3 ms
Number of encompassed events	0
Time lapse between encompassed events (mean)	0 ms
Time lapse between encompassed events (STD)	0 ms
Recall (fits of cough)	59.6%
Precision (fits of cough)	91.2%
Mean energy of the signal in explosive phases detected	0.1059 a. u.
Mean energy of the signal in 1st approx. cough epochs detected	0.0724 a. u.
Mean peak energy of the signal in 1st approx. cough epochs detected	0.5561 a. u.
Elapsed time	8298 s

Table 4.12: Results obtained for WPG+RPG analysis, with k= 11 and 8 ranked features.

	Value
Accuracy frame by frame	97.1%
Recall frame by frame	42.6%
Precision frame by frame	65.6%
Recall by 1st approx. cough epochs	84.6%
Precision by 1st approx. cough epochs	83.1%
True Positive Events	5689
False Positive Events	867
False Negatives Events	728
Onset (mean)	50.2 ms
Onset (STD)	284.6 ms
Offset (mean)	232.8 ms
Offset (STD)	222.8 ms
Number of encompassed events	0
Time lapse between encompassed events (mean)	0 ms
Time lapse between encompassed events (STD)	0 ms
Recall (fits of cough)	57.7%
Precision (fits of cough)	90.7%
Mean energy of the signal in explosive phases detected	0.1049 a. u.
Mean energy of the signal in 1st approx. cough epochs detected	0.0730 a. u.
Mean peak energy of the signal in 1st approx. cough epochs detected	0.5161 a. u.
Elapsed time	7854 s

Table 4.13: Results obtained for WPG analysis, with k=11 and 12 ranked features.

	Value
Accuracy frame by frame	95.7%
Recall frame by frame	40.5%
Precision frame by frame	63.2%
Recall by 1^{st} approx. cough epochs	88.8%
Precision by 1^{st} approx. cough epochs	81.9%
True Positive Events	1551
False Positive Events	304
False Negatives Events	108
Onset (mean)	47.4 ms
Onset (STD)	184.1 ms
Offset (mean)	227.0 ms
Offset (STD)	203.6 ms
Number of encompassed events	294
Time lapse between encompassed events (mean)	182.0 ms
Time lapse between encompassed events (STD)	21.2 ms
Recall (fits of cough)	58.4%
Precision (fits of cough)	88.5%
Mean energy of the signal in explosive phases detected	0.1034 a. u.
Mean energy of the signal in 1st approx. cough epochs detected	0.0659 a. u.
Mean peak energy of the signal in 1st approx. cough epochs detected	0.6715 a. u.
Elapsed time	1191 s

Table 4.14: Results obtained for RPG analysis, with k=7 and 19 ranked features.

The achieved measures of onset and offset (mean and STD) reveal some difficulties to pinpoint the initial and final instants of the approximated cough events detected. It can be assumed that the major contribution for these values comes from deviations in detection of the fits of cough approximate events, since those have much longer duration than explosive phases alone. Therefore, in the detection of fits, if one or more explosive phases from the beginning or the ending miss the detection, the flaw is in the order of more than a 3-phase cough sound duration, around 350.7 seconds [Olia et al.,2000]. The onset vagueness should represent the expected imperfection of the detection. However, the higher values of STD for the onset suggest that the flaw in missed detections of the first explosive phase of peal events is balanced with prior detections, which can be justified by detections in the initial forced expiration before the first explosive phase. The higher values in mean offset suggest difficulties in the detection of last explosive phases of peal events, which can be understood by the gradual decrease in the intensity of the signal in explosive phases during the peal event, leading to missed detections in the end of peals. The high value of STD in offset insinuate that this missed detections are also balanced with late detections, which can be caused by detections in intermediate and voicing phases of the last 3-phase cough sound of each peal event. Figure 4.2 illustrates the possible misclassification intervals of the signal. All this values suggest that the windowing of events should be improved.

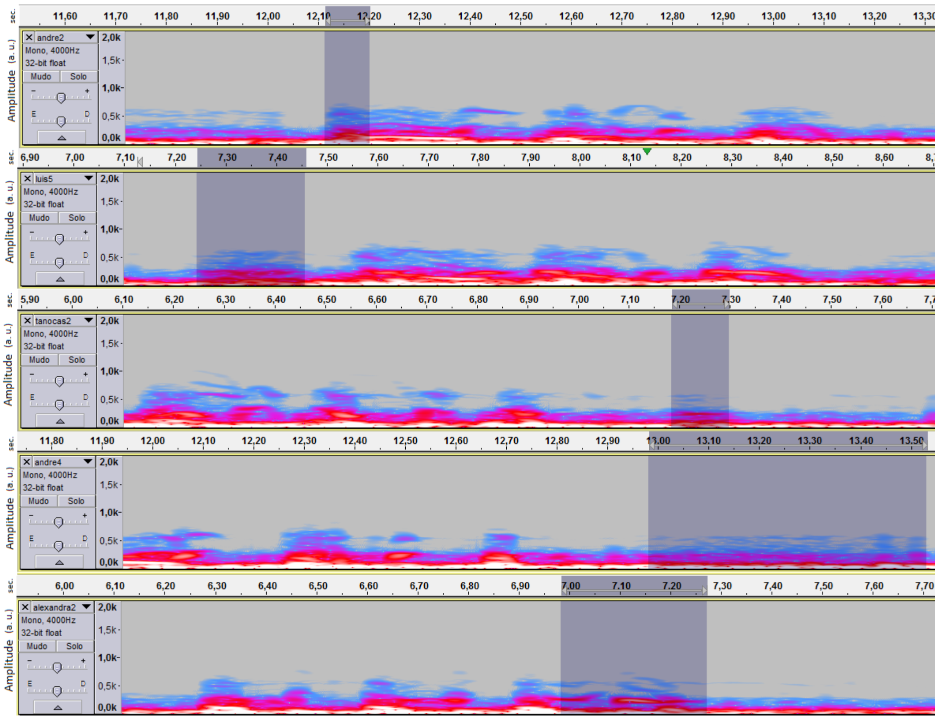

Figure 4.2: Possible justifications for misclassified detections: in the first signal it is highlighted the first explosive phase of one peal of three cough sounds, which can be missed, adding a positive onset error; in the second signal it is highlighted the initial forced expiration before the cough sound, which ca be classified as explosive phase, adding a negative onset error; in the third signal it is highlighted the last explosive phase of one peal of four cough sounds, which is weaker than the previous ones and can be missed, adding a positive offset error; in the fourth signal it is highlighted the final forced expiration after a peal of cough sounds, which ca be classified as explosive phase, adding a negative offset error; in the last signal it is highlighted the final intermediate and voicing phases, which ca be classified as explosive phase, also adding a negative offset error.

For the encompassed events, only WPG analysis obtained classified events that encompassed annotated ones. Around 18% of the all detected events were covering more than one real events. This can be the result of misclassifications between events due to high frequency-component forced inspirations and expirations of the subjects with respiratory perturbations, probably because of the presence of sputum or wheeze. In WPG+RPG and WPG analysis, no detected events encompassed the annotated ones.

The detections in all testing groups were very precise in terms of explosive phases detected on peals of cough, which evidences that the approach does not detect more explosive phases than the ones that really happened. This shows that the algorithm do not miss positive frames in the middle of an explosive phase, because if that happened, a real explosive phase could be interpreted as two or more, and there would be peals with more single events detected than the ones that really happened, increasing the rate of false positives. Since each 3-phase cough sound has explosive phase, the number of explosive phases detected can be interpreted as the same as the number of 3-phase cough sounds. So, the achieved results represent the capability to discriminate how many cough sounds can be detected in a peal event, providing important information about the pattern of coughing. However, the lower sensitivity values point out that some explosive phases of peal events are not detected, as predicted above.

The energy metrics intended to give a rough estimation of both the energy peak intensity and the overall energy of the cough sounds, one of the three characteristics of the cough sound which provide important clinical information. Its relevance is evidenced in the differences of the values for the WPG and RPG groups. The mean peak energy of the signal in the approximate cough epochs detected is substantially higher in the RPG group, which is due to a louder cough sound, with high frequencies resulting probably from the presence of sputum or wheeze, in this group. However, for mean energy of the signal in explosive phases and 1^{st} approximate cough epochs detected, the values are slightly higher for the WPG group. For the first metric, as referred in [Tracey et al.,2011], the explosive phase is characterized by a rapid increase in signal energy, and as we can see in Figure 3.2, the higher amplitudes can just be present in a small fraction of the complete explosive phase, so the differences showed in peak analysis can be attenuated by the whole contribution of the explosive phase interval. In the other metric, it can be mainly due to a significant contribution of the intermediate and voicing phases, present in the analyzed interval, which may attenuate the differences showed in peak analysis. The similar values for this metric in all groups may indicate that the mean energy of the internal signal in the cough epochs is similar to subjects with and without respiratory perturbations.

The achieved results used relatively small sets of features, and the time needed for processing one hour of recording (WPG+RPG) was around 2 hours and 7 minutes, in the processor system described in Chapter 3. We found no reference about computation times of the systems described in Chapter 2, in comparable analysis proceedings.

Our results are comparable to the mentioned works in Chapter 2, although they were not obtained on the same database. Drugman et al. [Drugman et al.,2012], with the similar explosive phase detection approach, obtained an event detection specificity and recall of 88%, for audio signal analysis in voluntary protocols performed by only 20 healthy subjects. The specificity, also known as true negative rate, is the ratio of the number of true negative events to the sum of the number of false positive events with true negative events. They also applied the same method to the signal of a contact microphone over trachea and over thorax, achieving both specificity and recall of 71%. In [Drugman et al.,2013] the signal of a contact microphone placed on the throat is compared to other signal analysis, in voluntary protocols performed by 32 healthy subjects, and results also fall short of expectations when compared with ambient audio signal analysis for same purposes, but they do not use the explosive phase detection approach. To the best of our knowledge, no other scientific works tried the internal chest sound signal alone for automatic cough counting, getting to know if our better results in a wider population are due to the acquisition site. However, this signal solves one of the disadvantages of the commonly used ambient audio signal, mentioned in [Drugman et al.,2013] - the external noise. In chest sounds signal, only the subject's internal acts can confuse the classification.

As we have seen in Chapter 2, LCM system achieved an overall recall and specificity of 91% and 99%[Birring et al.,2008], but it requires a human expert to develop a statistical model tailored for each recording. The fully automatic application of the HMM trained on ambient audio features in [Birring et al.,2008] achieved a low recall of 71%, with missed identifications in both candidates selection and cough event validation phases. Applying HMM seems promissing, but with some authors claiming that coughs acoustic differing characteristics need to be fully considered in the design of algorithms [Chunmei et al.,2013], and with this results for automatic detection, the signs may indicate a better way to solve the problem.

The HACC system achieved a specificity of 96% and a recall of 80% in ambient audio recordings from 30 smoking subjects. Although the system does not automatically count cough (just identifies, labels and shows the cough events in the graphical user interface), its identification seems promissing. The event detection logic based on adaptable thresholding select the candidates, which are then classified into cough and non-cough events by the use of PNNs trained in characteristic spectral coefficients calculated from 75 cough events and 75 non-cough events. The application of the PNNs in ambient audio signal shows acceptable results, but recall values should be increased and also needs to be assessed in a broader population.

A fully automated cough-counting industrial device, like the VitaloJAK system, seems to support this, since when experimented in continuous monitoring data from 10 subjects with respiratory perturbations achieved a recall of 97.5% and a specificity of 97.7% [McGuinness et al.,2007b]. This results are obtained with a physiological approach tailoring, with the use of acoustic parameters extracted from voluntary coughs performed by the subjects, getting to know more information about the system. It is also not revealed if it is used an explosive phase approach for the detection. What is certain is that it is not fully automatic, one of the specifications for the ideal cough monitor.

The PulmoTrack-CC system shows good performance (recall of of 96% and a specificity of 94% on voluntary cough from 12 volunteers), but requires a piezoelectric belt, one lapel microphone and two contact microphones, which despite being applicable to continuous monitoring, are not as comfortable as the use of a single signal. The other industrial device, LS system, had worse performances, with a recall of 78.1% and a precision of 84.6% for a study in continuous monitoring data from 8 COPD patients [Coyle et al.,2005].

Chapter 5

Conclusions and Future Work

The results achieved for quantification of cough showed that the explosive detection approach is a reliable method for identifying cough sounds, making use of this mandatory phase in cough, which is also the most homogeneous and predictive phase, and with continuous spectral characteristics during all the time interval for the internal sound. The explosive phase detection approach shows robustness across subjects with different respiratory perturbations and mitigates the inherent difficulty of the variety of patterns in cough. Future work on our method involves trying other classifiers, testing in ambulatory data and the use of other features. The use of inputs for tailoring parameters by each subject can also improve the efficiency, if we were willing to pay the price of loss of autonomy.

Our results refute the idea that chest sound signal analysis provide no significant performances when compared with ambient audio signal analysis, for automatic cough detection. Moreover, internal sound demonstrates advantages in one of the main challenges of audio signal analysis, the external-to-subject noise. As an important implication, we have a significant reduction on the captured sound and external speech, relevant in ethic terms. Besides, the subject internal events like cough, speech, laughter, throat clears, forced expirations, etc. seem to be better characterized by this signal. Here, some features related with tonality, pitch, timbre and spectral analysis reveal themselves very descriptive. It is worth noticing that the achieved results were originated from analyzing concise sets of features, not being necessary large loads of data and computation, something important in continuous monitoring for battery saving. Thus, the internal sound analysis through a contact microphone gathers attention for applications in new generation of smart clothing, seeming to be less inconvenient and less intrusive than a lapel microphone, and also suitable for portable and diminished batteries.

The measures of the onset and offset flaw reveal some difficulties on windowing the explosive phases. This metrics are more relevant for the onset, since the explosive phase is the first phase of the cough sound, and so, its onset is the onset of the cough sound itself. For the whole dataset, it was not detected events that covered more than one real 1st approximate cough epochs, and a high score for the number of cough sounds by peal events was attained. Along with the onset exactness, those metrics can contribute for better pattern assessing of the cough sounds.

The energy metrics of the squared sound signal are a preliminary attempt to have a rough estimation of the energy intensity invested in cough sounds, which may provide important clinical information. Further analysis between different pathologies are needed to assess the importance of this metrics.

For classification of the cough sound, no performance parameters analyzed the acoustic properties, because our approach only identifies the first explosive phase, while this characteristics are important to analyze in intermediate and voicing phases, for sputum and wheeze prediction, among other clinical information.

Improvements are needed in the post-processing methods to identify intermediate and voicing phase, and thereby obtain the complete cough event, which can then be analyzed in terms of acoustic properties. This can be solved with methods to identify the characteristic intermediate phase in the presence of sputum or finding speech related moments in a neighborhood of the explosive phase. A promising solution could be to integrate various events such as speech, throat clears, laughter and sneezes, as well as cough itself, in the classification, since chest sound signal seems to characterize those more minutely and with less ambiguity than ambient audio sound. Thereby, it could be obtained the complete cough sound, which can then be analyzed in terms of acoustic properties, as well as pattern and intensity, for better assessing.

Bibliography

[Agusti et al.,2012] A. Agusti and J. Vestbo, *Current controversies and future perspectives in chronic obstructive pulmonary disease*, American journal of respiratory and critical care medicine, vol. 184, no. 5, 2012.

[Barry et al.,2006] S. J. Barry, A. D. Dane, A. H. Morice, and A. D. Walmsley, *The automatic recognition and counting of cough*, Cough, vol. 2, no. 1, pp. 8–17, 2006.

[Birring et al.,2003] S. S. Birring, B. Prudon, A. J. Carr, S. J. Singh, L. Morgan and I. D. Pavord, *Development of a symptom specific health status measure for patients with chronic cough: Leicester Cough Questionnaire (LCQ)*, Thorax, vol. 58, no. 4, pp. 339–343, 2003.

[Birring et al.,2008] S. S. Birring, T. Fleming, A. A. Raj, S. Matos, D. H. Evans, I. D. Pavord, *The Leicester Cough Monitor: preliminary validation of an automated cough detection system in chronic cough*, European Respiratory Journal, vol. 31, no. 5, pp. 1013–1018, 2008.

[Brookes et al.,2002] M. Brookes, *Voicebox: Speech Processing Toolbox for Matlab*, Dept. Electrical and Electronic Eng., Imp. C., 2002.

[Burgel et al.,2009] P. R. Burgel, P. Nesme-Meyer, P. Chanez, D. Caillaud, P. Carr, T. Perez and N. Roche, *Cough and sputum production are* *associated with frequent exacerbations and hospitalizations in COPD subjects*, CHEST Journal, vol. 135, no. 4, pp. 975–982, 2009.

[Chang et al.,1997] A. B. Chang, R. G. Newman, P. D. Phelan and C. F. Robertson, *A new use for an old holter monitor: An ambulatory cough meter*, Eur. Respiratory J., vol. 10, pp. 1637–1639, 1997.

[Chang et al.,1999] A. B. Chang, *Cough, cough receptors, and asthma in children*, Pediatric pulmonology, vol. 28, no. 1, pp. 59–70, 1999.

[Chang et al.,2003] A. B. Chang, *Causes, assessment and measurement of cough in children*, in Cough: Causes, Mechanisms and Therapy, pp. 57–73, 2003.

[Chung et al.,1996] K. F. Chung, *Methods of assessing cough and antitussives in man*, Pulmonary Pharmacol., vol. 9, pp. 373–377, 1996.

[Chunmei et al.,2013] Z. Chunmei, T. Lianfang, L. Xiangyang, M. Hongqiang and Z. Zeguang, *Recognition of cough using features improved by sub-band energy transformation*, BMEI 6th Int. Conf., pp. 251–255, 2013.

[Coyle et al.,2005] M. A Coyle, D. B. Keenan, L. S. Henderson, M. L. Watkins, B. K. Haumann, D. W. Mayleben, and M. G. Wilson, *Evaluation of an ambulatory system for the quantification of cough frequency in patients with chronic obstructive pulmonary disease*, Cough Journal, vol. 1, no. 1, 2005.

[Decalmer et al.,2007] S. C. Decalmer, D. Webster, A. A. Kelsall, K. McGuinness, A. A. Woodcock and J. A. Smith, *Chronic cough: how* *do cough reflex sensitivity and subjective assessments correlate with objective cough counts during ambulatory monitoring?*, Thorax, vol. 62, no. 4, pp. 329–334, 2007.

[Drugman et al.,2012] T. Drugman, J. Urbain, N. Bauwens, R. Chessini, A. S. Aubriot, P. Lebecque and T. Dutoit, *Audio and Contact Microphones for Cough Detection*, 13th Interspeech An. Conf. Int. Speech Com., 2012.

[Drugman et al.,2013] T. Drugman, J. Urbain, N. Bauwens, R. Chessini, C. Valderrama, P. Lebecque, T. Dutoit, *Objective Study of Sensor Relevance for Automatic Cough Detection*, IEEE J. Biomed. Hea. Inf., vol. 17, no. 3, pp. 1637–1639, 2013.

[Georgopoulas et al.,1991] D. Georgopoulas and N. R. Anthonisen, *Symptoms and signs of COPD*, Chronic obstructive pulmonary disease. Toronto: WB Saunders Co, vol. 1, pp. 1, 1991.

[Goldsobel et al.,2010] A. B. Goldsobel and B. E. Chipps, *Cough in the pediatric population*, The Journal of pediatrics, vol. 156, no. 3, pp. 352–358, 2010.

[Hamutcu et al.,2002] R. Hamutcu, J. Francis, F. Karakoc and A. Bush, *Objective monitoring of cough in children with cystic fibrosis*, Pediatric pulmonology, vol. 34, no. 5, pp. 331–335, 2002.

[Hirtum et al.,2002] A. V. Hirtum and D. Berckmans, *Automated recognition of spontaneous versus voluntary cough*, Medical engineering and physics, vol. 24, pp. 541–545, 2002.

[Hsu et al.,1994] J. Y. Hsu, R. A. Stone, R. B. Logan-Sinclair, M. Worsdell,

C. M. Busst and K. F. Chung, *Coughing frequency in patients with persistent cough: assessment using a 24 hour ambulatory recorder*, Eur. Respiratory J., vol. 7, no. 7, pp. 1246–1253, 1994.

[Irwin et al.,1990] R. S. Irwin, F. J. Curley and C. L. French, *Chronic cough. The spectrum and frequency of causes, key components of the diagnostic evaluation, and outcome of specific therapy*, A. Rev. Respiratory Disease, vol. 141, no. 3, pp. 640–647, 1990.

[Irwin et al.,1991] R. S. Irwin and F. J. Curley, *The treatment of cough - A comprehensive review*, CHEST Journal, vol. 99, no. 6, pp. 1477–1484, 1991.

[Irwin et al.,1998] R. S. Irwin, L. P. Boulet, M. M. Cloutier, R Fuller, P. M. Gold, V. Hoffstein, A. J. Ing, F. D. Mccool, P. O'Byrne, R. H. Poe, U. B. Prakash, M. R. Pratter and B. K. Rubin, *Managing cough as a defense mechanism and as a symptom. A consensus panel report of the American College of Chest Physicians*,Chest., vol. 114, pp. 133S–181S, 1998.

[Irwin et al.,2014] R. S. Irwin, C. T. French, S. Z. Lewis, R. L. Diekemper and P. M. Gold, *Overview to the Management of Cough: CHEST Guideline and Expert Panel Report*, CHEST Journal, 2014.

[Kelsall et al.,2008] A. Kelsall, S. Decalmer, D. Webster, N. Brown, K. McGuinness, A. Woodcock and J. Smith, *How to quantify coughing: correlations with quality of life in chronic cough*, Eur. Respir. J., vol. 32, pp. 175–179, 2008.

[Klapuri et al.,2006] A. Klapuri and M. Davy, *Signal processing methods for music transcription*, vol. 1, New York: Springer, 2006.

[Korpas et al.,1979] J. Korpas and Z. Tomori, *Cough and other respiratory reflexes*, S. Karger (Basel and New York), 1979.

[Korpas et al.,1987] J. Korpas, J. Sadlonova, D. Salat and E. Masarova, *The origin of cough sounds*, Bull. Eur. Physiopathol. Respir., vol. 23, pp. 47–50, 1987.

[Korpas et al.,1996] J. Korpas, J. Sadlonova and M. Vrabec, *Analysis of the cough sound: An overview*, Pulmonary Pharmacol., vol. 9, pp. 261–268, 1996.

[Lalloo et al.,1996] U. G. Lalloo, P. J. Barnes and K. F. Chung, *Pathophysiology and clinical presentations of cough*, Journal of allergy and clinical immunology, vol. 98, no. 5, pp. S9197, 1996.

[Lartillot et al.] O. Lartillot, P. Toiviainen and T. Eerola, *MIRtoolbox v1.5*, Finnish C. of Exc. in Interdiscipl. Music Res..

[Loddenkemper et al.,2003] R. Loddenkemper, *European Lung White Book. The first comprehensive survey on respiratory health in Europe*, European Respiratory Society, 2003.

[Loudon et al.,1966] R. G. Loudon, *Cough in health and disease*, Aspen Emphysema Conference, vol. 114, pp. 41–53, 1966.

[Matos et al.,2006] S. Matos, S. S. Birring, I. D. Pavord, D. H. Evans, *Detection of cough signals in continuous audio recordings using hidden Markov models*, IEEE Transactions on Biomedical Engineering, vol. 53, no. 6, pp. 1078–1083, 2006.

[Matos et al.,2007] S. Matos, S. S. Birring, I. D. Pavord, D. H. Evans, *An Automated System for 24-h Monitoring of Cough Frequency: The*

Leicester Cough Monitor, Biomed. Eng., IEEE Transactions on, vol. 54, no. 8, pp. 1472–1479, 2007.

[McGuinness et al.,2007a] K. McGuinness, J. Morris, A. Kelsall, A. Woodcock, J. A. Smith, *The relationship between cough acoustics and the volume inspired prior to coughing*, Am. J. Resp. Crit. Care Med., vol. 175, pp. A381, 2007.

[McGuinness et al.,2007b] K. McGuinness, A. Kelsall, J. Lowe, A. Woodcock, J. A. Smith, *Automated cough detection: a novel approach*, Am. J. Resp. Crit. Care Med., vol. 175, pp. A381, 2007.

[Marsden et al.,2008] P. A. Marsden, J. A. Smith, A. A. Kelsall, E. Owen, J. R. Naylor, D. Webster and A. A. Woodcock, *A comparison of objective and subjective measures of cough in asthma*, Journal of Allergy and Clinical Immunology, vol. 122, no. 5, pp. 903–907, 2008.

[Minkoff et al.,2005] N. B. Minkoff, *Analysis of the current care model of the COPD patient: a health outcomes assessment and economic evaluation*, Journal of Managed Care Pharmacy, vol. 11, no. 6, pp. S3, 2005.

[Morice et al.,1991] A. H. Morice, L. McGarvey and I. Pavord, *Recommendations for the management of cough in adults*, Thorax, vol. 61, pp. 1–24, 1991.

[Morice et al.,2007] A. H. Morice, G. A. Fontana, M. G. Belvisi, S. S. Birring, K. F. Chung, P. V. Dicpinigaitis, J. A. Kastelik, L. P. McGarveye, J. A. Smith, M. Tatar and J. Widdicombe, *ERS guidelines on the assessment of cough*, Eur. Respiratory J., vol. 29, pp. 1256–1276, 2007.

[Munyard et al.,1994] P. Munyard, C. Busst, R. LoganSinclair and A. Bush, *A new device for ambulatory cough recording*, Pediatric pulmonology, vol. 18, no. 3, pp. 178–186, 1994.

[Munyard et al.,1996] P. Munyard and A. Bush, *How much coughing is normal?*, Archives of disease in childhood, vol. 74, no. 6, pp. 531–534, 1996.

[Olia et al.,2000] P. M. Olia, P. Sestini and M. Vagliasindi, *Acoustic parameters of voluntary cough in healthy non-smoking subjects*, Respirology., vol. 5, pp. 271–275, 2000.

[Pauwels et al.,2004] R. A. Pauwels and K. F. Rabe, *Burden and clinical features of chronic obstructive pulmonary disease (COPD)*, The Lancet, vol. 364, pp. 613–620, 2004.

[Prevention et al.,2007] P. Prevention, *Expert Panel Report 3 (EPR-3): Guidelines for the Diagnosis and Management of Asthma-Summary Report 2007*, The Journal of allergy and clinical immunology, vol. 120, no. 5, pp. S94, 2007.

[Rabe et al.,2007] K. F. Rabe, S. Hurd, A. Anzueto, P. J. Barnes, S. A. Buist, P. Calverley and J. Zielinski, *Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary*, American journal of respiratory and critical care medicine, vol. 176, no. 6, pp. 532–555, 2007.

[Robnik-Sikonja et al.,2003] M. Robnik-Sikonja, I. Kononenko, *Theoretical and Empirical Analysis of ReliefF and RReliefF*, Machine Learning, vol. 53, no. 1–2, pp. 23–69, 2003.

[Smith,2008] J. Smith, *Cough: Assessment and equipment*, The Buyers Guide to Respiratory Care Products, pp. 96–101, 2008.

[Stone,1993] R. Stone, *Chronic coughmechanisms and management*, Respiratory medicine, vol. 87, no. 4, pp. 249–251, 1993.

[Sumner et al.,2013] H. Sumner, A. Woodcock, U. Kolsum, R. Dockry, A. L. Lazaar, D. Singh, et al.,*Predictors of objective cough frequency in chronic obstructive pulmonary disease*, American journal of respiratory and critical care medicine, vol. 187, no. 9, pp. 943–949, 2013.

[Thorpe et al.,1992] C. W. Thorpe, L. J. Toop and K. P. Dawson, *Towards a quantitative description of asthmatic cough sounds*, Eur. Respir. J., vol. 5, pp. 685–692, 1992.

[Toy et al.,2010] E. L. Toy, K. F. Gallagher, E. L. Stanley, A. R. Swensen and M. S. Duh, *The economic impact of exacerbations of chronic obstructive pulmonary disease and exacerbation definition: a review*, COPD: Journal of Chronic Obstructive Pulmonary Disease, vol. 7, no. 3, pp. 214–228, 2010.

[Tracey et al.,2008] B. H. Tracey, G. Comina, S. Larson, M. Bravard, J. W. Lopez and R. H. GilmanA.Kelsall, S. Decalmer, D. Webster, N. Brown, K. McGuinness, A. Woodcock and J. Smith, *How to quantify coughing: correlations with quality of life in chronic cough*, European Respiratory Journal, vol. 32, no. 1, pp. 175–179, 2008.

[Vizel et al.,2010] E. Vizel, M. Yigla, Y. Goryachev, E. Dekel, V. Felis, H. Levi and N. Gavriely, *Validation of an ambulatory cough detection and counting application using voluntary cough under different conditions*, Cough Journal, vol. 6, no. 1, pp. 3, 2010.

[Tracey et al.,2011] B. H. Tracey, G. Comina, S. Larson, M. Bravard, J. W. Lopez and R. H. Gilman, *Cough detection algorithm for monitoring pa-*

tient recovery from pulmonary tuberculosis, In Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, vol. 1, New York: Springer, pp. 6017–6020, 2011.

URLgraphy

[57] Bronchitis Definition, (2012, December 27). In Mayo Clinic Staff. Retrieved September 3, 2014, from http://www.mayoclinic.org/diseases-conditions/bronchitis/basics/definition/con-20014956

[58] LifeShirt, a new era in ambulatory monitoring, (2005, November 17). In PDA cortex. Retrieved August 4, 2014, from http://www.pdacortex.com/VivoMetrics.htm

[59] How Is COPD Diagnosed?, (2013, July 31). In National Heart, Lung and Blood Institute. Retrieved September 7, 2014, from http://www.nhlbi.nih.gov/health/health-topics/topics/copd/diagnosis.html

[60] Theories on Cough, (2014). In Physicians Herbal Formula. Retrieved September 7, 2014, from http://www.physiciansherbalformula.com/theories-on-cough/

[61] What Is Asthma?, (2014, August 4). In National Heart, Lung and Blood Institute. Retrieved September 7, 2014, from http://www.nhlbi.nih.gov/health/health-topics/topics/asthma/

Appendices

Appendix A

Testing Population - Biometric Data

Name	Age	Sex	Height	Weight	HealthStatus
ID1	16	Fem	1.65	70	WPG
ID2	37	Fem	1.65	59	WPG
ID3	43	Fem	1.65	70	WPG
ID4	16	Masc	1.72	59	Cold
ID5	28	Masc	1.75	79	WPG
ID6	16	Fem	1.65	71	WPG
ID7	26	Fem	1.66	66	WPG
ID8	24	Fem	1.67	57	WPG
ID9	26	Fem	1.72	61	WPG
ID10	16	Masc	1.83	82	WPG
ID11	43	Masc	1.78	100	Cold
ID12	23	Masc	1.72	73	WPG
ID13	42	Masc	1.78	75	WPG
ID14	25	Fem	1.69	62	Asthma
ID15	27	Masc	1.7	66	WPG

Name	Age	Sex	Height	Weight	HealthStatus
ID16	25	Masc	1.92	81	WPG
ID17	25	Fem	1.63	48	Rhinitis
ID18	18	Masc	1.74	72	WPG
ID19	10	Masc	1.38	29	Asthma
ID20	24	Fem	1.62	55	WPG
ID21	39	Fem	1.7	90	WPG
ID22	73	Masc	1.7	78	Cold
ID23	42	Fem	1.64	82	Cold
ID24	42	Fem	1.63	88	WPG
ID25	41	Masc	1.73	90	Cold
ID26	40	Masc	1.52	70	WPG
ID27	72	Fem	1.56	74	Cold
ID28	42	Fem	1.56	54	WPG
ID29	48	Fem	1.55	70	WPG
ID30	26	Masc	1.71	70	Cold
ID31	25	Masc	1.72	70	WPG
ID32	8	Fem	1.4	50	Asthma
ID33	24	Fem	1.6	50	WPG
ID34	24	Fem	1.6	55	Cold
ID35	24	Masc	1.83	90	WPG
ID36	52	Masc	1.7	92	WPG
ID37	26	Fem	1.68	65	WPG
ID38	26	Masc	1.7	78	WPG
ID39	43	Masc	1.82	98	WPG
ID40	22	Masc	1.79	73	WPG

Name	Age	Sex	Height	Weight	HealthStatus
ID41	24	Masc	1.86	100	WPG
ID42	22	Fem	1.7	56	WPG
ID43	22	Masc	1.69	74	WPG
ID44	26	Fem	1.54	47	WPG
ID45	25	Fem	1.67	60	WPG
ID46	45	Fem	1.68	65	WPG
ID47	41	Masc	1.68	87	WPG
ID48	78	Fem	1.52	70	Bronchitis
ID49	24	Masc	1.65	60	Smoker
ID50	27	Masc	1.78	73	WPG

Table A.1: Biometric Data for the whole testing population. For ethical reasons, the names are fictitious. Parameters: Name, Age (years), Sex (Feminine/Masculine), Height(meters), Weight (kilograms), Health Status (WPG - without respiratory perturbations, other - with respiratory perturbations)