University of Southampton

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

Institute of Sound and Vibration Research

A novel method for automatic separation of pulmonary crackles from normal breath sounds

by

Ravi Pal

Thesis for the degree of Doctor of Philosophy

March 2021

University of Southampton

Abstract

FACULTY OF ENGINEERING AND PHYSICAL SCIENCES

Institute of Sound and Vibration Research

Doctor of Philosophy

A novel method for automatic separation of pulmonary crackles from normal breath sounds

by

Ravi Pal

Pulmonary crackles are an important physiological parameter for evaluating lung condition of an individual and usually determined at auscultation by conventional stethoscope. The presence of crackles is generally an early indication of the disease and their number per breath cycle can indicate the severity of the disease. A conventional stethoscope placed on the chest wall can identify the presence of crackles, but this approach is subjective and the accurate detection of crackles and the identification of their type (fine or coarse) is highly dependent on clinician hearing ability and expertise. The misinterpretation of crackles may lead to inappropriate treatment of the patient. Computer aided lung sound analysis (CALSA) using advanced signal processing techniques can provide an objective way of analysing recorded lung sounds and hence can play important role in diagnosing or monitoring pulmonary diseases.

In this study, a novel crackle separation technique: iterative envelope mean fractal dimension (IEM-FD) filter is developed for automatically separating crackles from normal breath sounds. The separation of crackles from normal breath sounds is an initial processing stage which can lead to better estimation of crackle features such as number of crackles and two-cycle deflection. To test the crackle separation ability of the IEM-FD filter, a dataset was generated. The performance of the IEM-FD filter was compared with the selected previously published crackle separation techniques using the developed dataset. The experimental results show the proposed method can achieve high accuracy for the number of crackles identified with low computational cost, better quality of crackle separation (less over or underestimation), and good preservation of crackle morphology and hence it may be useful in a clinical setting for determining number of crackles and characteristics of crackles in a recorded lung sound.

The proposed IEM-FD filter is applied to two different datasets: (a) longitudinal dataset recorded from 19 idiopathic pulmonary fibrosis (IPF) patients in 7 visits (every visit in 2 months) over a 1

year time period and (b) Cross-sectional dataset recorded from 55 subjects who were referred for a high-resolution computed tomography (HRCT) scan of the chest for various clinical indications.

In the longitudinal study application of the IEM-FD filter prior to counting the number of crackles present, allowed evaluation of the association between number of crackles per breath cycle (NOC/BC) and reproducible acoustic features directly generated from the original signal. In this study, it was found that some of these acoustic features were directly associated with NOC/BC therefore might be useful for monitoring progression of IPF.

In the cross-sectional study, the IEM-FD filter was applied as a first stage of an automatic crackle counting system which can be used for differentiating idiopathic pulmonary fibrosis patients from patients with other types of pulmonary pathology based on the average NOC/BC. The diagnosis given by two radiologists using the HRCT scan was used as a gold standard for classifying IPF and non-IPF groups. The ability of the automatic system to differentiate IPF patients from non-IPF patients was compared with the individual and average assessment by two respiratory physicians based on listening for the presence of Velcro crackles. Velcro crackles are generally considered as an early clue to the presence of fibrosis. The results show that the automatic system can perform as well as the expert physicians' assessments and hence could support the auscultatory findings of lung sounds in less specialist clinics.

In both the longitudinal and cross-sectional studies, in each recorded lung sound file the number of breathing cycles was audio-visually marked by the Author with the help of open access Audacity software. Audio-visual marking is a highly time-consuming process, therefore a new automatic breath cycle detection method based on the estimation of breathing phases was developed. The performance of the method was tested using both the longitudinal and cross-sectional datasets, and a dataset recorded from 10 healthy subjects in 7 visits (each visit in 2 months) over a period of 1 year. The promising results show the possibility of the developed algorithm as an automatic method for breath cycle detection in lung sound recordings.

Table of Contents

Table of Tables

Table of Tables

Table of Figures

Table of Figures

Table of Figures

Table of Figures

Research Thesis: Declaration of Authorship

Print name: Ravi Pal

Title of thesis: A novel method for automatic separation of pulmonary crackles from normal breath sounds

I, Ravi Pal declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:-

    - *R. Pal and A. Barney. (2021). Iterative envelope mean fractal dimension filter for the separation of crackles from normal breath sounds. Biomedical Signal Processing and Control, 66, 1–12.*

    - *R. Pal and A. Barney. (2020). Pulmonary Crackle Detection using the Hilbert Energy Envelope. Proc. 8th European Medical and Biological Conf. - EMBES'20, Slovenia, pp. 994-1003.*

    - R. Pal and A. Barney. (2019). A dataset for systematic testing of crackle separation techniques*. Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. - EMBS'19, Berlin, pp. 4690–4693.*

Signature: ..... Date: 26/03/2021

Acknowledgements

I would first and foremost like to thank my supervisor Professor Anna Barney for her expertise, ideas, suggestions, and encouragement throughout this journey. I am grateful for your guidance and valuable feedback, which helped me in all the time of this research. This work would not have been possible without your continuous support.

Secondly, I would like to acknowledge Dr. Giacomo Sgalla whose collected datasets during his PhD work and research helped in many chapters of this thesis. I would like to thank Dr. Jordan Cheer for reviewing my 18-month report and special thanks to Professor Paul White for reviewing my 9- and 18- month report.

I am grateful to my family, especially my parents and sisters, for their patience, support and unconditional love. I would also like to thank my friends for their help and motivation during this long journey.

I would like to thank the National Institute for Health Research (NIHR) Southampton Biomedical Research Centre, the Engineering and Physical Sciences Research Council (EPSRC), and the Asthma, Allergy & Inflammation research (AAIR) Charity for funding this research.

Main definitions and Abbreviations

NOC/BC	Number of crackles per breath cycle
IDW	Initial deflection width
2CD	Two cycle deflection
IPF	Idiopathic pulmonary fibrosis
CALSA	Computer aided lung sound analysis
HRCT	High-resolution computed tomography
BE	Bronchiectasis
IEM-FD	Iterative envelope mean fractal dimension
LDW	Largest deflection width
CFA	Cryptogenic fibrosing alveolitis
COPD	Chronic obstructive pulmonary disease
HF	Heart failure
FA	Fibrosing alveolities
SG	Savitzky-Golay
FD	Fractal dimension
FDPP	Fractal dimension peak peeling
TVAR	Time varying autoregressive
FIR	Finite impulse response
DIP	Diffuse Interstitial Pneumonia
ST-NST	Stationary non-stationary
WTST-NST	Wavelet transform stationary non-stationary
MRD	Multiresolution decomposition
MRR	Multiresolution reconstruction
WPST-NST	Wavelet packet stationary non-stationary

WT-FD	Wavelet transform fractal dimension
EMD	Empirical mode decomposition
IMFs	Intrinsic mode functions
EMD-FD	Empirical mode decomposition fractal dimension
FST-NST	Fuzzy based stationary non-stationary
ANFIS	Adaptive-network-based fuzzy inference system
GFST-NST	Generalized fuzzy rule based stationary-non-stationary
OLS-FF	Orthogonal least squares based fuzzy filter
ICA	Independent component analysis
IC	Independent component
Infomax	Information maximization
ACC	Accuracy
UOE	Under-, over-estimation
POC	Preservation of crackle morphology
CCX	Analysis speed (computational complexity)
NRB	Robustness to additive/environmental noise
OBJ	Objectivity
SFC	Simulated fine crackles
SCC	Simulated coarse crackles
RFC	Real fine crackles
RCC	Real coarse crackles
RBFC	Real breath sound with fine crackles
RBCC	Real breath sound with coarse crackles
NOC	Number of crackles

D_g	Diagnosis
BN	Background noise
SNR	Signal to noise ratio
ND	Not defined
NA	Not applicable
BR_N	Breath noise
W_N	Gaussian white noise
NBS	Normal breath sound
f_s	Sampling frequency
W_{FD}	Fractal dimension window
W	Duration of the window in seconds
F_D	Fractal dimension estimate of the waveform
F_{DPP}	Estimated fractal dimension peak peeling sequence
SD	Standard deviation
σ	Standard deviation
QMF's	Quadrature mirror filters
m	WT scale index
M	Number of wavelet transform scales
n	Sample index
N	Number of samples
F_{adj}	Multiplication factor
$\beta 1, \beta 2$, and $\beta 3$	Accuracy level
$STC1, STC2$, and $STC3$	Stopping criterion
k, q, and l	Iteration index
K, Q, and L	Total number of iterations
U	Filter order

b	Filter length index
B	Filter length
Th	Threshold
IEM	Iterative envelope mean
CCI	Cross correlation index
D_R	Rate of detectability
TD_R	Total performance
QFs	Quality factors
PE_{2CD}	Two cycle deflection percentage error
NOTS	Number of test samples
Min	Minimum value
Max	Maximum value
S_T	Separation time
CORSA	Computerized respiratory sound analysis
ICC	Intra correlation coefficient
FVC	Forced vital capacity
ANST	Absolute non-stationary
PCW	Potential crackle window
PC	Potential crackle
BW	Before window
AW	After window
BC	Breath cycle
ROC	Receiver operating characteristic
I	Inspiratory phase
E	Expiratory phase
SE	Sensitivity

PPV	Positive predictive value
F_1	F-score
TP	True positive
FN	False negative
FP	False positive

Chapter 1 Introduction

This thesis presents a novel automatic method for separating pulmonary crackles from normal breath sounds. When air is drawn into the lungs for the purposes of respiration, normal breath sounds are generated because of the turbulent airflow in the bronchial tree. Pulmonary crackles which overlap with normal breath sounds, can be an early first indication of the presence of a pulmonary disease.

Auscultation is a medical term, which refers to the process of listening to the sounds generated from inside the body using a stethoscope or any other devices (Pramono et al., 2017). Auscultation provides a non-invasive way of lung examination and is useful in diagnosing various pulmonary diseases (Sarkar et al., 2015). In recent years, the advancement in electronic stethoscopes has opened the field of computer-aided auscultation (Leng et al., 2015). The electronic stethoscope provides an opportunity to capture, record, playback, and analyse the recorded lung sounds through advanced digital signal processing techniques (Malik et al., 2017).

Crackles are short duration, non-musical lung sounds, which may occur during inspiration or expiration (Nath & Capel, 1974). The number of crackles per breath cycle (NOC/BC) and features related to the shape of the crackle signature in a recorded lung sound can have clinical significance. The NOC/BC can be used for assessing severity of the disease, whereas initial deflection width (first deflection of a crackle) and two-cycle deflection (duration of the first 5 zero crossing of a crackle) can be used for classifying crackle types (fine or coarse); where fine crackles are generally associated with interstitial lung diseases and coarse crackles are more often related with obstructive airway diseases (Du et al., 1997).

The existence of normal breath sounds may introduce errors: when estimating the number and shape of crackles in a recording, which can lead to incorrect assessment of disease type or severity leading to sub-optimal treatment or management of the condition. Therefore, it can be advantageous to separate crackles from normal breath sounds before analysing them.

This thesis starts with a brief introduction of lung sounds, different types of pulmonary crackles, the clinical importance of pulmonary crackles, limitations of crackle analysis using conventional stethoscope and motivation behind selecting the research topic. Research aim and objectives, an overview of the thesis, research contributions, and publications from this study will also be presented in this chapter.

1.1 Lung Sounds

Lung sounds can provide useful information for assessing and monitoring pulmonary patients (Marques et al., 2009). Lung sounds are divided into two categories: normal lung sounds (or normal breath sounds) and adventitious lung sounds. Normal breath sounds are heard in healthy as well as pathological lungs and result from the flow of air through the airways.

Normal breath sounds are produced due to turbulent airflow along the trachea-bronchial tree during the respiration process. Turbulent airflow is caused by the high velocity of flow travelling through a large diameter airway, particularly along an airway with irregular walls, such as the trachea and bronchi or in the airway with sudden branching. The nature of turbulent airflow is disorganised and chaotic. On the other hand, the flow in the small airways is laminar and silent in nature. The mechanism of noise generation due to turbulence involves the collision of air molecules with the airway wall and with each other (Sarkar et al., 2015).

Adventitious lung sounds are additional sounds, which usually occur with respiratory disorders (Yeginer & Kahya, 2008; Serbes et al., 2013). The adventitious lung sounds can be continuous (wheezes) and discontinuous (crackles) (Reichert et al., 2008; Dinis et al., 2012; Jacome & Marques, 2017). In this context, the word ‘continuous’ refers to a duration of more than 250 ms rather than a sound that continues throughout the respiratory cycle (Meslier et al., 1995).

The musical sound of wheezes can easily be recognized by simple hearing (Taplidou & Hadjileontiadis, 2007). Wheezes are generated due to airway limitations caused by the narrowing of airways (Nagasaka, 2012; Pramono et al., 2019). Wheezes are superimposed on normal breath sounds and usually louder than the underlying breath sounds. They can appear during the inspiration or expiration (Polat & Guler, 2004). In some patients, they may be audible at some distance from the patient (Loudon & Murphy, 1984). Wheezes can be heard in patients with asthma and chronic obstructive pulmonary disease (Bohadana et al., 2014; Henry & Royston, 2018).

On the other hand, crackles are known as discontinuous adventitious lung sounds (Yeginer & Kahya, 2005). They are short duration (less than 20 ms); explosive sounds of a non-musical character (Earis et al., 1992; Loudon & Murphy, 1984; Speranza et al., 2020; Ellington et al., 2012, Nath & Capel, 1974; Grzywalski et al., 2019). Crackles can be heard on the chest wall in patients with cardiopulmonary diseases such as cystic fibrosis, pneumonia, fibrosing alveolitis, bronchiectasis (BE), sarcoidosis, congestive heart failure, and asbestosis (Kiyokawa et al., 2001). During deep inspiration, crackles may occasionally occur in healthy subjects (Sovijarvi et al.,

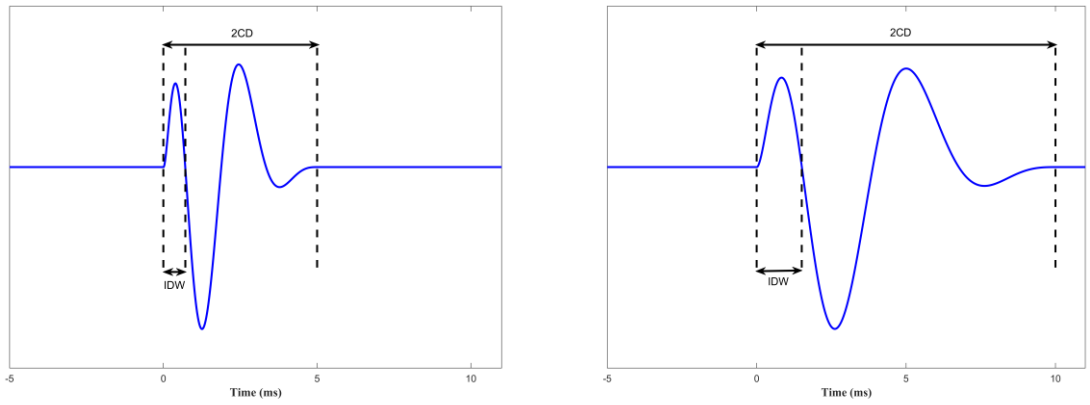
2000a). The frequent occurrence of crackles in a patient is usually an early indication of lung abnormality.

1.1.1 Different types of pulmonary crackles

The first deflection of a crackle is known as the initial deflection width (IDW) and the time to complete the first five zero crossing is known as the two cycle deflection (2CD). Pulmonary crackles may be characterized as either fine or coarse sounds based on IDW and 2CD (Charbonneau et al., 2000). According to the American Thoracic Society, the average IDW and 2CD of fine crackles are 0.7 ms and 5 ms, and for coarse crackles are 1.5 ms and 10 ms, respectively (Charbonneau et al., 2000).

Fine crackles are thought to be generated due to the explosive reopening of small airways that closed during the previous expiration (Munakata et al., 1991). Fine crackles are usually mid- to late-inspiratory events that follow a similar pattern for each consecutive inhalation. These crackles can be an indication of pneumonia, congestive heart failure and various pulmonary fibrotic diseases (Pramono et al., 2017). Fine crackles have a high pitch, greater than 400 Hz (Vyshedskiy & Murphy, 2012).

On the other hand, coarse crackles are typically early inspiratory and expiratory events. Coarse crackles may be generated from fluid in small or medium airways and have a popping quality. These crackles can change pattern after coughing and are a symptom of, for example, chronic bronchitis, BE and cystic fibrosis (Kraman, 1993). The sound of coarse crackles is low pitched, less than 400 Hz (Vyshedskiy & Murphy, 2012). The waveform of typical fine and coarse crackles with their time domain features are shown in Figure 1 (a) and Figure 1(b), respectively.



(a) (b)

Figure 1 Fine and coarse crackles: (a) fine crackle with IDW=0.7 ms and 2CD= 5 ms, (b) coarse crackle with IDW=1.5 ms and 2CD=10 ms.

1.2 Clinical significance of crackles

Crackles are clinically important for several reasons:

(1) The crackle sounds superimposed on the normal breath sounds can be used for diagnosis (Speranza et al., 2020). The classification of crackle type has important clinical significance: the fine crackles are associated with interstitial lung diseases and coarse crackles are common in obstructive airway diseases (Du et al., 1997). Moreover, crackle features can be used for differentiating different pulmonary diseases. For example, Flietstra et al., (2011) showed that based on different crackle features (number of crackles in inspiratory phase, number of zero-line crossings, and first half period of the crackle) idiopathic pulmonary fibrosis (IPF) patients can be differentiated from patients with pneumonia and congestive heart failure.

(2) The NOC/BC may play an important role in early detection or monitoring the prognosis of interstitial lung disorders. In the initial phase of IPF, crackles are generated in the base of the lungs and as the disease progresses crackles start to be produced in upper zones of the lungs (Cottin et al., 2012). Therefore, the NOC/BC is associated with the disease severity in patients with interstitial lung disorders (Sovijarvi et al., 2000a; Rocha et al., 2019).

(3) The timing of crackles within the breathing cycle allows the direct estimation of sound origin (Kompis et al., 2001). Smaller airways have been shown to produce late inspiratory crackles of high frequency, short duration less than 10 ms (fine crackles) whereas larger airways tend to produce early (inspiratory/expiratory) crackles with low frequency, longer duration greater than 10 ms (coarse crackles) (Marques et al., 2009).

1.3 Limitations of crackle analysis using a conventional stethoscope

In 1816, Rene Laennec started the science of auscultation (Alvarado & Arce, 2016; Piirila & Sovijarvi, 1995). This was the first stethoscope for listening to lung sounds and was made of wood and paper. It allowed him to assess a patient's lung condition without touching the patient (Andres et al., 2018). Since 1800s, significant improvement has been made in the stethoscope (Andres et al., 2018). The conventional stethoscope is a most popular tool in clinics for assessing lung condition and can provide a non-invasive way of examining lung diseases. However, interpretation of the sound is subjective (Jacome & Marques, 2015; Hafke-Dys et al., 2019; Guler et al., 2005). The ability to differentiate lung sound patterns is highly dependent on expertise and hearing ability of the observer (Kandaswamy et al., 2004; Oliveira & Marques, 2014). Furthermore, this method is limited by high inter-observer variability (Gurung et al., 2011; Rao et

al., 2019; Spieth & Zhang, 2011). If the intensity of the normal breath sounds is very high it can mask the additive crackle sounds and, in that case, accurate detection of the number of crackles or their type (fine or coarse) using the convention stethoscope can be very challenging.

The limitations of the human hearing system mean that auscultation in the hospital environment, when background noise is often very high, can result in crackles whose intensity is low being missed. Furthermore, auscultation using a conventional stethoscope cannot provide continuous monitoring (Rocha et al., 2019). The technology has changed much over the years and computer based lung sound analysis (CALSA) can provide an objective way of analysing lung sounds (Betiencourt et al., 1994; Pasterkamp et al., 2016). CALSA minimises the inter-observer variability of standard lung sounds auscultation and can provide an objective and automatic way of analysing recorded lung sounds (Kaisla et al., 1991; Emmanouilidou et al., 2018).

An automatic crackle separation method using advanced signal processing techniques can assist with diagnosing lung diseases and in monitoring disease progression. By separating, the normal breath sounds from the crackles both the large amplitude crackles and the small amplitude crackles, which are often significantly masked by the breath sounds, may be revealed. This can provide better estimation of number of crackles present. Furthermore, crackle time domain features such as IDW and 2CD can be accurately estimated. These can be used for crackle classification but, due to the background normal breath sounds, their waveform can be distorted which can mislead about the values of IDW and 2CD (Yeginer & Kahya, 2008). Therefore, for the better estimation of the crackle features it is important to eliminate normal breath sounds before analysing the crackle characteristics.

An automatic crackle separation technique with high accuracy for number of crackles identified which is robust to noise could be used in a clinical environment for analysis of recorded crackle sounds. Therefore, the large part of this research focuses on developing a novel automatic crackle separation technique, suitable for use in a clinical setting for recorded lung sounds analysis, which may be useful to support clinical decision-making.

1.4 Research aim and objectives

1.4.1 Aim

(1) The aim of this study is to develop a new automatic crackle separation technique, which can separate crackles from normal breath sounds with high accuracy for number of crackles identified, low computational complexity, good quality of crackle separation (low over or under estimation), high noise robustness, preservation of crackle morphology

after separation, and few requirements to make decisions about process based on the data (high objectivity).

1.4.2 Objectives

(2) Develop a dataset, which can be used for systematic testing of crackle separation techniques and published to encourage standardized testing between studies.

(3) Provide systematic comparison between the proposed method and the selected previously published crackle separation methods using the developed dataset.

(4) Test the new method on real data and explore its potential to identify clinical outcomes.

(5) Explore the association between NOC/BC and global reproducible acoustic features directly extracted from the original signal (Sgalla, 2017).

(6) Assess the ability of an automatic system to differentiate IPF patients from patients with other types of pathology who were referred for high resolution computed tomography (HRCT) scan of the chest for various clinical indications based on the average NOC/BC calculated using the lung sounds recorded at the lung bases.

(7) Compare the classification performance (IPF or non-IPF) of an automatic system with the individual and average assessment of two experienced physicians who listened to recorded lung sound files to classify them as containing Velcro crackles (IPF) or not (Non-IPF).

(8) Develop an automatic breath cycle detection algorithm based on the estimation of the breathing phases.

1.5 Research contributions and Thesis overview

The review of the literature on different methods used for crackle lung sounds analysis is presented in Chapter 2.

Contribution 1: A dataset was developed which can be used for systematic testing of different crackle separation techniques. The dataset not only contains real fine and coarse crackles but also has simulated fine and coarse crackles with different IDW and 2CD, and real lung sounds with fine and coarse crackles recorded from a person with IPF and a BE patient. The dataset also provides the option of two types of background noise: Breath noise and Gaussian white noise.

The detailed description of the dataset is provided in Chapter 3.

From the literature, three previously published crackle separation techniques were selected with which to compare the crackle separation performance of the proposed iterative envelope mean fractal dimension (IEM-FD) filter. The reason for selecting these methods and their detailed working process is presented in Chapter 4.

Contribution2: A novel automatic crackle separation technique i.e. IEM-FD filter was developed, which can provide high accuracy for the number of crackles identified with low computational cost, good quality of crackle separation (less under or over estimation), good preservation of crackle morphology after separation and high noise robustness. The filter can be used for analysing real lung sounds recorded in clinical environment.

The detailed working process of the proposed method is provided in Chapter 5.

The crackle separation performance evaluators used in this study for assessing the proposed and previous methods are presented in Chapter 6. The systematic comparison between the proposed IEM-FD filter and the selected previously published crackle separation techniques using the test dataset is made in terms of crackle identification accuracy, quality of crackle separation (over or underestimation), computational cost and the ability to preserve crackle morphology after separation. This evaluation is presented in Chapter 7.

Contribution 3: The proposed IEM-FD filter was used in two case studies:

(a) Longitudinal study: In this study, the dataset recorded from 19 IPF patients in 7 visits over a 1-year time was analysed using the proposed IEM-FD filter. From the analysis, it was found that the number of fine 'Velcro' crackles per breath cycle (NOC/BC) was highly correlated with the reproducible global acoustic features directly extracted from the original acoustic signal. Therefore, these reproducible acoustic features might be used for monitoring IPF. This study makes the first link between those global features and the underlying crackle sounds.

(b) Cross-sectional study: in this study, a dataset recorded from 55 subjects who were referred to a specialist pulmonary clinic for a HRCT scan of the chest for various clinical indications was analysed using the new IEM-FD filter. From this study, it was found that the average NOC/BC can be used for differentiating IPF patients from patients with other types of pathology.

The methodology and results of the two case studies are presented in Chapter 8.

Contribution 4: When analysing the datasets in the patient case studies the breath cycles were audio-visually marked in each lung sound signal. The manual marking of a large number of breath cycles is highly time consuming. Therefore, an automatic breath cycle detection algorithm based

on the estimation of the breathing phases was developed. The algorithm was evaluated against the manually marked data for the longitudinal, cross-sectional studies, and a dataset recorded from healthy subjects. However, this algorithm was not used in this thesis for analysing the two case studies.

The detailed description of the algorithm is provided in Chapter 9.

The discussion of the findings of this research with conclusions and future possibilities are provided in Chapter 10.

1.6 Publications

Journal publications

- *R. Pal and A. Barney. (2021). Iterative envelope mean fractal dimension filter for the separation of crackles from normal breath sounds. Biomedical Signal Processing and Control, 66, 1–12.*

Conference papers

- R. Pal and A. Barney. (2020). Pulmonary Crackle Detection using the Hilbert Energy Envelope*, Proc. 8th European Medical and Biological Conf. - EMBES'20, Slovenia, pp. 994-1003.*
- R. Pal and A. Barney. (2019). A dataset for systematic testing of crackle separation techniques*, Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. - EMBS'19, Berlin, pp. 4690–4693.*

Chapter 2 Literature review on crackle lung sounds analysis

2.1 Introduction

In recent years, lots of work has been done in the field of computer based crackle lung sounds analysis. Automatic crackle detection and automatic separation of crackles from normal breath sounds are two areas of significant activity in the computer-based crackle analysis literature. Automatic separation of crackles from the lung sound signal is a first step in crackle analysis which can facilitate better estimation of the number of crackles and of their time domain features such, IDW or 2CD. This chapter will provide a review of the different methods proposed in the literature for crackle separation and analysis. The topics covered in this chapter are outlined below:

2.2 Different methods for crackle lung sounds analysis

Human hearing is an unreliable way of crackle detection. Kiyokawa et al., (2001) presented a study to check the human ear's capability to detect crackles in an auscultation signal. The audibility of crackles was tested by superimposing simulated crackles (fine, medium and coarse) with large and small amplitude on recorded breath sounds at 0 L/s (breath hold), 1 L/s and 2 L/s airflows. They showed that failed detections are more frequent in the following conditions: (1) crackle amplitude is small, (2) background breath sound is of higher intensity (2 L/s), where 2 L/s shows volume flow rate, and (3) crackle type is either coarse or medium.

Murphy et al., (1977) proposed time-expanded waveform analysis for visually analysing the recorded lung sounds. This approach showed that digitally recorded lung sounds waveform can be analysed by viewing the waveform on a screen and providing the user to control the magnification of the image. Although visual lung sound analysis can help to differentiate between the normal breath sounds and crackles, this approach is highly time-consuming and limited by large inter observer variability (Hadjileontiadis & Panas, 1997).

Use of CALSA with the advanced signal processing techniques may improve reliability. Computer based crackle counting was compared with visual crackle counts and audible counts in a study by Murphy et al., (1989). In this study, 100 samples of lung sounds from 41 subjects were recorded using an electret microphone air coupled to the chest wall. The recorded sounds were examined by a trained technician using the crackle identification criteria: (i) the crackle waveform has to cross the baseline between three and sixteen times, (ii) the amplitude of the largest deflection

width (LDW) has to be two times greater than the amplitude of the background sound (the width of the largest absolute peak of a crackle is known as LDW), (iii) the beginning of the event had a sharp deflection in either a positive or a negative direction, and (iv) the baseline crossings after the initial deflection have to be progressively further apart. After the visual analysis two physicians separately counted the number of crackles by listening to the recorded sounds. Now, using the Spearman's rank correlation computer based crackle counts were compared with visual counts and auditory counts. The correlation coefficient between the computer counts and the visual counts was 0.85 ($p<0.001$) while the correlation coefficient between computer counts and audible counts was 0.74 ($p<0.001$). For auditory counting, crackles which are highly masked by high intensity background sound can be missed (Kiyokawa et al., 2001). In addition, multiple crackles, occurring in a short interval are difficult to count for humans by listening (Murphy et al., 1989). On the other hand, the visual analysis is a highly time-consuming approach (Hadjileontiadis & Panas, 1997). Although, in Murphy's (1989) study the computer based lung sound crackle counter was compared with the visual and audible counts, the ability of the computer algorithm to count crackles in different noisy conditions when crackle morphology is completely hidden by the background normal breath sounds was not considered.

In another study, a multi-channel lung sound analyser (model STG-1602, Stethographics, Westborough, Massachusetts) was used to test whether recorded lung sounds differed significantly between pneumonia patients and subjects who had no clinical evidence of pneumonia (Murphy et al., 2004). In this method, 14 microphones were used for simultaneously collecting the lung sounds data and one microphone was used to record tracheal sounds. All the microphones were incorporated into a soft foam pad which the subject lay supine upon. The STG system software was used to automatically identify the adventitious sounds in accordance with published definitions (Murphy et al., 2004) and the visually-based time expanded waveform analysis was used to verify the automatic analysis. In this study, they selected 100 patients who were diagnosed by their physicians as having pneumonia and 100 subjects who had no clinical evidence of pneumonia (control subjects) but were age-matched to those of the pneumonia group (age>60). All of these 100 subjects were patients who came to an internist for annual physical examination. After selection of the patients, they divided subjects into two categories: - learning sample (50 patients from each group) and test sample (remaining 50 patients in each group). Now for each subject, an acoustic pneumonia score was generated by adding crackle score and rhonchus score. Rhonchus is a low frequency wheeze which contains a rapidly damping periodic waveform with a duration of > 100 ms (Sovijarvi et al., 2000b). Crackles and rhonchus were used for calculating the pneumonia score because their rate per breath cycle is associated with the higher likelihood of pneumonia. The rhonchus score was estimated using the proportion

of the breath cycle occupied by rhonchi. A rhonchus rate of 4-5 % was assigned a score of 3, a rate of 6-10 % was received a score of 5, and a rate of 11-100 % was assigned a score of 6. Crackle score was calculated using the number of crackles per respiratory phase. The maximum crackle score was 10 for inspiratory phase and 10 for expiratory phase. The minimum and maximum pneumonia scores were therefore 0 and 26, respectively. The performance of pneumonia score was first tested on the learning sample and then examined on the test sample. An average pneumonia score of 13 in the learning sample and 11 in the test sample of pneumonia patients was found. An average pneumonia score of 2 in the learning sample and 3 in the test sample of control subjects was found. Furthermore, by investigation, they found that almost all (91%) of the subjects had some adventitious sounds and that crackles were the most common finding. Further, they observed significant differences in lung sounds between pneumonia patients and subjects who had no clinical evidence of pneumonia. This study proposed that computer based analyses could be used to quantify the auscultatory abnormalities related to pneumonia and help to provide important clinical information for diagnosing pneumonia.

Piirila et al., (1991) studied the crackling lung sounds in patients with cryptogenic fibrosing alveolitis (CFA), BE, chronic obstructive pulmonary disease (COPD) and heart failure (HF). The waveform, frequency, and timing of crackles in 10 patients with CFA, 10 patients with BE, 10 patients with COPD, and 10 patients with HF were analyzed. The key findings of the study were: (a) In CFA the upper frequency limit of the inspiratory sounds was higher compared to that in HF or in COPD, (b) In COPD period of crackling within a respiratory cycle was shorter compared to in CFA or in BE, (c) In COPD inspiratory crackling terminated significantly earlier in the breath phase than in CFA, in BE or in HF, (d) The IDW and 2CD were shorter in CFA than in COPD, BE, or HF, (e) In CFA the LDW was smaller compared to in COPD, BE, or HF. In another study, Baughman et al., (1991) found that the crackles are more frequent in fibrosing alveolitis (FA) compared to in sarcoidosis. Furthermore, Ponte et al., (2013) found that maximum frequency and 2CD of crackles may allow differentiation between crackles generated by fibrosis from those generated by HF and pneumonia. Flietstra et al., (2011) demonstrated that crackle features can be used for differentiating IPF from patients with pneumonia and congestive heart failure. These studies clearly indicate that the crackles may have different features in different pulmonary diseases. However, the presence of normal breath sounds may mislead about the number of crackles and their time domain features such as IDW and 2CD. The incorrect estimation of crackle features due to the background normal breath sounds may increase the chances for misdiagnosis. Hence, better estimation of crackling features may help in differentiating different pulmonary diseases. An automatic crackle separation technique can separate crackles from normal breath sounds and may help in better estimation of crackle features. In recent years, much effort has been made to

automate separation of crackles from normal breath sounds. Before talking about published crackle separation techniques automatic crackle detection methods proposed in the literature are discussed below.

2.2.1 Automatic crackle detection techniques

Different methods have been proposed in the literature for automatic detection of crackles in recorded lung sounds. Kaisla et al., (1991) presented an automatic crackle detection technique based on analysing the spectral stationarity of the lung sound. This method was validated using 10 patients with FA and 10 patients with BE. For patients with FA the method achieved a sensitivity of 89 % and a positive predictive value of 88 % and for patients with BE the method obtained a sensitivity of 80 % and positive predictive value of 83 %. The performance of the automatic method was compared with the number of crackles counted by two observers using time expended waveform analysis (Murphy et al., 1977). The linear correlation coefficients between the automatic method and the number of crackles counted by the observers was 0.86 ($p<0.001$) and 0.93 ($p<0.001$) for the patients with FA and BE, respectively.

Vannuccini et al., (1998) proposed an automatic method for detecting and analysing crackles in respiratory sounds. This method is based on two steps. In the first step, a threshold value is applied to the first derivative absolute value of the input signal, with the goal to identify the crackle's location. In the next step, the identified location can be considered the location of a crackle if it conforms to certain conditions. The crackle conditions (height of the peaks and their distance from the starting point of the crackle) are verified within a temporal window. A Savitzky-Golay (SG) filter is used for finding the first derivative absolute value of the input signal (Savitzky & Golay, 1964). As a reference, an expert observer scans the recordings and identifies visually where the crackles are using Murphy's criteria (Murphy et al., 1989) and by simultaneously listening to the signal. The algorithm was tested on a sample of 200 inspiratory crackles from 15 patients with CFA and achieved sensitivity of 84 % and specificity of 89 %.

Hadjileontiadis & Rekanos, (2003) proposed a fractal dimension (FD) technique for automatically detecting crackles in recorded lung sounds. The FD technique is based on the FD of the recorded lung sounds. The FD evaluates the complexity of the recorded lung sounds in the time domain and provides a way to track the time location of crackles (Hadjileontiadis & Rekanos, 2003). In the FD technique, an overlapped temporal window is applied to the input signal for estimating FD. Once the FD of the signal is estimated, the fractal dimension peak peeling (FDPP) algorithm is applied to the estimated FD vector for automatically detecting the location and duration of the crackles in the input signal through FD peaks. The FDPP algorithm is an iterative process, which iteratively

peels the estimated FD vector. In this way, the FDPP algorithm not only searches for the high peaks of the estimated FD vector, which may correspond to high amplitude crackles but also looks for the low peaks of the estimated FD vector, which may correspond to low amplitude crackles. The detailed working of the FD technique is provided in section 4.3. The algorithm was tested on 54 pulmonary fibrosis and 19 interstitial fibrosis fine crackles, and 41 chronic bronchitis coarse crackles, and achieved a 100 % accuracy for the number of crackles detected for both fine and coarse crackles.

Pinho et al., (2015) proposed a technique based on filtering the FD. In this algorithm, firstly a window of interest of a potential crackle was extracted. A window of interest was estimated using different signal processing steps: (a) In the first step, SG filter was used for smoothing an input signal. Smoothing was done for removing the high frequency noise from an input signal. (b) In the second step, FD of the smoothed signal was calculated using a sliding window. (c) In the third step, box filtering (average filtering, with a sliding window was used on the estimated FD vector for calculating the trend (smoothed version of the estimated FD vector). (d) Now the estimated FD vector was compared to a threshold to identify the window of interest of a potential crackle, the threshold used is proportional to the smoothed FD. After extraction of a window of interest of a potential crackle, the probability of the potential crackle is verified using respiratory sound analysis established criteria and conditions established empirically by Pinho et al., (2015): (i) the amplitude of the different peaks of the crackles had to be progressively lower than the LDW peak. (ii) peaks had to be progressively wider after the IDW. (iii) minimum of 5 zero crossings, to guarantee the calculation of 2CD, and a maximum of 16 zero crossings. (iv) the mean absolute amplitude of the crackle had to be higher than two times the mean absolute amplitude of the background noise (v) crackle's IDW had to be higher than 1/8 of the LDW. The algorithm was tested on twenty-four 10-second files, acquired in clinical settings, from 10 patients, 6 with pneumonia and 4 with cystic fibrosis. Here, three-experienced respiratory researchers performed the manual annotation. All three researchers were experienced in visual-auditory crackle identification and independently annotated the beginning and end of each crackle in each sound file using the respiratory sound annotation software V1.1 (Dinis et al., 2012) and a crackle was considered as present when at least two researchers agreed. The performance of the algorithm was evaluated by comparing the maximum absolute peak within each identified crackle with the multi-annotation gold standard obtained from the annotations of the three researchers. The algorithm achieved a sensitivity of 89 %, positive predictive value of 95 % and F-score of 92 %.

Recently, Reyes et al., (2018) presented an automatic crackle detection algorithm based on time varying autoregressive (TVAR) modelling. The idea behind the algorithm was that non-stationary events such as crackles may produce abrupt changes in the coefficients of a TVAR model

compared to the basal respiratory sounds. In this algorithm, firstly, an input signal was pre-processed: (a) a 500th order finite impulse response (FIR) band pass filter between 75 Hz and 1000 Hz was used to minimize the presence of heart sounds and other muscular noises, and (b) to account for different amplitude variations between recordings all filtered signals were normalized in amplitude in the range [-1, 1]. Secondly, a 4th order TVAR model was used and the TVAR coefficients were calculated using the recursive least squares algorithm. Now, after estimating the TVAR coefficients, the derivative of the each TVAR coefficient time series was estimated to enhance the abrupt changes due to crackles. In the next step, a sliding window of length 4 ms was used for segmenting each estimated derivative and the standard deviation under each window was calculated. A threshold value was used inside each window for finding the presence of a crackle. In each window, if the standard deviation of the all derivative coefficient time series was above the threshold value, then the absolute values of all derivatives were added together. Now, in each window, from the added absolute values of derivatives the maximum point was calculated which indicated the starting point of the detected crackle. The algorithm was tested on: (a) simulated fine and coarse crackles randomly inserted in the basal lung sounds recorded from 10 healthy subjects with different SNRs, and (b) lung sounds recorded from 9 Diffuse Interstitial Pneumonia (DIP) patients. In healthy subjects lung sounds were recorded at a posterior left basal location. On the other hand, for each DIP patient lung sounds were recorded at three pulmonary zones indicated by the physician. The physician performed a pulmonary auscultation using a mechanical stethoscope and indicated pulmonary zones in each patient based on the presence of crackle sounds. The lung sounds were recorded using an electret subminiature microphone (BT-2159000, Knowles Electronics, Itasca, IL, USA) enclosed in a plastic bell. In the case of lung sounds recorded from DIP patients, crackles were manually counted by pneumologists via time-expanded waveform analysis and listening to the signals (Audio-visual marking). Due to cumbersome and time demanding manual crackle detection by the pneumologists, from each recording only one breathing cycle was analysed. The algorithm obtained on average, an accuracy ranging from 84.86% to 89.16%, a sensitivity ranging from 93.45 % to 97.65 %, and a specificity ranging from 99.82% to 99.84%, in the inserted simulated fine crackles scenarios. On the other hand, the algorithm achieved on average, an accuracy ranging from 57.95 % to 85.18 %, a sensitivity ranging from 64.02% to 94.68%, and a specificity ranging from 99.83 % to 99.85 %, in the inserted simulated coarse crackles scenarios. However, the performance of the algorithm was drastically decreased in the real data case (DIP patients lung sound recordings) where on average, accuracy of 51 % and specificity of 63 % were found.

2.2.2 Limitations of automatic crackle detection techniques

Automatic crackle detection techniques can detect crackles in the lung sound signal; however, they are not able to separate crackles from normal breath sounds. As a result:

(a) Detection directly from unprocessed recordings may fail to detect crackles which are deeply buried in the background normal breath sounds (small amplitude crackles) or may detect false crackles due to the presence of background normal breath sounds, hence may provide poor estimation of the number of crackles.

(b) Crackle time-domain characteristics (IDW, 2CD) may be different in different pulmonary diseases but the presence of background normal breath sounds may distort these measures and hence may lead to misclassification of crackle types (fine or coarse) and hence, potentially, to misdiagnosis.

On the other hand, the separation of crackles from normal breath sounds before counting or characterising can provide better estimates of the number of crackles and their time domain features. However, an optimal process for crackle separation, useable in a clinical context, should have: low computational complexity, high accuracy for the number of crackles identified, good separation quality with neither under- nor over-estimation of the crackle waveform, ability to preserve crackle morphology after separation, high robustness to noise, and less requirement to make decisions about process based on the data (high objectivity). Note that the failure to extract all crackles or loss of some portion of the crackle in the output signal is referred as under-estimation, and the inclusion of non-crackle components as over-estimation. The published crackle separation techniques in the literature are discussed below.

2.2.3 Automatic crackle separation techniques

There has been much effort in the field of automatic separation of crackles from normal breath sounds. The level slicer and high pass filter can separate crackles from normal breath sounds to some extent but they fail to separate small amplitude crackles, which are deeply buried in the background normal breath sounds and further can distort the crackle morphology in the process. To analyse the effect of high pass filtering on the morphology of the crackles, a case study was presented by Katila et al., (1991). In this study, a digital filter (Kaiser FIR) and an analogue filter (4th order Butterworth) with seven different cut-off frequencies: 25, 50, 75, 100, 150, 200 and 400 Hz were used to analyse the effect of high pass filtering on the crackle waveform. Lung sounds with crackles were recorded from one patient with silicoasbestosis. According to the analysis, it was noticed that the effect of cut-off frequency on the crackle IDW and the duration of

the crackle was very low up to cut-off frequencies of 100 Hz but at higher cut-offs the IDW and duration both gradually decreased with increase in cut-off frequency up to the 400 Hz cut-off level. The duration and IDW of the crackle were more influenced by the analog filter compared to the digital filter, although the digital filter influenced the polarity of the IDW more. It was observed from the study not only the high pass filter cut-off frequency but also the filter type influenced the crackle waveform. If the high pass filter cut-off frequency is too high, crackle waveforms (especially coarse crackles) can be distorted and for the low cut-off frequency, background noise will not be eliminated from the crackle (Yeginer & Kahya, 2008).

Ono et al., (1989) presented a nonlinear digital filter for the automatic separation of crackles from normal breath sounds, which is known as stationary non-stationary (ST-NST) separating filter. This method used the autoregressive prediction and the coefficients of the autoregressive model were adaptively updated using the least mean square algorithm (Widrow et al., 1976). However, the authors note that although the ST-NST filter can be useful for counting the number of crackles after separating them from normal breath sounds, this method requires empirical setting of many of its parameters (Hadjileontiadis & Panas, 1996). A requirement of empirical setting of the filter parameters according to the characteristics of the input signal is not ideal for a clinical setting.

The time duration of crackles is typically in the range of 10 – 20 ms and on this time scale, normal breath sounds may be considered quasi-stationary and therefore separation of lung sounds into estimates of the non-stationary and stationary parts generally sends crackle components of the lung sound signal to the non-stationary signal estimate and the majority of components associated with normal breath sounds to the stationary signal estimate.

Hadjileontiadis & Panas, (1996) provided a modified version of the ST-NST filter i.e. mST-NST filter by combining the ST-NST filter with third order statistics. In this method, autoregressive prediction was performed based on third order statistics. The mST-NST filter provided more robust results in the noisy case, but as for the ST-NST filter this filtering method requires empirical setting of the set of its parameters (Hadjileontiadis & Panas, 1996).

In another study, the wavelet transform stationary non-stationary (WTST-NST) filter was proposed by (Hadjileontiadis & Panas, 1997) for automatically separating crackles from normal breath sounds. This method is based on an iterative multiresolution decomposition (MRD) and multiresolution reconstruction (MRR) scheme which separates wavelet transform coefficients related to the signal of interest from the background normal breath sounds using a threshold value at each decomposition level. The WTST-NST filter is based on the contention that explosive peaks have large components over many wavelet scales, but coefficients related to background normal breath sounds reduce with increasing wavelet scale. Although the WTST-NST algorithm

showed good results in terms of crackle identification accuracy, due to the non-adaptive characteristics of the threshold value (hard thresholding) and high computational complexity, it is still not ideal to use in a clinical environment or for analyzing datasets of recorded crackle lung sounds where high processing speed is advantageous for decision-making.

Lu & Bahoura, (2008) presented the wavelet packet stationary non-stationary (WPST-NST) filter for separating crackles from normal breath sounds. In the wavelet packet transform not only the approximation coefficients vector but also the detail coefficients vector is decomposed into two parts at each wavelet scale. Similar to the WTST-NST filter, this filtering method is also based on the fact that explosive peaks have large components over many wavelet scales, while coefficients related to the background normal breath sounds decreases with increasing wavelet scale. Unlike the WTST-NST filter, the WPST-NST filter is not an iterative process. The WPST-NST filter uses two thresholds for separating wavelet transform coefficients corresponding to crackles from background normal breath sounds. The two thresholds are defined in time domain and frequency domain respectively. The WPST-NST filter may provide fast crackle separation compared to the WTST-NST filter, but the non-adaptive characteristics of its two thresholds limits overall generalization of the algorithm.

In order to overcome the empirical setting of the threshold for separating the wavelet transform coefficients related to crackles and normal breath sounds in the WTST-NST filter, the wavelet transform fractal dimension (WT-FD) filter was proposed (Hadjileontiadis, 2005(I), Hadjileontiadis, 2005(II)). In this method the FD technique (Hadjileontiadis & Rekanos, 2003) was used for thresholding. Firstly, the input signal is decomposed into approximation and detail coefficient vectors using the wavelet transform, and after decomposing the signal, the FD technique is applied to separate crackle coefficients from normal breath sounds. The FD of the signal is estimated within an overlapped window shifted repeatedly by one sample along the signal's length. The combination of the WT and the FD can overcome the limitation of the non-adaptive threshold used in the WTST-NST filter but here also the accurate selection of base wavelet type and number of WT decomposition scales may be critical for its overall performance (Hadjileontiadis, 2007). A need to optimize these parameters for any given data set prior to use in a clinical setting would be less than ideal.

The empirical mode decomposition (EMD) based analysis of real respiratory data and simulated crackles was presented by (Charleston-Villalobos et al., (2007). Different cases of crackles (fine, coarse, overlapped crackles, a combination of fine and coarse crackles etc.) were investigated using the EMD technique (Huang et al., 1998). In this study, the simulated crackles were generated using the mathematical functions proposed by Kiyokawa et al., (2001). After generation

of simulated crackles, they were inserted in the normal inspiratory sound (only the inspiratory phase was used for the analysis) recorded from the posterior right clavicular location of the healthy subject. Next the input signal was decomposed into different intrinsic mode functions (IMFs) and a residual signal using the EMD technique. For the next step in EMD, all the IMFs were visually analysed (qualitative analysis) and crackle information related IMFs were identified. It was observed that mostly IMF1 to IMF4 contained relevant information about the crackles, and after IMF4, the respiratory sound became dominant, and it was impossible to identify or extract further information about crackles from the IMFs. Moreover, it was observed that the EMD technique does not have the capacity to individually separate the individual crackles that overlap by a large amount. Further, these analyses were carried out using just a single inspiratory sound from a single healthy subject and the Influence of pulmonary diseases on the number of IMFs containing elements related to crackles was not considered.

The EMD technique provides a solution to denoising the explosive lung sounds but is not sufficient on its own for separating the non-stationary (crackles) and stationary parts (normal breath sounds) of the input signal. After combining the IMFs obtained from the EMD technique, which are related to the non-stationary parts of the input signal, some of the noise, (normal breath sounds) is still present in the reconstructed signal. To overcome this problem, the combination of FD and EMD based EMD-FD filter was proposed by Hadjileontiadis, (2007) for separating the stationary and non-stationary parts from the input signal. In this denoising method, the input signal was first decomposed into different IMFs and the residual component. After decomposition, a number of IMFs, are selected using an energy criterion, as being related to the crackles. Next, the FD technique was applied to the selected IMFs for separating out background normal breath sounds. Once background normal breath sounds is separated from the selected IMFs, the remaining parts are combined to estimate the explosive lung sounds. The separated background normal breath sounds, a combination of remaining IMFs and residual component, provides the overall background normal breath sounds in the input signal. Although this method gives good separation of crackles from normal breath sounds, the non-adaptive characteristics of the energy criterion lead to challenges in the selection of how many independent mode functions contain crackle information. Too many IMFs may lead to overestimation and too few to underestimation.

Different fuzzy based filters were also tested for their ability to separate crackles and normal breath sounds. Tolias et al., (1997) presented a fuzzy based stationary non-stationary (FST-NST) filter. This method is based on training a fuzzy inference system constructed according to the adaptive-network-based fuzzy inference system (ANFIS) approach (Jang, 1993). In this filtering method, the ANFI system was trained using the stationary and non-stationary outputs of the

WTST-NST filter (Hadjileontiadis & Panas, 1997), obtained after applying the WTST-NST filter to the recorded lung sounds. Furthermore, Tolias et al., (1998) proposed a modified version of the FST-NST filter i.e. generalized fuzzy rule based stationary-non-stationary (GFST-NST) filter. In this filtering method, instead of using two ANFIS's operating in parallel as in the FST-NST filter, a serial combination of two ANFIS's is used. As for the FST-NST filter in this method the stationary and non-stationary outputs of the WTST-NST filter were used for training the fuzzy inference system. In another study, Mastorocostas et al., (2000) presented an alternative fuzzy model, the orthogonal least squares based fuzzy filter (OLS-FF). The OLS-FF improved the ANFIS structure used in the FST-NST filter (Tolias et al., 1997) and the GFST-NST filter (Tolias et al., 1998) by introducing a more flexible structure, which uses a lower number of rules. The OLS-FF also used two fuzzy inference systems operating in parallel. Furthermore, unlike the FST-NST filter (Tolias et al., 1997) and the GFST-NST filter (Tolias et al., 1998) the OLS-FF does not require a training phase for calculating the optimum model parameters. Recently, the computational intelligence based filter has been proposed by Kandilogiannakis & Mastorocostas, (2018). This filtering method uses two dynamic fuzzy neural networks that operate in parallel for separating crackles from normal breath sounds. In this method, the simulated annealing dynamic resilient propagation algorithm is used for training the algorithm.

All four fuzzy filters described above can provide separation of crackles from background normal breath sounds with low computational load, but amongst all four fuzzy filters the GFST-NST filter provided the best performance in terms of high accuracy of number of crackles identified. The crackle separation performance of all four fuzzy filters was not tested in terms of noise robustness.

Recently, Garcia et al., (2020) proposed a methodology for automatic extraction of fine and coarse crackles by independent component analysis (ICA). The efficiency of three ICA algorithms, i.e. FastICA, Information maximization (Infomax) and temporal decorrelation source separation were evaluated using the Amari index, the total relative distortion index, and the signal to interference ratio. From their analysis, it was found that in simulated, multichannel signal scenarios, the Infomax algorithm provided the best separation of crackles from normal breath sounds compared to the other two algorithms. The presence of crackles in the independent component (IC) obtained from the Infomax algorithm was determined using the skewness and kurtosis and the spectrogram of the selected IC was used to find the type of crackles. However, this methodology was only tested on simulated crackles and with breath sounds recorded from healthy subjects. The real crackle scenario was not tested in this study; the real crackles may have different temporal morphology compared to simulated crackles used in this study. Moreover, this

Table 1 Technical characteristics for good crackle separation compared for different published separation methods.

Methods	**ACC**	**UOE**	**POC**	**CCX**	**NRB**	**OBJ**
VTEWA (Murphy et al., 1977)	+	x	x	--	--	--
LS & HPFs (Katila et al., 1991)	-	-	--	+	x	+
ST-NST (Ono et al., 1989)	-	-	-	x	x	--
mST-NST (Hadjileontiadis & Panas, 1996)	+	x	x	x	+	--
WTST-NST (Hadjileontiadis & Panas, 1997)	+	+	-	-	+	-
WT-FD (Hadjileontiadis, 2005(I), Hadjileontiadis, 2005(II))	++	+	+	+	+	-
EMD-FD (Hadjileontiadis, 2007)	+	+	+	--	+	--
FST-NST (Tolias et al., 1997)	+	+	-	+	x	-
GFST-NST (Tolias et al., 1998)	+	+	-	+	x	-
OLSF (Mastorocostas et al., 2000)	+	+	-	+	x	-
Neurofuzzy filter (Kandilogiannakis & Mastorocostas, 2018)	+	x	x	+	x	-
WPT (Lu & Bahoura, 2008)	+	x	x	+	+	-
ICA (Garcia et al., 2020)	+	x	x	x	x	-

ACC: Accuracy (number of crackles correctly separated); UOE: Under-, over-estimation; POC: Preservation of crackle morphology; CCX: Analysis speed (computational complexity); NRB: Robustness to additive/environmental noise; OBJ: Objectivity (need to set hard thresholds and/or make decisions about process based on the data and/or requirement of training phase for estimation of the optimum model parameters); ++ = strong attribute; + = acceptable attribute; - = weak attribute; - - = very weak attribute; x = attribute not reported.

study requires specialist equipment with 25 recording channels and adoption of such bespoke systems in clinical settings is unlikely.

2.2.4 Different published separation methods

Despite the various crackle separation techniques proposed in recent years; their use in a clinical setting has so far been quite limited due to several factors. A summary of the strengths and weaknesses of each method is given in Table 1. As it can be seen in Table 1, the main limitation with existing automatic crackle separation methods is the high computation complexity and/or failure to preserve crackle morphology after separation and/or low objectivity. These factors limit their utility in the clinical environment. The high computational cost is not advantageous in clinical setting where fast processing is desirable for decision-making (Tolias et al., 1998; Kandilogiannakis & Mastorocostas, 2018). On the other hand, failure to preserve crackle morphology after separation may lead to incorrect estimation of crackle time domain features such as IDW and 2CD. Furthermore, need to set hard thresholds and/or make decisions about process based on the data and/or the requirement for a training phase for estimation of the optimum model parameters reduces their objectivity. The above mentioned limitations clearly indicate that there is need of a new automatic crackle separation technique which can incorporate all the factors of an optimum process and can be useful in clinical setting for recorded lung sounds analysis.

2.3 Summary

Chapter 2 presented a review of the different methods published in the literature for crackle lung sound analysis. Automatic crackle detection techniques can identify crackles in the lung sound signal but they are not able to separate crackles from normal breath sounds. The separation of crackles from normal breaths sounds as a first step can lead to better estimation of crackle time domain features such as number of crackles, IDW, 2CD etc.

In this thesis, considering all the points for an optimal process for crackle separation, a novel IEM-FD filter is presented for automatic crackle separation. The performance of the IEM-FD filter is compared with three previously published methods (WTST-NST filter, WT-FD filter and EMD-FD filter) and applied to two different case studies. To compare the crackle separation performance of the IEM-FD filter with the previously selected methods a dataset was generated. Before discussing the IEM-FD filter and the reason for selecting the previously published methods for comparison with their detailed working process; a dataset developed for systematic testing of crackle separation techniques is presented in Chapter 3.

Chapter 3 Dataset for systematic testing of crackle separation techniques

3.1 Introduction

In the last chapter, we saw various different methods proposed in the literature for recoded crackle lung sounds analysis and considered their limitations. In this chapter, a dataset generated for systematic testing of crackle separation techniques is described.

3.2 Dataset

Different pulmonary diseases may have different type of crackle (fine or coarse). As Du et al., (1997) showed, fine crackles are common in interstitial lung diseases and coarse crackles are associated with obstructive airway diseases. Therefore, the performance of any crackle separation technique needs to be evaluated using a dataset containing different types of crackles. It is also important to test the crackle separation ability of an algorithm when crackles are deeply buried in background normal breath sounds to test the noise robustness of an algorithm. Due to the absence of publicly available lung sound datasets, testing of crackle separation of the separation methods is often fairly limited in terms of range and/or number of test samples. Furthermore, every research group has used a different dataset for testing their algorithm hence crackle separation performance of different algorithms published by different research groups can't be compared with each other especially when there is a need to make decisions about process based on the data. Therefore, a dataset is developed in this study, which can be used for systematic testing of a crackle separation technique. We have published this dataset (10.1109/EMBC.2019.8857928), which is available at https://doi.org/10.5258/SOTON/D0801 to encourage more consistency in testing regimes in future. This open access dataset provides a common platform to all research groups so that they can test their algorithm. The dataset consists of: (a) simulated fine and coarse crackles, and real fine and coarse crackles with a range of values for IDW and 2CD, (b) noise with the spectral characteristics of breath noise (referred to here after as 'breath noise') and (c) a real breath sound with fine crackles recorded from a patient with IPF and a real breath sound with coarse crackles recorded from a patient with BE, these signals were recorded with an electronic stethoscope and sampled at 44,100 Hz. In this thesis, we have selected the sampling rate of 44, 100 Hz because it is recommended by Cheetham et al., (2000) for respiratory sound recordings. The detail description of the dataset is presented below.

The dataset consists of three subsets.

- Test crackles comprising: (1) simulated fine crackles (SFC), (2) simulated coarse crackles (SCC), (3) real fine crackles (RFC), and (4) real coarse crackles (RCC).
- Test noise comprising: (1) Gaussian white noise, (2) colored noise with the same long-term spectral properties matching that of a healthy breath sound recorded at the posterior right chest location.
- Test samples comprising: (1) real breath sound with fine crackles (RBFC), and (2) real breath sound with coarse crackles (RBCC).

3.2.1 Simulated fine and coarse crackles

The simulated fine and coarse crackles are generated using the mathematical function defined by (Kiyokawa et al., 2001), where the crackle waveform $x(t)$ is given by:

$$x(t) = x_0(t)m_1(t) \tag{1}$$

With

$$x_0(t) = sin(4\pi t^{\alpha}), \alpha = \frac{log\ (0.25)}{log\ (t_0)} \tag{2}$$

$$m_1(t) = 0.5(1 + cos\ [2\pi(t^{0.5} - 0.5)]) \tag{3}$$

where $x_0(t)$ is a progressively wider sinusoidal function with first positive zero interception time at t_0 and $m_1(t)$ is a modulating function to shift the power of $x_0(t)$ to the beginning of the crackle waveform. On the basis of IDW and 2CD, three cases of fine and coarse crackles are generated: (1) American Thoracic Society definition (Charbonneau et al., 2000) with: fine crackle IDW = 0.7 ms an 2CD = 5 ms, coarse crackle IDW = 1.5 ms and 2CD = 10 ms; (2) based on Hoevers (Hoevers & Loudon, 1990) with fine crackle IDW = 0.5 ms and 2CD = 3.3 ms, coarse crackle IDW = 1 ms and 2CD = 5.1 ms and (3) based on Cohen (Cohen, 1990) with fine crackle IDW = 0.9 ms and 2CD = 6 ms and coarse crackle IDW = 1.25 ms and 2CD = 9.5 ms.

The time domain features of all three cases of simulated fine and coarse crackles are shown in Figure 2 & Figure 3. In the dataset for each simulated crackle case, a set of ten identical crackles was generated.

3.2.2 Real fine and coarse crackles

We selected 10 RFC from our lung sound database recorded from a patient with IPF and 10 RCC recorded from a patient with BE. All selected crackles followed at least

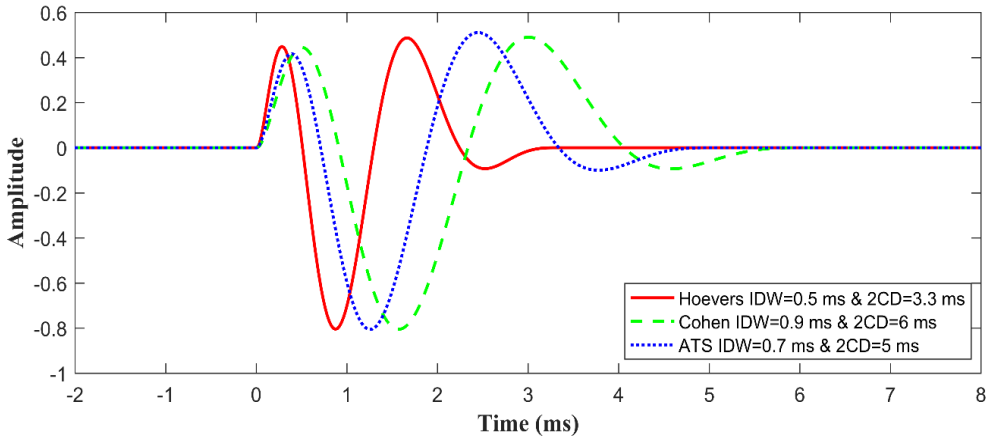


Figure 2 Simulated fine crackles.

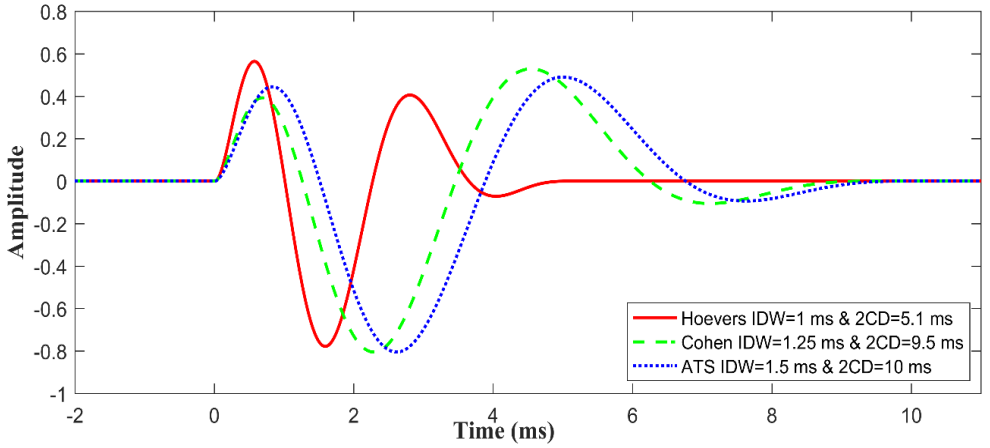


Figure 3 Simulated coarse crackles.

three of the criteria defined by Murphy (Murphy et al., 1989) as characteristic of a pulmonary crackle following the characteristics outlined in section 2.2. Figure 4 displays the 10 RFC selected from a patient with IPF and the 10 RCC selected from a patient with BE are presented in Figure 5.

3.2.3 Simulated test signals

Each set of crackles was buried in two types of noise: 1) Gaussian white noise and 2) noise with the same spectrum as breath noise from a healthy subject measured over the lung bases on the right-hand side of the back and SNR ranged from -10 to 10 dB in steps of 1 dB. Note that the interference of heart sounds on the normal breath sounds is minimum at the posterior base locations so this is the location used for generating the breath noise. Figure 6 shows the normal breath sound with no crackles recorded from a healthy subject at the location of posterior right, which was used for generating breath noise. A 10^{th} order autoregressive model is used for generating a noise signal with the same long-term spectral properties as the input signal. Figure 7 and Figure 8 show the generated noise signal and power spectral density plot, respectively.

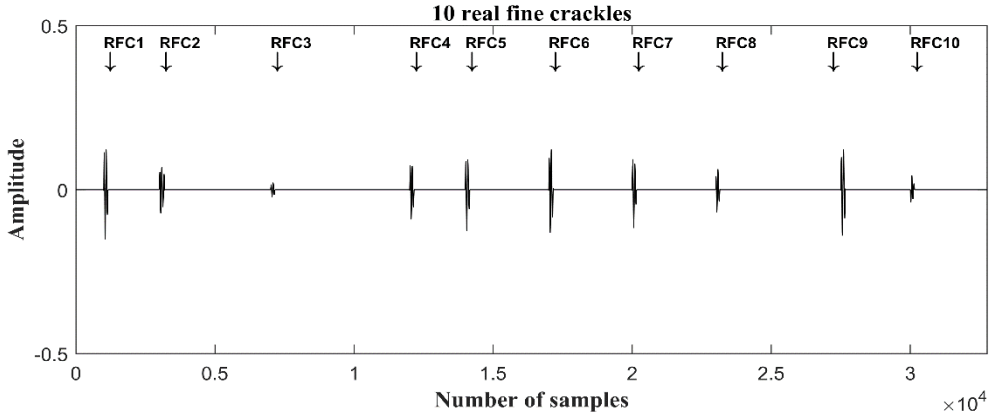


Figure 4 10 Real fine crackles (RFC1-RFC10) selected from a patient with IPF.

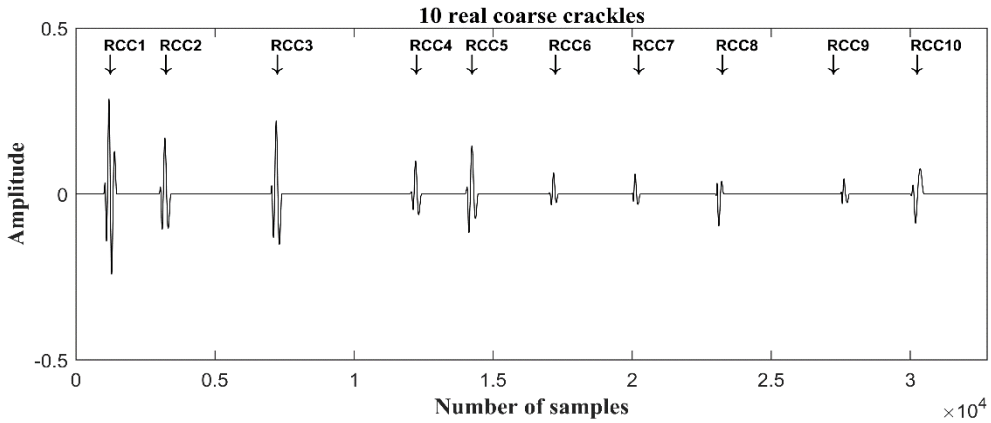


Figure 5 10 Real coarse crackles (RCC1-RCC10) selected from a patient with BE.

3.2.4 Real breath sound with fine and coarse crackles

In addition to the simulated data, two examples of real breath sound, one from a patient with IPF with predominantly fine late-inspiratory crackles and one from a patient with BE with coarse, mainly expiratory crackles, are included in the dataset. These sound files can provide useful comparisons between separation techniques to evaluate their performance in separating crackles from normal breath sounds in real lung sounds. The real breath sound signals are displayed in Figure 9 and Figure 10.

Table 2 summarizes the different test cases. Note that the developed dataset provides test cases of (a) crackles embedded on background Gaussian white noise, (b) crackles buried in background breath noise, and (c) RBFC recorded from a IPF patient and RBCC recorded from a BE patient. However, the Gaussian white noise was not used in this study because it is not at all realistic. Therefore, only test cases with crackles embedded on background breath noise, RBFC recorded from a patient with IPF, and a RBCC recorded from a patient with BE were used for analyzing the crackle separation performance of the proposed and previously selected methods.

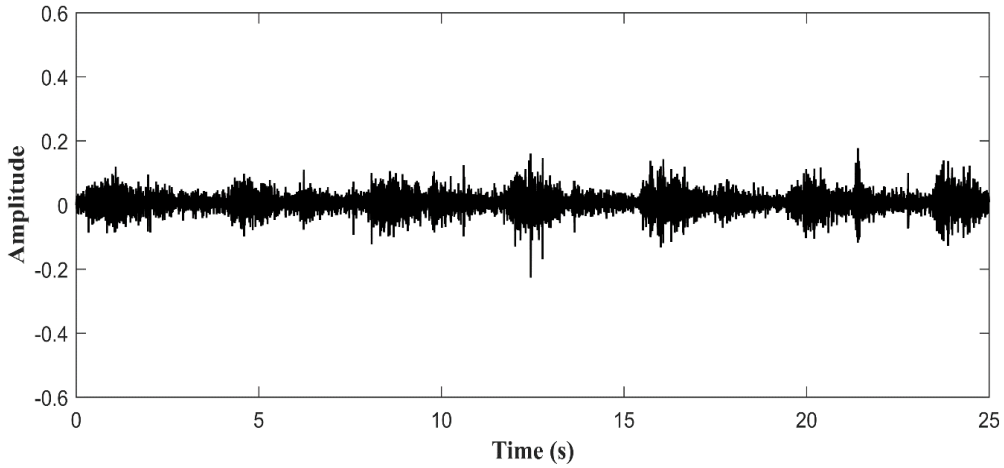


Figure 6 Breath signal used for generating breath noise in the range of -10 to 10 dB signal to noise ratio.

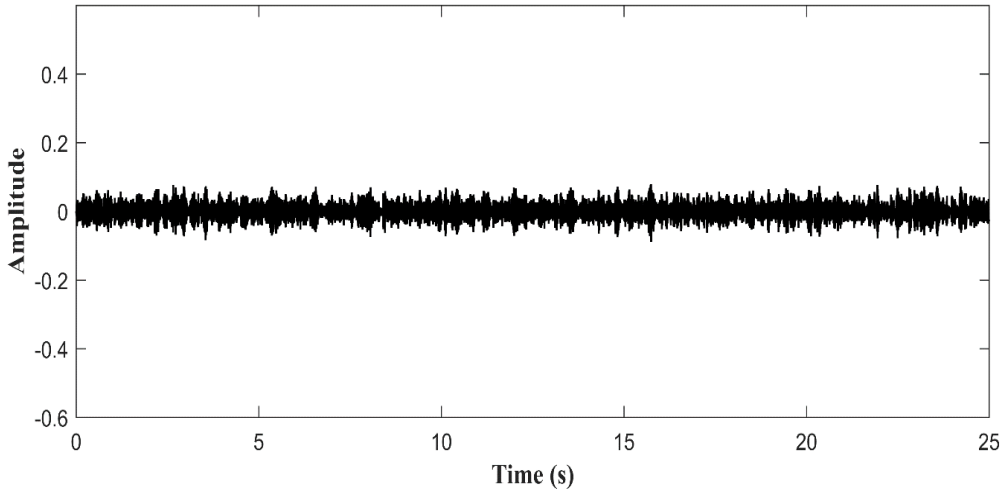


Figure 7 Time series of the synthetic breath signal with same long-term spectrum as breath signal.

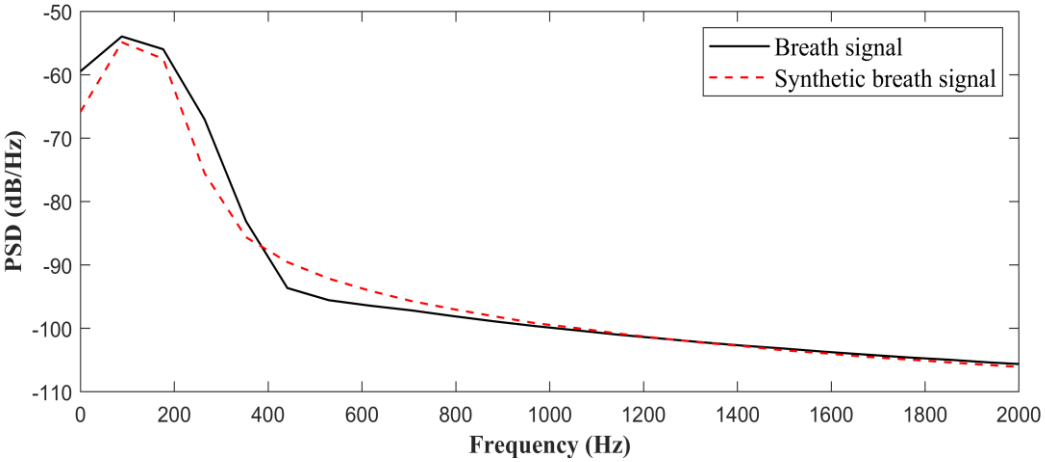


Figure 8 Power spectral density of the breath signal and synthetic breath signal.

Table 2 Summary of the test dataset

Cases	**NOC**	D_g	**IDW & 2CD (ms)**	**BN**	**SNR**
SFC	10	NA	0.7&5 (Charbonneau et al., 2000)	BR_N	-10 to 10 dB
				W_N	
			0.5&3.3 (Hoevers & Loudon, 1990)	BR_N	
				W_N	
			0.9&6 (Cohen, 1990)	BR_N	
				W_N	
SCC	10	NA	1.5&10 (Charbonneau et al., 2000)	BR_N	
				W_N	
			1&5.1 (Hoevers & Loudon, 1990)	BR_N	
				W_N	
			1.25&9.5 (Cohen, 1990)	BR_N	
				W_N	
RFC	10	IPF	ND	BR_N	
				W_N	
RCC	10	BE	ND	BR_N	
				W_N	
RBFC	ND	IPF	ND	NBS	ND
RBCC	ND	BE	ND	NBS	ND

SFC: Simulated fine crackles; SCC: Simulated coarse crackles; RFC: Real fine crackles; RCC: Real coarse crackles; RBFC: Real breath sound with fine crackles; RBCC: Real breath sound with coarse crackles; NOC: Number of crackles; D_g: Diagnosis; BN: Background noise; SNR: Signal to noise ratio; ND: Not defined; NA=Not applicable; IPF: Idiopathic pulmonary fibrosis; BE: Bronchiectasis; BR_N: Breath noise; W_N: Gaussian white noise; NBS: Normal breath sound.

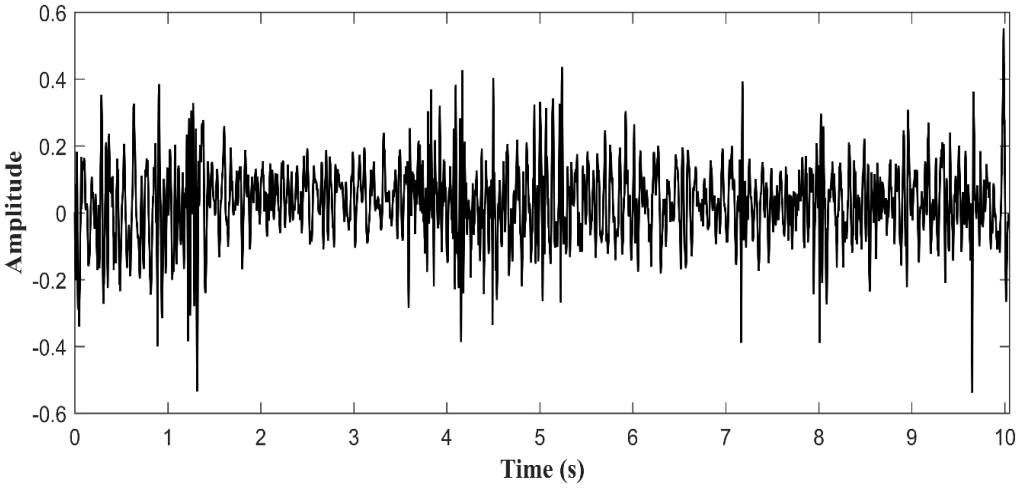


Figure 9 Real breath sound recorded from a patient with IPF.

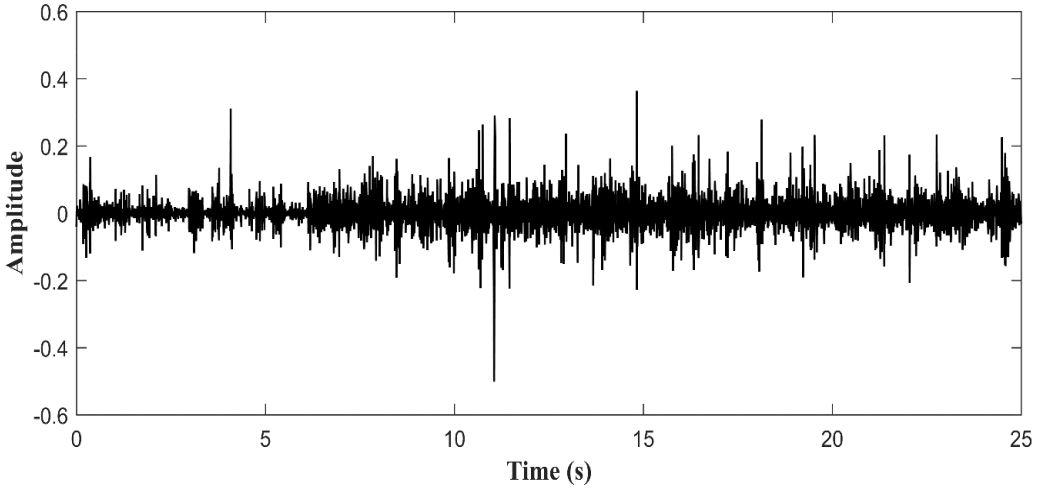


Figure 10 Real breath sound recorded from a patient with BE.

3.3 Summary

This chapter presented a dataset for systematic testing of crackle separation techniques. The next chapter will discuss the details of the best of the currently available methods, before an enhanced scheme is proposed.

Chapter 4 Previous crackle separation techniques

4.1 Introduction

Three crackle separation techniques: WTST-NST filter (Hadjileontiadis & Panas, 1997), WT-FD filter (Hadjileontiadis, 2005(I), Hadjileontiadis, 2005(II)), and EMD-FD filter (Hadjileontiadis, 2007) were selected from the literature for comparing the crackle separation performance of the IEM-FD filter. This chapter will provide the reason for selecting these three previously published crackle separation techniques and detail how they work.

4.2 Selection of previous crackle separation techniques for comparison

As discussed in Chapter 2, in the literature, many different crackle separation techniques have been proposed. From different separation techniques, the WTST-NST filter, the WT-FD filter, and the EMD-FD filter are selected for comparing the performance of the proposed IEM-FD filter. Tolias et al., (1998) and Kandilogiannakis & Mastorocostas, (2018), both mentioned that among all the separation methods presented in the literature, the WTST-NST filter obtains best separation results. The WT-FD filter and the EMD-FD filter are also selected for comparison because the idea of the FD is already used in the fields of the wavelet transform and EMD.

The WTST-NST filter separates wavelet transform coefficients corresponding to crackles and normal breath sounds using a threshold value at each decomposition scale (Hadjileontiadis & Panas, 1997). The number of wavelet transform scales $M = log_2(N)$, where *N* is the number of samples in the signal. In the WTST-NST filter appropriate selection of the threshold for separating WT coefficients related to crackles and normal breath sounds is very important for good results. Selecting the threshold too low leads to overestimation and a higher value may lead to underestimation. To overcome the requirement of empirical setting of the WTST-NST filter threshold, the WT-FD filter was presented (Hadjileontiadis, 2005 (I); Hadjileontiadis, 2005 (II)). In the WT-FD filter, firstly the input signal is decomposed into approximation and detail coefficients using the wavelet transform MRD process and after decomposition, wavelet transform coefficients related to crackles and the wavelet transform coefficients corresponding to normal breath sounds are separated using their FD (Hadjileontiadis & Rekanos, 2003). The idea of FD is further extended in the field of EMD with the EMD-FD filter (Hadjileontiadis, 2007). In this filtering method, those IMFs which contain contributions from the crackle part of the signal are selected using an energy criterion. On the selected IMFs, individually the FD technique is used to separate the crackle part from the normal breath sounds. The detailed working process of the FD technique

and all three selected crackle separation techniques (WTST-NST filter, WT-FD filter and EMD-FD filter) are presented below.

4.3 Fractal dimension technique

The FD is a noninteger, fractional dimension, of a geometric object (Esteller et al., 2001). The FD is a common tool in biomedical signal processing for transient detection (Goh et al., 2005). The FD of a time varying signal or waveform is calculated directly in the time domain, where the signal is considered a geometric figure (Esteller et al., 2001), and is a measure of the signal complexity in the time domain (Esteller et al., 1999). The fractal complexity of a signal in the time domain can vary hence the FD can be used to track the location of sudden changes in a time series. The FD technique based on Katz's definition (Katz, 1988) for detecting crackles in the recorded lung sound signals was proposed by (Hadjileontiadis & Rekanos, 2003).

The detailed working process of Katz's definition (Katz, 1988) for estimating the FD of a curve defined by sequence as follows.

$$F_D = \frac{\log_{10}(n_{stp})}{\log_{10}\frac{d}{L_c} + \log_{10}(n_{stp})} \tag{4}$$

where L_c represents the total length of the curve defined as the sum of the distances between successive points and d is the diameter of the curve (Sevcik, 2010):

$$L_c = \sum_{i=1}^{N-1} dist(i, i+1) \tag{5}$$

$$d = \max[dist(i,j)]\,;\ i \neq j;\ i,j \in [1, N\,] \tag{6}$$

Usually, for the curves that do not cross themselves, the diameter, *d*, of the waveform is evaluated as the distance between the first point of the sequence and the point of the sequence that gives the greatest Euclidean distance (Hadjileontiadis, 2005 (I)):

$$d = \max[dist(1,i)]\,;\ i \in [2, N\,] \tag{7}$$

where $dist(i,j)$, the Euclidean distance between two points i, j is calculated by

$$dist(i,j) = \sqrt{\left(y_i - y_j\right)^2 + \left(n_i - n_j\right)^2} \tag{8}$$

with n_i and n_j the values of the abscissa (i.e. sample numbers) at points *i* and *j,* respectively and y_i and y_j the values of the ordinate (i.e. signal amplitude) corresponding to n_i and n_j, respectively.

and n_{stp} the number of steps in the curve, defined as

$$n_{stp} = \frac{L_c}{\bar{a}} \tag{9}$$

where $\bar{a}$ is the average distance between successive points such that:

$$\bar{a} = \frac{L_c}{N-1} \tag{10}$$

The FD of the signal is estimated using a sliding window, the length of the FD window is calculated $W_{FD} = int(W \times f_s)$, where f_s is the sampling frequency of the signal, 'int' is the operator rounding to the nearest integer and W is the duration of the window in seconds. For signals used in this thesis, the sampling frequency of the input signal is $f_s = 44{,}100\ Hz$ and the duration of the window is 6 ms (Hadjileontiadis & Rekanos, 2003), so $W_{FD} = 264$. The effect of the window length on the performance of the FD technique is discussed in section 4.3.1.

In order to estimate the point-to-point value of the FD, a sliding window is shifted by one sample at a time along the signal and the estimated value of the FD for each segment of the input signal is assigned to the midpoint of the sliding window. In this way, we get the FD vector of length $N - W_{FD} + 1$ for an input signal of length N. The difference in length between the input signal and estimated FD is $W_{FD} - 1$, so to make the estimated FD length equal to the length of the input signal, the first value of the fractal dimension $F_D(1)$ and last value of the fractal dimension $F_D(N - W_{FD} + 1)$ are assigned to the first and last half of the missing value of length $W_{FD} - 1$, respectively. Equation (11) is used to make the minimum value of the F_D always equal to '1'.

$$F_D = F_D - mF_D + 1 \tag{11}$$

where F_D is the FD estimate of the waveform using Eq. (4) and mF_D is the minimum value of F_D.

After estimating the FD of the input signal, the iterative FDPP algorithm (Hadjileontiadis & Rekanos, 2003) is used to automatically identify the location and duration of the FD peaks. The iterative FDPP algorithm not only searches for the high peaks of the FD vector, which may correspond to the high amplitude crackles but also look for the low peaks within the FD vector, which may correspond to the small amplitude crackles (Hadjileontiadis & Rekanos, 2003). The working process of the FDPP algorithm is described in the following steps. In this method, first the standard deviation of the estimated FD is calculated.

$$SD_l = \sigma\left(F_{D_l}\right), \tag{12}$$

where l is the iteration index i.e. $l = 1,2,3 \ldots L$ and SD_l is the standard deviation of the FD estimate at iteration l.

The standard deviation at each iteration, SD_l is compared with each coefficient of the $F_{D_l}(n)$ to estimate the number of peaks related to the sudden occurrence of transient signals such as crackles. If $F_{D_l}(n) > 1 + SD_l$, peak fractal dimension $P_{FD_l}(n) = F_{D_l}(n)$(valid peaks), else $P_{FD_l}(n) = 1.0$ (no peaks). Now the peak peeling process is started in which peak fractal dimension coefficients are subtracted from FD coefficients and using the residual signal $(z_l(n))$ the stopping criterion is calculated:

$$z_l(n) = F_{D_l}(n) - P_{FD_l}(n) + 1 \tag{13}$$

$$STC3_l = |\overline{(z_{l-1}^2(n))} - \overline{(z_l^2(n))}| \tag{14}$$

where $(\bar{.})$ denotes the average value, $|.|$ represents the absolute value, and $z_0 = 0$.

In the last step, the stopping criterion, $STC3$, is compared with constant β_3 (accuracy level).

If $STC3_l \geq \beta_3$, fractal dimension $F_{D_{l+1}}(n) = z_l(n)$ until $l = L$ and end the loop.

$$F_{DPP}(n) = \sum_{l=1}^{L} P_{FD_l}(n) - (L - 1) \tag{15}$$

where the accuracy level β_3 is between the 0 and 1 i.e. $0 < \beta_3 < 1$.

For the better understanding of the process, a schematic representation of the FDPP algorithm is shown in Figure 11 and a worked example of the FD technique is shown in Figure 12. Figure 12 (a) displays a section of the lung sound signal (0.743 s) with coarse crackles recorded from a BE patient (Chapter 3, Table 2 case RBCC). The location of the crackles has been audio-visually identified by an experienced pulmonary acoustics researcher and marked with arrowheads. Note that this example will be used repeatedly as a typical example. From the Figure 12 (b-c) it is clear that the FD technique not only tracks the location and time duration of the high amplitude crackles but also identifies the small amplitude crackles. Although, the FD technique can detect crackles in the recorded lung sounds, it cannot separate them from normal breath sounds. In the FD technique, the selection of window length (W_{FD}) is critical for its overall performance. The effect of window length on performance is discussed below.

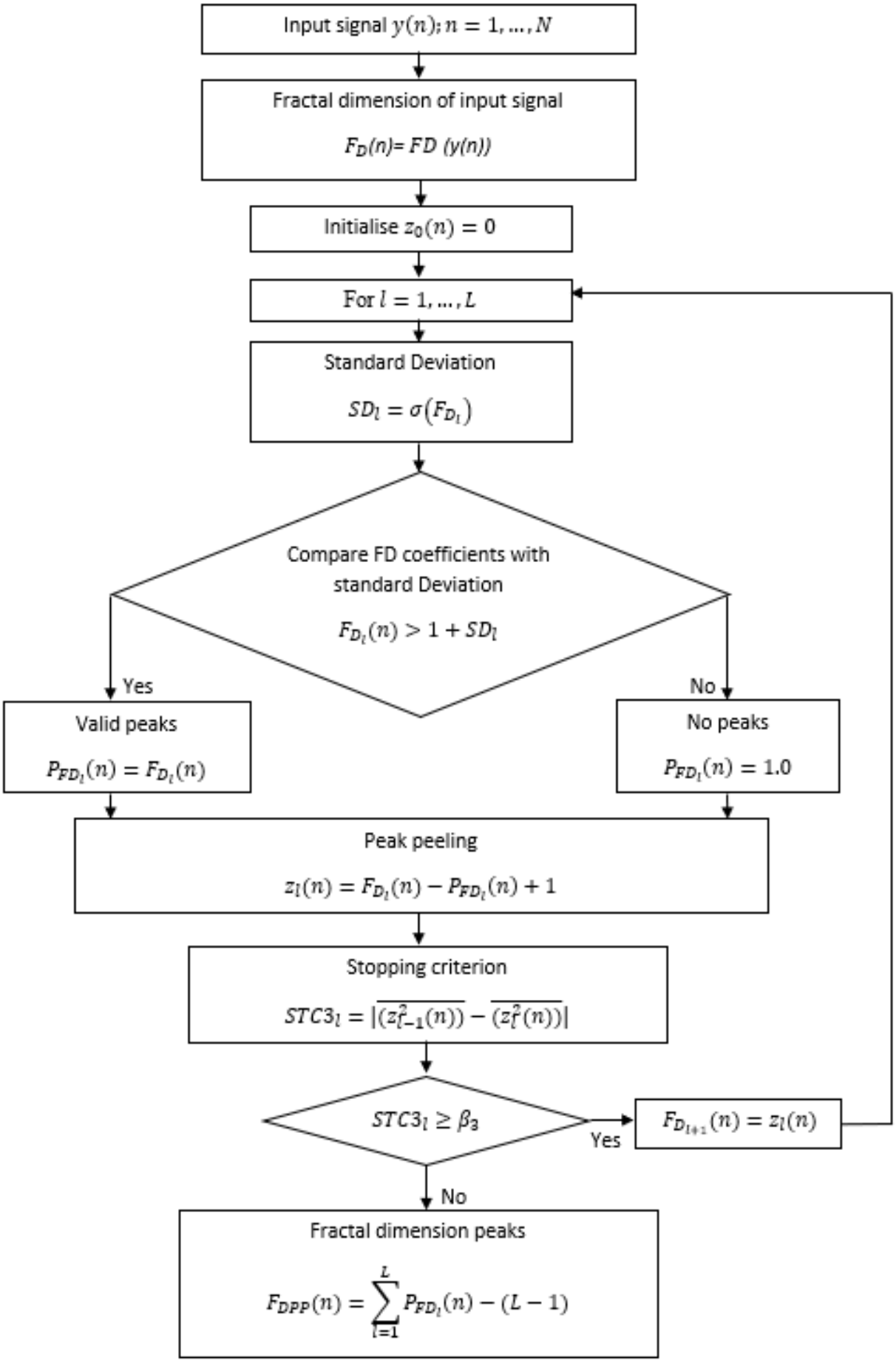
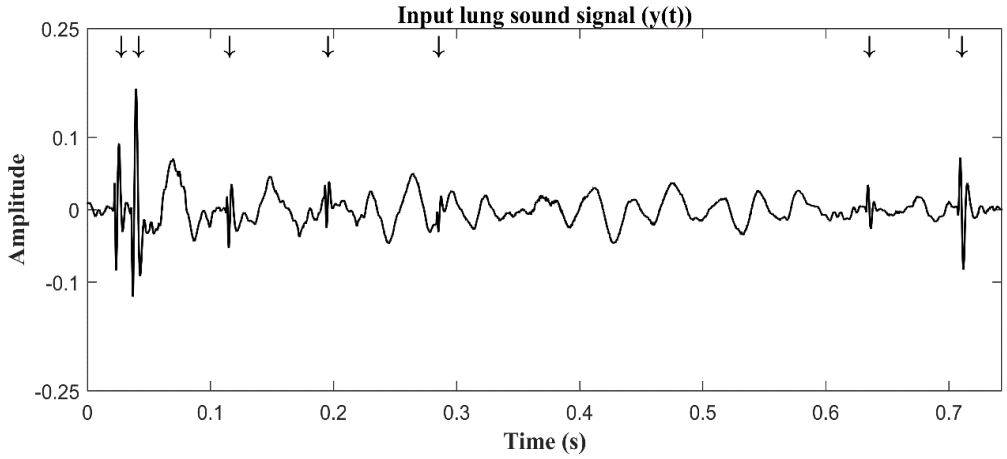
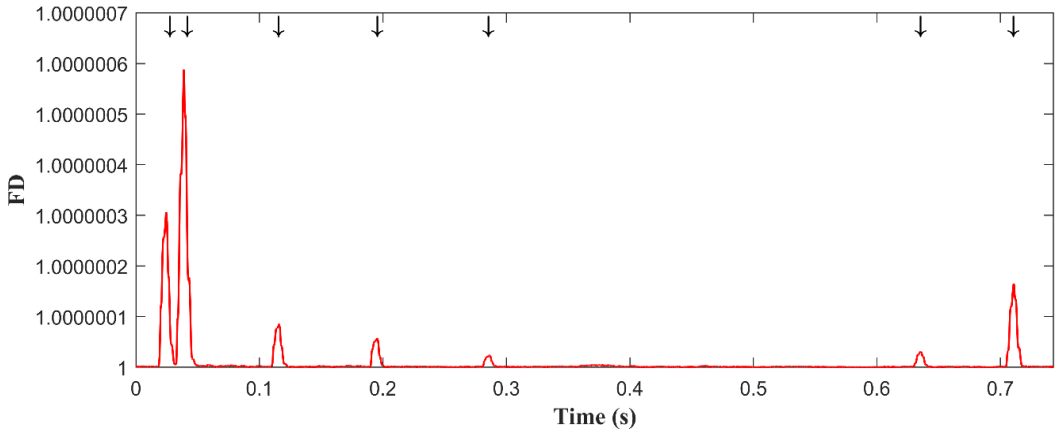


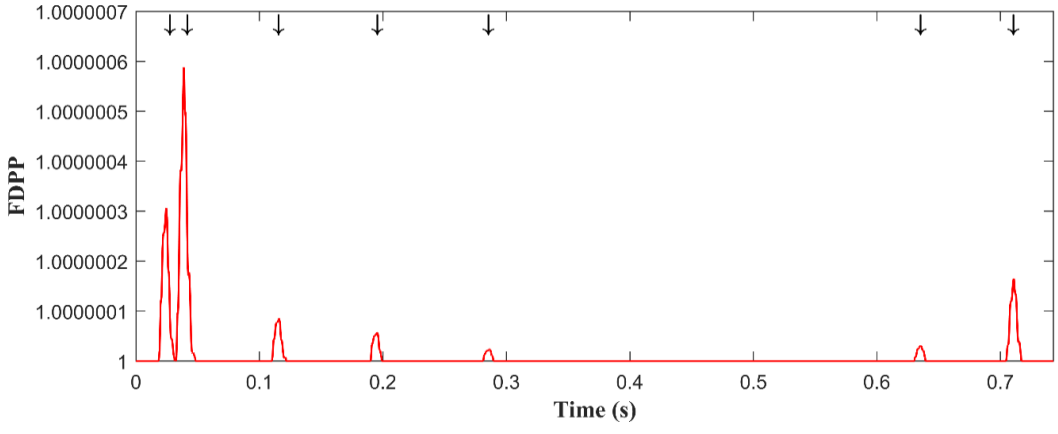
Figure 11 A schematic diagram of the FDPP algorithm.



(a)

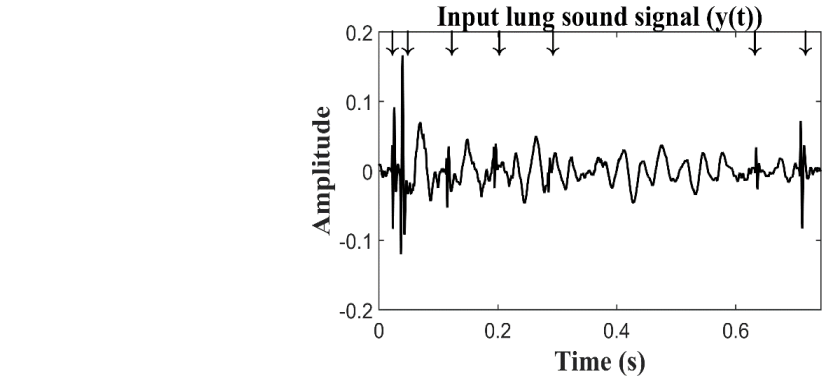


(b)

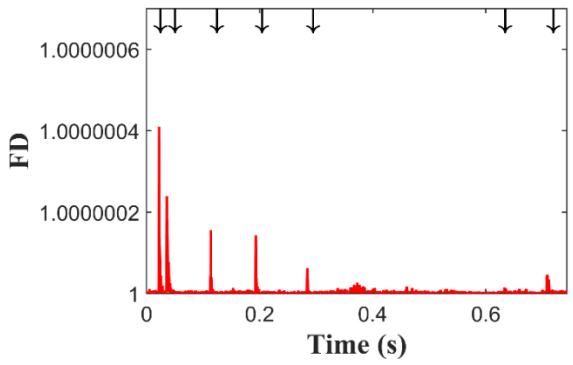


(c)

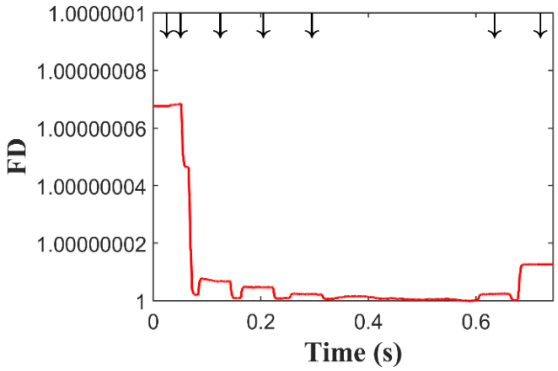
Figure 12 Result of applying the FD technique (a) A time section of 0.743 s lung sound recorded from a BE patient (Chapter 3, Table 2 case RBCC), (b) FD estimated using the Katz's definition (c) output of the FDPP algorithm.



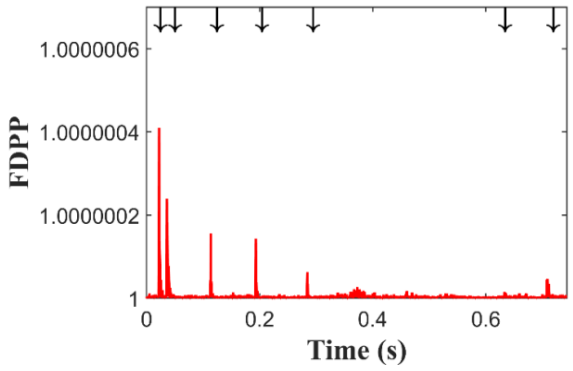
(a)



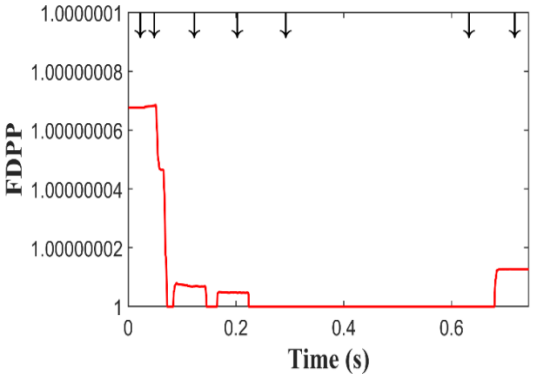
(b)



(d)



(c)



(e)

Figure 13 The FD technique response. (a) A time section of 0.743 s lung sound recorded from a BE patient (Chapter 3, Table 2 case RBCC). When window length is low i.e. W_{FD} = 26 (W_{FD}=int(0.0006f_s) i.e. 0.6 ms); (b) FD estimated using the Katz's definition and (c) output of the FDPP algorithm. When window length is high i.e. W_{FD} = 2646 (W_{FD}=int(0.06f_s) i.e. 60 ms); (d) FD estimated using the Katz's definition and (e) output of the FDPP algorithm.

4.3.1 Effect of fractal dimension window length

In the FD technique, the selection of the appropriate window length is very important. Figure 13 (a) indicate an input lung sound signal with coarse crackles (marked with arrowheads). Figure 13 (b-c) show the response of the FD technique when window length is low (e.g. $W_{FD} = int(0.0006f_s)$=26, i.e. 0.6 ms): (b) estimated FD of the input lung sound signal, (c) output of the FDPP algorithm; and Figure 13 (d-e) show the result of the FD technique when the window length is too high (e.g. $W_{FD} = int(0.06f_s)$ =2646, i.e. 60 ms): (d) estimated FD of the input lung sound signal, (e) output of the FDPP algorithm. In these figures, we can notice that if the window length is low, too many false FD peaks are generated and for the higher value of the window length the FD time series is smoothed.

4.4 Wavelet transform stationary non-stationary (WTST-NST) filter

Hadjileontiadis & Panas, (1997) proposed the wavelet transform based crackle separation technique (WTST-NST) filter. The WTST-NST filtering method separates stationary from non-stationary parts of a signal and hence can be used to separate crackles from normal breath sounds. Lung sounds can be divided into stationary and non-stationary parts where normal breath sounds are generally stationary and crackles non-stationary. The idea behind this filtering method is that explosive peaks in the time domain (crackles) have large components over many wavelet scales, but most components related to background noise (normal breath sounds) reduce with increasing wavelet scale (Hadjileontiadis & Panas, 1997). This fact permits the separation of wavelet transform coefficients with respect to their amplitude related to crackles and normal breath sounds at each wavelet scale using some threshold value. At each wavelet scale the threshold value is based on the standard deviation of the wavelet transform coefficients and an empirical multiplication factor. In this technique, an iterative MRD-MRR scheme are applied to the input signal. A schematic diagram of the WTST-NST filter is shown in Figure 14. In this filtering method firstly, the input signal is decomposed into multiple scales of approximation and detail coefficient vectors using an MRD scheme. The wavelet transform is calculated using the db4 quadrature mirror filters (QMF's) of eight coefficients (Daubechies, 1988; Hadjileontiadis & Panas, 1997).

$$WT_k^m = MRD[y_k(n)]_{m\ scales}\ ;\ n = 1,2,\dots.N; \quad m = 1,\dots,M, M = log_2(N) \tag{16}$$

Where k is the iteration index i.e. $k = 1,2,3 \dots K$, m is the WT scale index, N is the number of samples, and $y_k(n)$ is the input signal at iteration k. In the MRD firstly, the input signal is convolved with the mother wavelet (Daubechies-4 wavelet is taken as the mother wavelet) low

pass filter and high pass filter coefficients followed by dyadic decimation (down sampling:-keep the even indexed elements) which gives an approximation coefficients vector and a detail coefficients vector, respectively. In the next scale, the approximation coefficients vector is further decomposed in approximation and detail coefficient vectors. In this algorithm, number of samples *N*= 32,768 i.e. $M = 15$. The length of the low pass and high pass filters is *B* (for Daubechies-4 wavelet, filter order *u=4* and filter length *B=2u=8*). The low pass and high pass filters in WT are known as quadrature mirror filters and Eq. (17) shows the relationship between the coefficients of the MRD low pass filter (*h*(*b*)) and high pass filter (*g*(*b*)); the correlation between the MRR low pass filter ($h_1(b)$) and high pass filter ($g_1(b)$) is shown in Eq. (18).

$$g(B-1-b) = (-1)^b . h(b) \tag{17}$$

$$g_1(b) = (-1)^b . h_1(B-b-1) \tag{18}$$

where b denotes the filter length index i.e. $b = 0,1,\dots,B-1$. In the wavelet transform, only the approximation coefficients vector is decomposed at every scale and successive detail coefficients are never reanalyzed. Figure 15 shows the MRD for WT scales.

After decomposition, a threshold value is applied at each scale to separate WT coefficients related to crackles from the normal breath sounds.

$$Th_k^m = SD_k^m F_{adj} \tag{19}$$

Where SD_k^m is the standard deviation of the wavelet transform coefficients at iteration *k* and scale *m*, and F_{adj} is the adjusting multiplication factor. Here, $F_{adj} = 3$ (Hadjileontiadis & Panas, 1997) is used. The wavelet transform coefficients $WT_k^m > Th_k^m$ are related to crackles and the remaining wavelet coefficients are related to normal breath sounds. After separating the coefficients, using a threshold value, reconstruction of the non-stationary and stationary parts $NSTS_k(n)$ and $STS_k(n)$, respectively is made using the MRR scheme. In the MRR scheme, firstly, approximation and detail coefficients vectors at scale *m,* are up-sampled (insert zeros at even indexed elements) and the resulting signals are convolved with the Daubechies-4 wavelet reconstruction low pass filter and high pass filter. After convolution, the centre part of the obtained signal with length equal to the length of scale *m-1* detail coefficients vector is taken which provides scale *m-1* approximation coefficients vector. In the wavelet transform MRR process at each scale the approximation coefficients vector and the detail coefficients vector are used to obtain the previous scale approximation coefficients vector and at last i.e. scale *M-(M-1)* approximation coefficients vector and detail coefficients vector provide input signal. Note that here at each scale wavelet transform coefficients are separated into those related to the crackles

(WTC_k^m) and those related to the normal breath sounds ($WTNB_k^m$). MRR is applied to both the wavelet transform coefficients related to crackles and the wavelet transform coefficients related to normal breath sounds to create estimates of the non-stationary ($NSTS_k(n)$) and stationary ($STS_k(n)$) parts of the input signal. Figure 16 shows the MRR for wavelet transform coefficients related to crackles and the MRR for wavelet transform coefficients related to normal breath sounds is shown in Figure 17. The iteration procedure stops after satisfying the stopping criterion:

$$STC2_k = |\overline{(STS_{k-1}^2(n))} - \overline{(STS_k^2(n))}| < \beta_2, \quad 0 < \beta_2 < 1 \tag{20}$$

where β_2 is the accuracy level and $STS_0 = 0$.

After the last iteration K, STS_K provides the stationary part and the summation of estimated non- stationary parts separated at each iteration gives the total non-stationary part of the signal $NS(n)$.

$$NS(n) = \sum_{k=1}^{K} NSTS_k\,(n) \tag{21}$$

$$ST(n) = STS_K(n) \tag{22}$$

A worked example of the WTST-NST filter is shown in Figure 18.

Figure 18 (a) shows an input lung sound signal with coarse crackles (marked with arrowheads). The non-stationary ($NS(n)$) and stationary ($ST(n)$) parts after applying the WTST-NST filter are shown in Figure 18 (b) and (c), respectively. Comparing these results with the input signal we can notice that all the crackles are separated into the non-stationary part. However, it is important to notice that the non-stationary part not only consists the crackles but also contains some portion of the background noise (normal breath sounds) or residue of the normal breath sounds, which shows overestimation. For any crackle separation technique, it is not only important to separate crackles in the non-stationary part, but it is also important to minimize the non-crackle contribution to the non-stationary output. Estimation of the number of crackles and their time domain features (IDW, LDW, 2CD) can be more accurate following crackle separation with minimal over or under estimation.

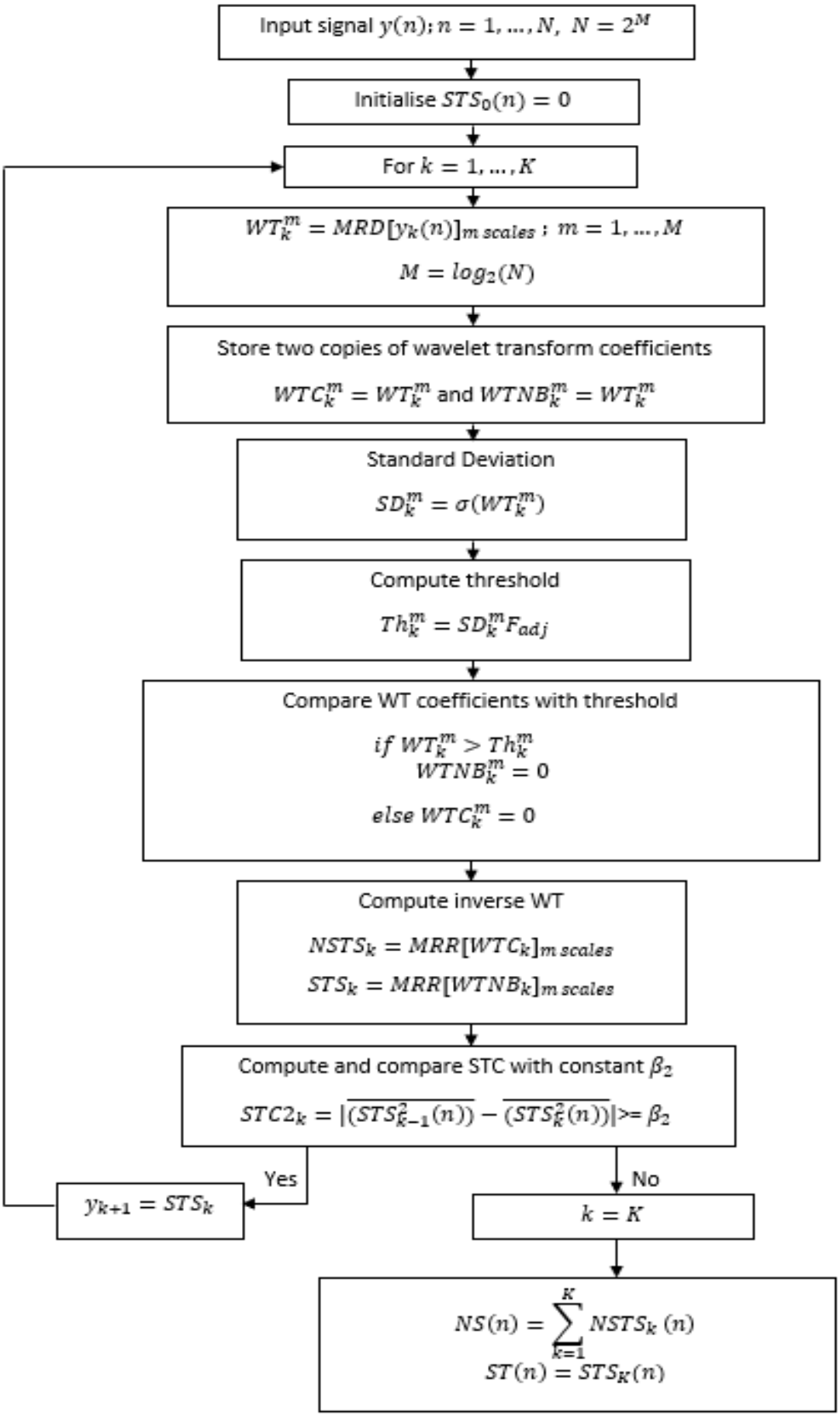


Figure 14 A schematic diagram of the WTST-NST filter.

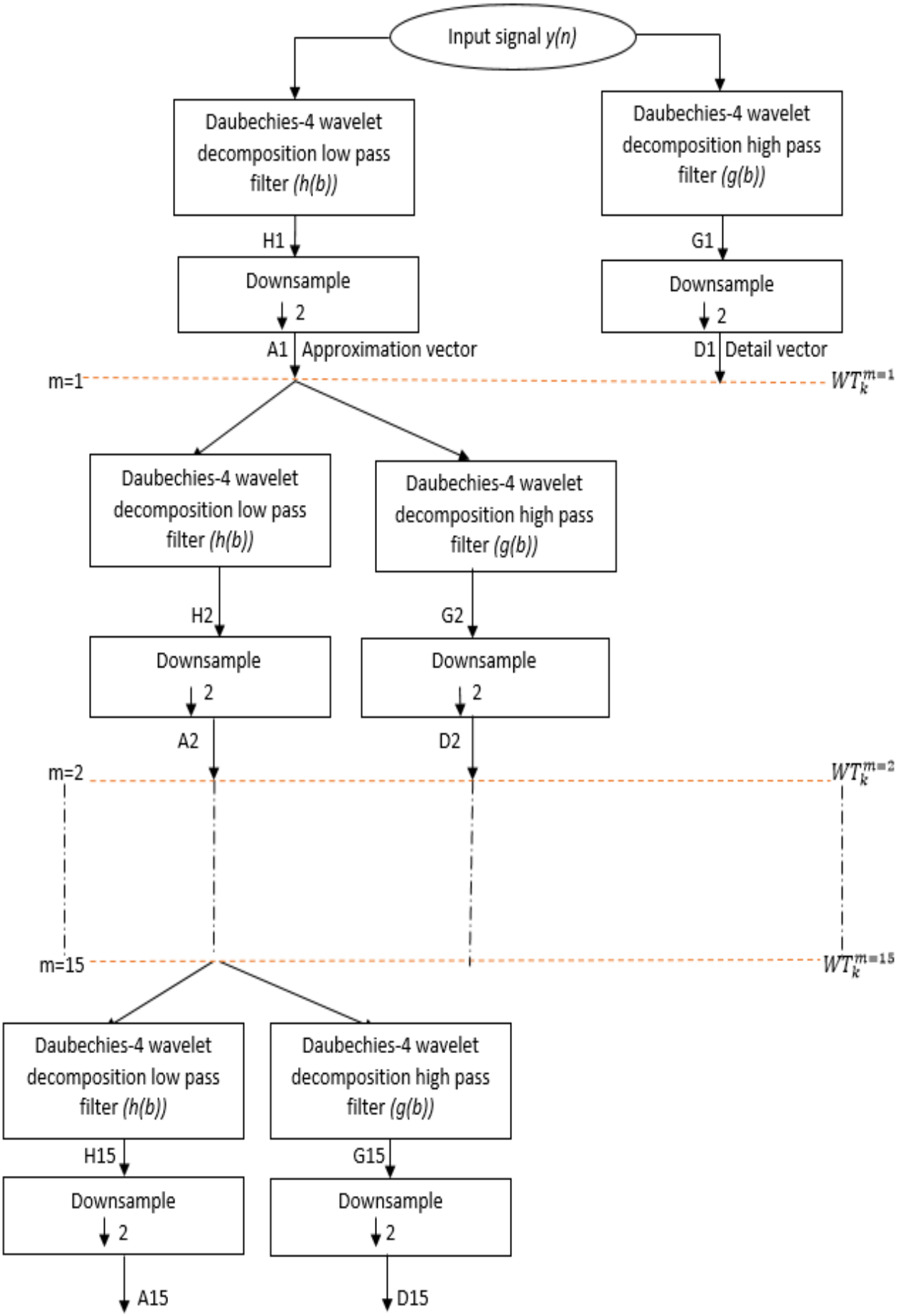


Figure 15 Wavelet transform MRD of scale $m = 1,2, \dots, M,\ M = 15$.

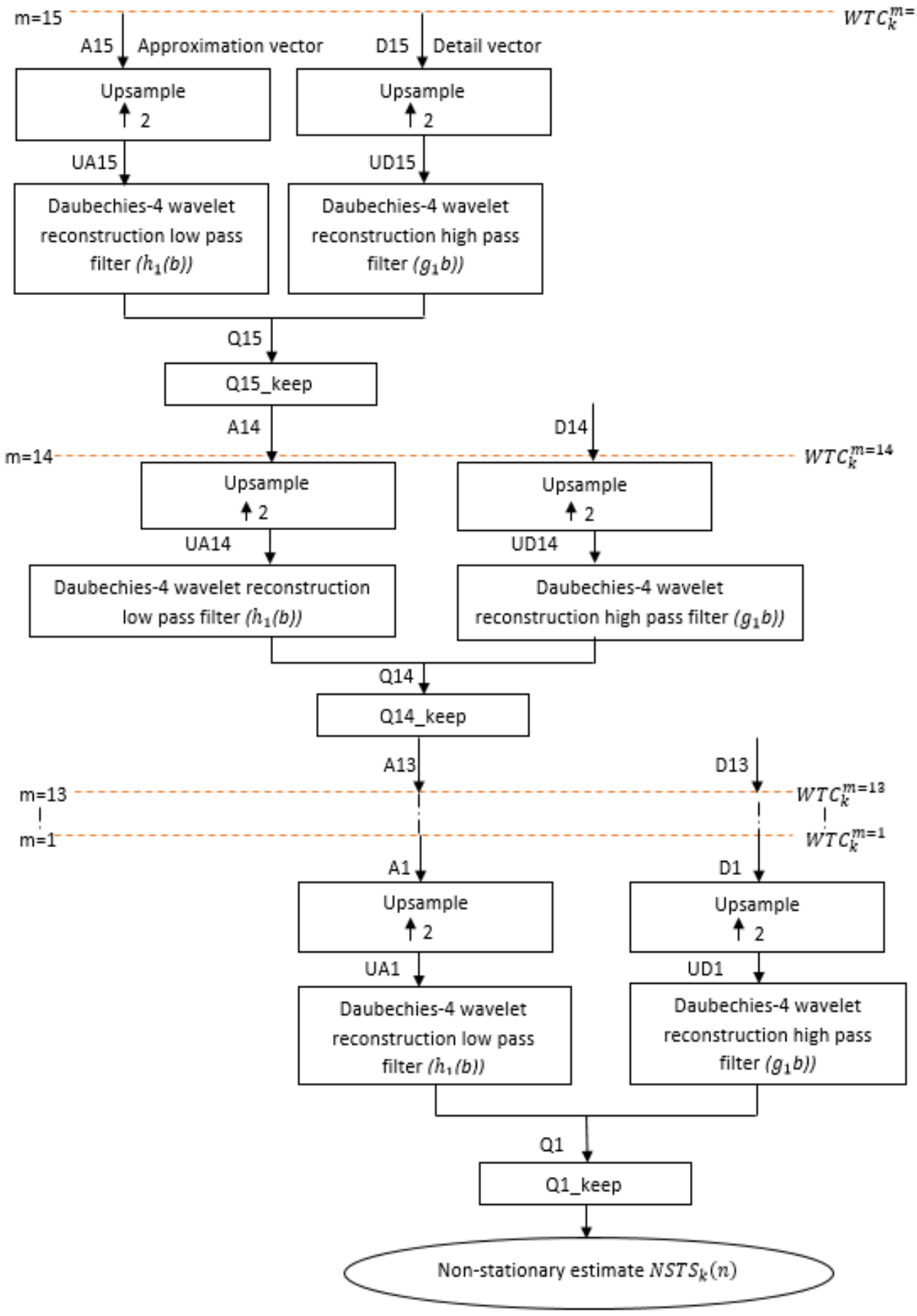


Figure 16 MRR for wavelet transform coefficients related to crackles (WTC_k^m) scales $m = 1,2,\dots,M,\ M = 15$.

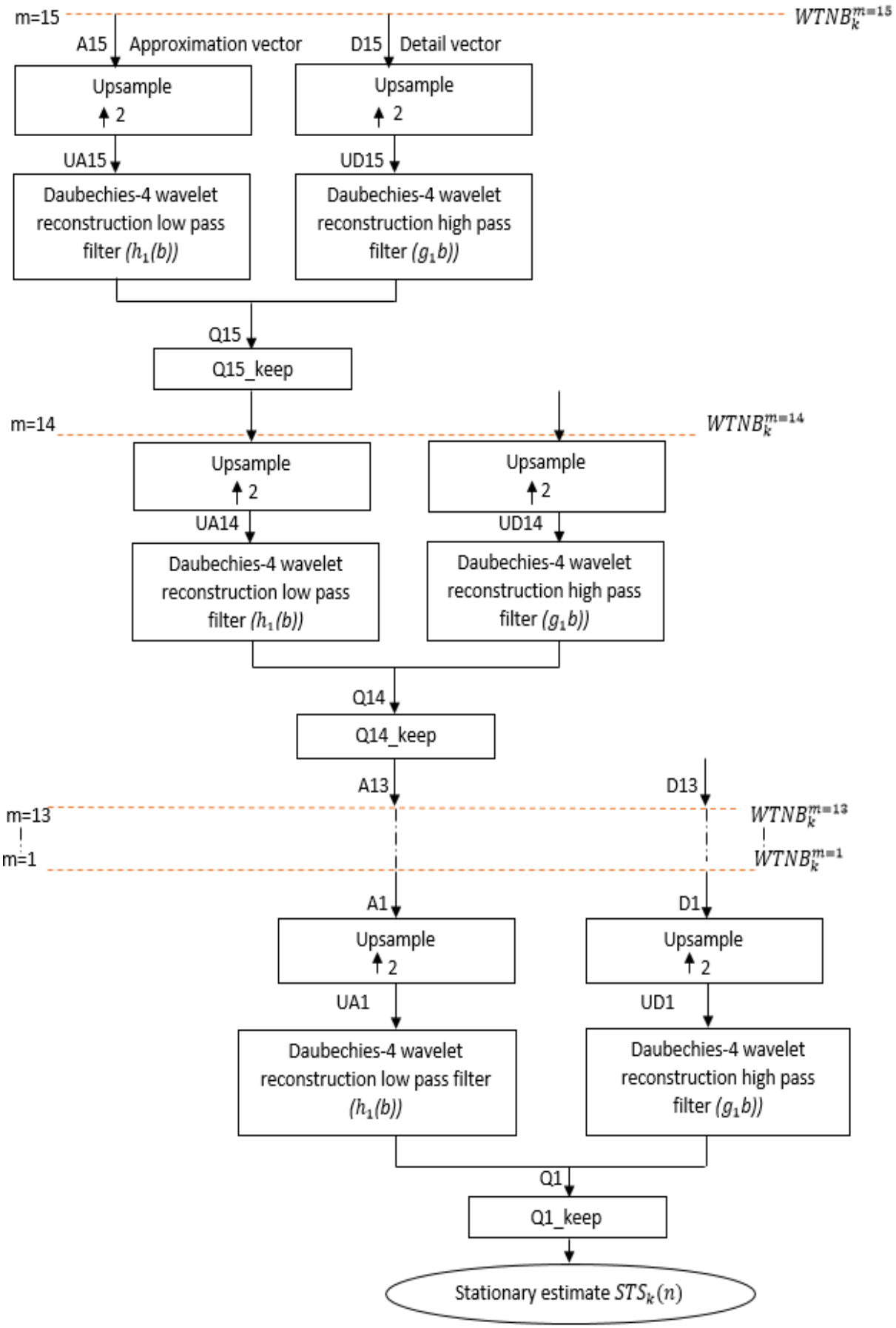
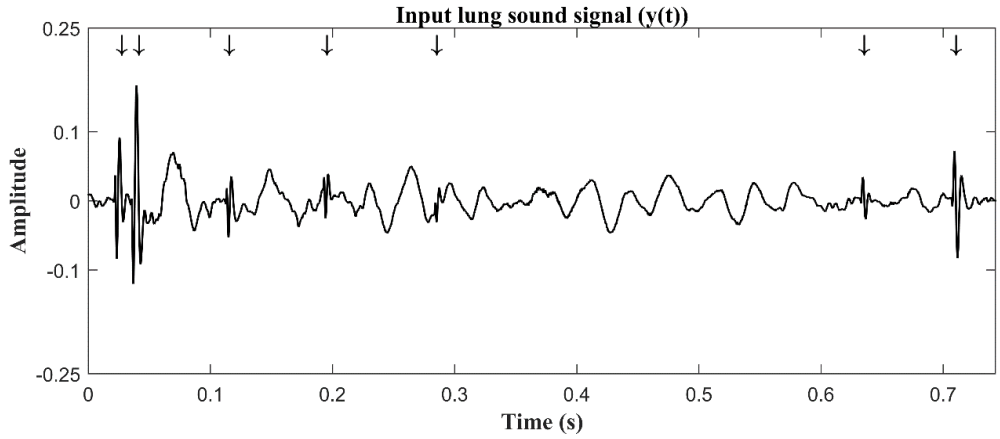
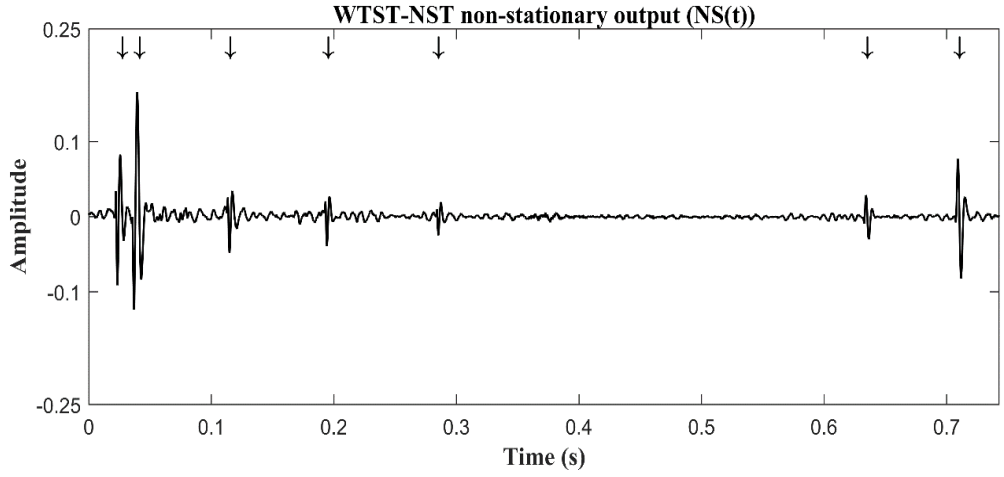


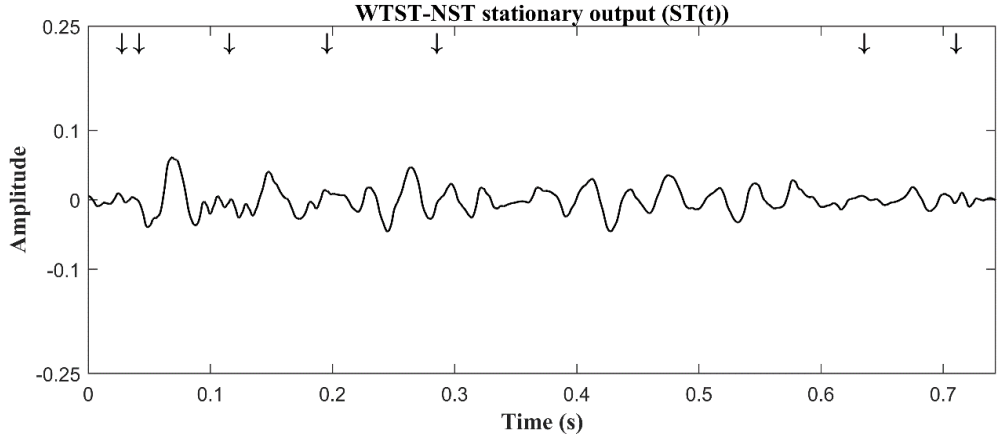
Figure 17 MRR for wavelet transform coefficients related to normal breath sounds ($WTNB_k^m$) scales $m = 1,2,\dots,M,\ M = 15$.



(a)

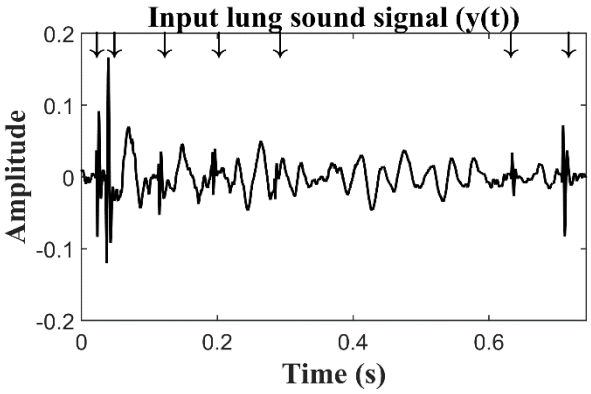


(b)

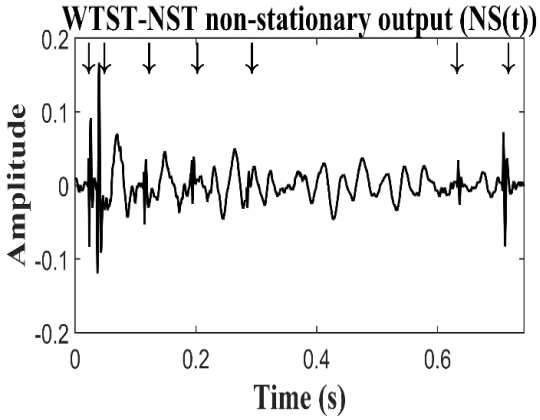


(c)

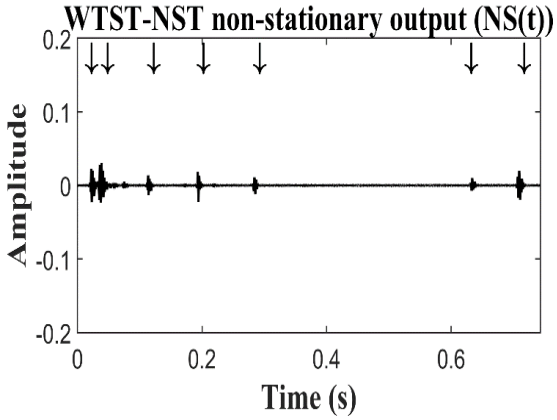
Figure 18 Result of applying the WTST-NST filter (a) A time section of 0.743 s lung sound recorded from a BE patient (Chapter 3, Table 2 case RBCC), (b) WTST-NST filter non-stationary output (c) WTST-NST filter stationary output.



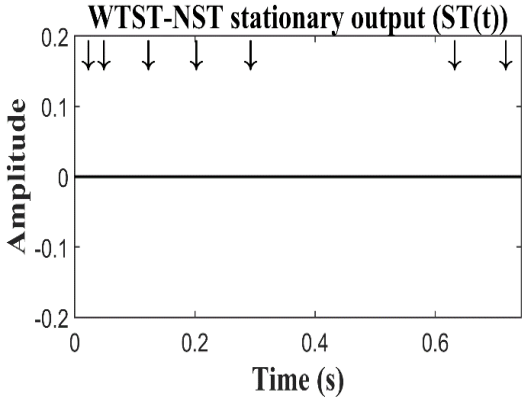
(a)



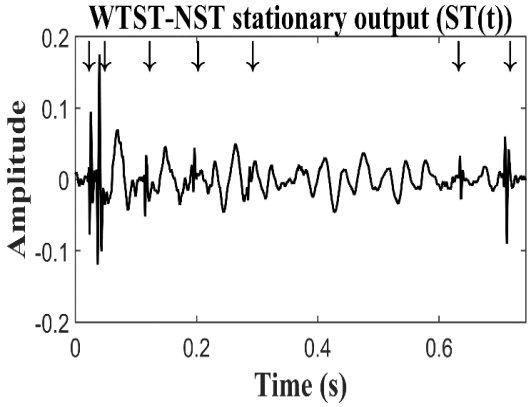
(b)



(d)



(c)



(e)

Figure 19 The WTST-NST filter response. (a) A time section of 0.743 s lung sound recorded from a BE patient (Chapter 3, Table 2 case RBCC). When low threshold ($F_{adj} = 0.3$); (b) WTST-NST filter non-stationary output (c) WTST-NST filter stationary output. When high threshold ($F_{adj} = 8.6$); (d) WTST-NST filter non-stationary output (e) WTST-NST filter stationary output.

4.4.1 Effect of threshold

The WTST-NST filter is a threshold-based method, and the appropriate selection of the threshold is very important for good performance by the WTST-NST filter. Figure 19 (a) shows an input lung sound signal with coarse crackles (marked with arrowheads). From Figure 19 (b-c) it can be noticed that for a low value of the threshold (in this example $F_{adj} = 0.3$), the non-stationary part includes a portion of the normal breath sounds. On the other hand, in Figure 19 (d-e) we can see if the threshold value is high ($F_{adj} = 8.6$), the separation of crackles is not accurate.

4.5 Wavelet transform fractal dimension (WT-FD) filter

To overcome the requirement of empirical setting of a threshold in the WTST-NST filter, the FD technique (Hadjileontiadis & Rekanos, 2003) is applied to the WT domain for automatically separating the WT coefficients related to crackles and to normal breath sounds (Hadjileontiadis, 2005(I); Hadjileontiadis, 2005(II)). In this filtering method firstly, the input signal is decomposed into an approximation coefficients vector and a detail coefficients vector using the MRD process similar to the WTST-NST filter using the db4 QMF's of eight coefficients but here WT scale is used $m = 1$ (Hadjileontiadis, 2005(II)).

$$WT_k^1 = MRD[x(n)]_{m\ scales} \tag{23}$$

After decomposing the input signal, the FD of the wavelet transform coefficients is calculated using a overlapped sliding window 6 ms duration (Hadjileontiadis, 2005(II)) which is shifted one sample along the length input signal for estimating the point to point FD. In this work, the sampling frequency chosen is $f_s = 44{,}100$ Hz.

$$F_{D_k}^m = FD[WT_k^m] \tag{24}$$

Then the FDPP algorithm is applied to the estimated FD.

$$F_{DPP_k}^m = FDPP[F_{D_k}^m] \tag{25}$$

For detail of the working of the FD technique see section 4.3.

After automatically estimating the FD peaks, two binary thresholds are computed one is related to crackle sounds $NBTH_k$ and the other one is related to normal breath sounds $SBTH_k$.

$$NBTH_k^m = \begin{cases} 1\ if\ F_{DPP_k}^m \neq 1 \\ 0\ if\ F_{DPP_k}^m = 1 \end{cases} \tag{26}$$

$$SBTH_k^m = [1 - NBTH_k^m] \tag{27}$$

When both thresholds are constructed they are multiplied by the WT coefficients calculated from the MRD scheme i.e. WT_k^m (see Figure 20), where k is the iteration index i.e. $k = 1,2, \ldots, K$ and m is the WT scale ($m = 1$).

The result of these two multiplications provides the WT coefficients related to crackles i.e. WTC_k^m, and WT coefficients related to normal breath sounds i.e. $WTNB_k^m$.

Now by using a MRR procedure section 4.4 a first version of the non-stationary part i.e. $NSTS_k(n)$, and the stationary part i.e. $STS_k(n)$ are estimated. The iteration procedure stops when it satisfies the stopping criterion:

$$STC2_k = |\overline{(STS_{k-1}^2(n))} - \overline{(STS_k^2(n))}| < \beta_2, \quad 0 < \beta_2 < 1 \tag{28}$$

where $STS_0 = 0$.

At iteration $k = K$, the non-stationary and stationary parts of the input signal are calculated using Eq. (29) and Eq. (30).

$$NS(n) = \sum_{k=1}^{K} NSTS_k(n) \tag{29}$$

$$ST(n) = STS_K(n) \tag{30}$$

A schematic diagram of the WT-FD filter is shown in Figure 20 and the working example of the WT-FD filter is displayed in Figure 21. Figure 21 (a) shows an input lung sound signal with coarse crackles (marked with arrowheads). The non-stationary and stationary parts after applying the WT-FD filter are illustrated in Figure 21 (b) and (c), respectively. Comparing these results with the input lung sound signal it can be observed that all the crackles are separated into the non-stationary part. Here, it is also important to notice that in the stationary output some normal breath sounds segments are missing at the location of crackles. This indicates that the WT-FD filter can separate normal breath sounds from non-crackle locations but it is not able to perfectly separate normal breath sounds at the location of crackles. This effect occurs due to magnitude domination of WT coefficients related to crackles over the WT coefficients corresponding to normal breath sounds (Hadjileontiadis, 2005(II)). For any crackle separation technique it is not only important to separate normal breath sounds from where crackles are not present but it is also important to separate normal breath sounds where crackles are present, to reveal the actual morphology of the crackles.

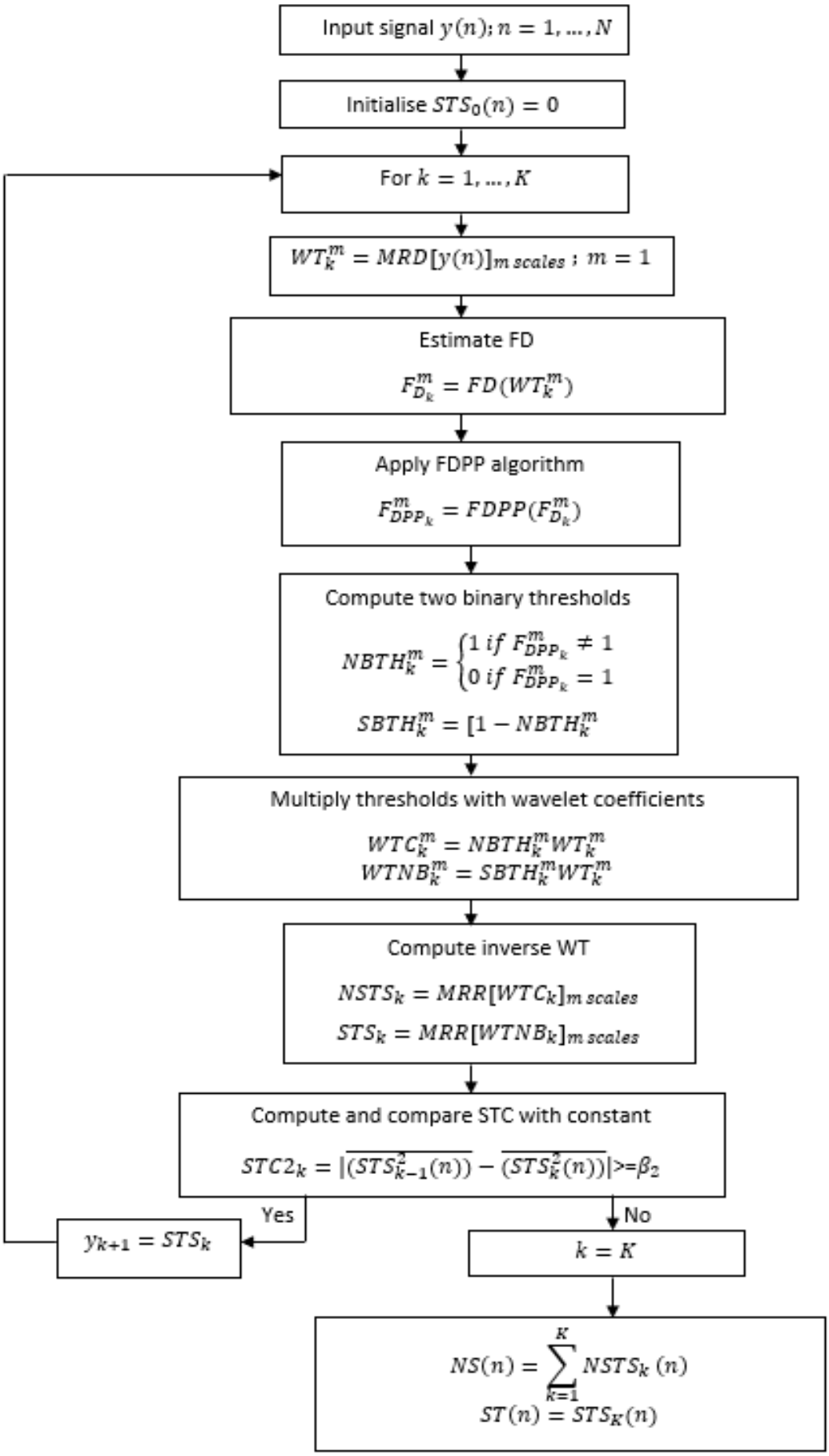
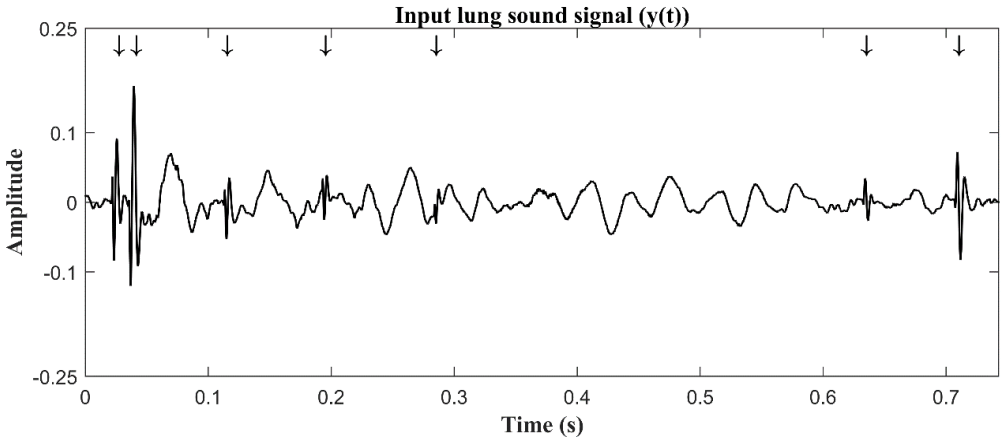
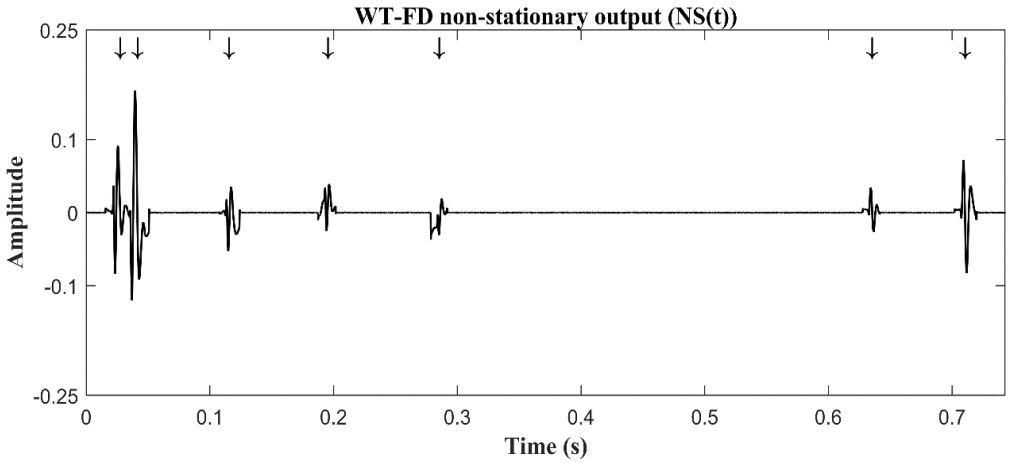


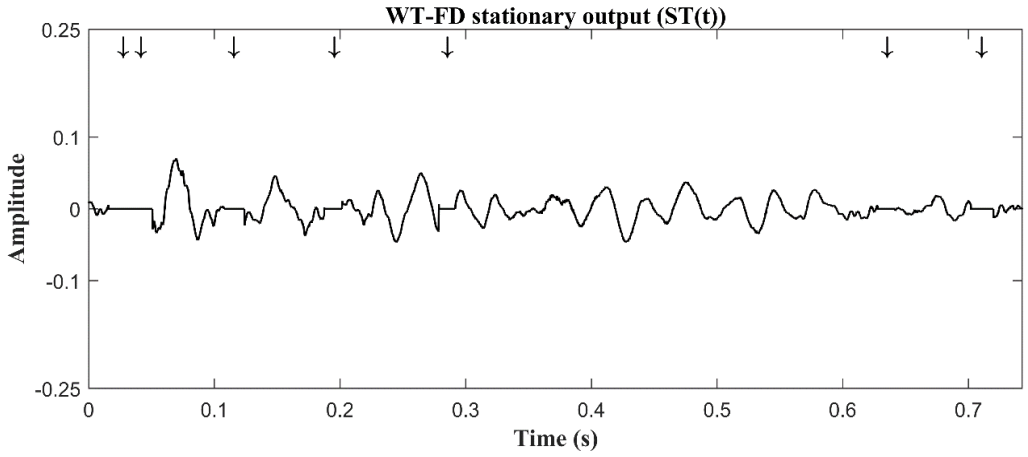
Figure 20 A schematic diagram of the WT-FD filter.



(a)



(b)



(c)

Figure 21 Result of applying the WT-FD filter (a) A time section of 0.743 s lung sound recorded from a BE patient (Chapter 3, Table 2 case RBCC), (b) WT-FD filter non-stationary output (c) WT-FD filter stationary output.

4.6 Empirical mode decomposition fractal dimension (EMD-FD) filter

The crackle identification ability of the FD technique is further applied in the field of EMD, Hadjileontiadis (Hadjileontiadis, 2007), proposed the EMD-FD filter for automatically separating crackles from normal breath sounds. In this section, firstly, the detailed working process of the EMD method (Huang et al., 1998) is presented and then the EMD-FD filter is discussed.

4.6.1 Empirical mode decomposition (EMD)

Huang (Huang et al., 1998) proposed the EMD method for adaptively decomposing a signal into its IMFs and a residual component in decreasing order of frequency. Each estimated IMF follows two conditions: (i) The number of extrema points and the number of zero crossing either be equal or differ at most by one, (ii) At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero (Huang et al., 1998; Omitaomu et al., 2011; Bajaj & Pachori, 2012; Lei et al., 2013; Chauhan et al., 2010; Zhang et. al., 2018) i.e. its envelopes are symmetric with respect to local zero mean. On the other side, the residual component contains only one extremum value. The procedure of extracting IMFs from an input signal using sifting process is explained in following steps.

Step (1):- The decomposition process starts with identifying extrema points i.e. local maxima and local minima of the input signal.

In the input signal ($y(n)$), any sample is a maximum if its preceding slope ($n-1$) is positive and succeeding slope ($n+1$) is negative. A sample is a minimum point if its preceding slope is negative and succeeding slope is positive.

Step (2):- After finding the extrema points of the input signal, all the maxima are connected with each other to estimate an upper envelope ($UP_{env}(n)$) and all the minima are connected to each other to extract a lower envelope ($LW_{env}(n)$) using cubic spline interpolation.

Step (3):- Once, the upper and lower envelopes are estimated their mean is calculated.

$$m(n) = \frac{UP_{env}(n) + LW_{env}(n)}{2} \tag{31}$$

Step (4):- The difference between the input signal ($y(n)$) and mean ($m(n)$) provides the first proto IMF $p_1(n)$.

$$p_1(n) = y(n) - m(n) \tag{32}$$

Step (5):- Now, the proto-IMF is tested with the IMF conditions using the sifting criterion.

$$S_C = \sum_{n=1}^{N} \left[\frac{|(p_{1_{i_1-1}}(n) - p_{1_{i_1}}(n))|^2}{p_{1_{i_1-1}}^2(n)} \right] \quad (33)$$

where i_1 is the number of times the sifting criterion is repeated for searching an IMF. The value of S_C lies between 0.2-0.3 (Huang et al., 1998). If the proto-IMF satisfies the sifting criterion then it is considered as an IMF ($c_1(n)$) otherwise steps 1 to 4 are repeated.

Each IMF satisfies two conditions; Figure 22 (a) shows the first condition of the IMF, in which the number of zero crossings is 14 and the number of extrema (maxima (7) + minima points (6)) is 13 (one less than the number of zero crossings) and Figure 22 (b) shows the second condition of the IMF in which the mean of the point on the upper and lower envelopes is zero i.e. (0.0159 + (-0.0159))/2=0.

Step (6):- Once the first IMF is calculated it is subtracted from the input signal and the resultant output is used as an input for finding the next IMF.

$$r_1(n) = y(n) - c_1(n) \quad (34)$$

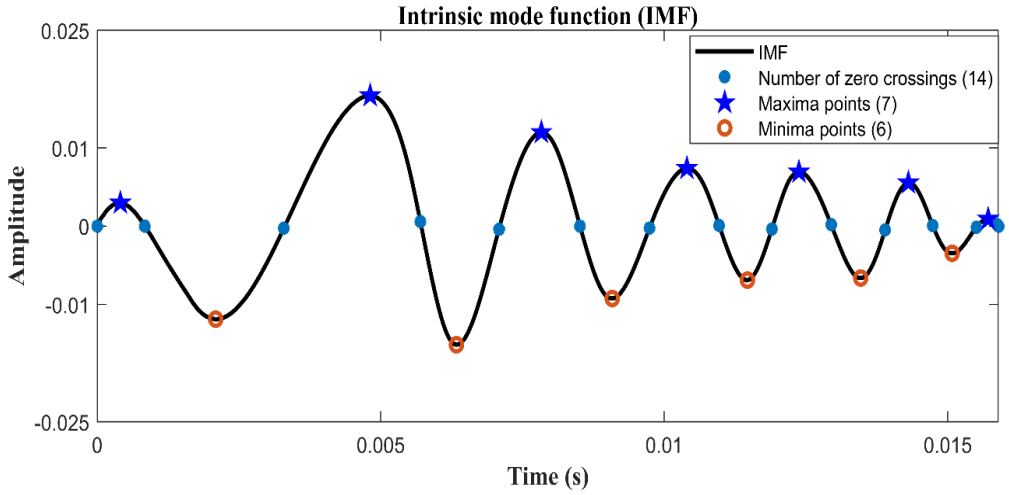
Step (7):- If the signal contains only one extremum or the signal becomes a monotonic function from which no more IMFs can be obtained this signal is assigned as a residual component; otherwise repeat step 1 to 6. Figure 23 shows the residue component which contains only one extremum.

Step (8):- We can estimate the input signal back by adding together all the IMFs and residue component.

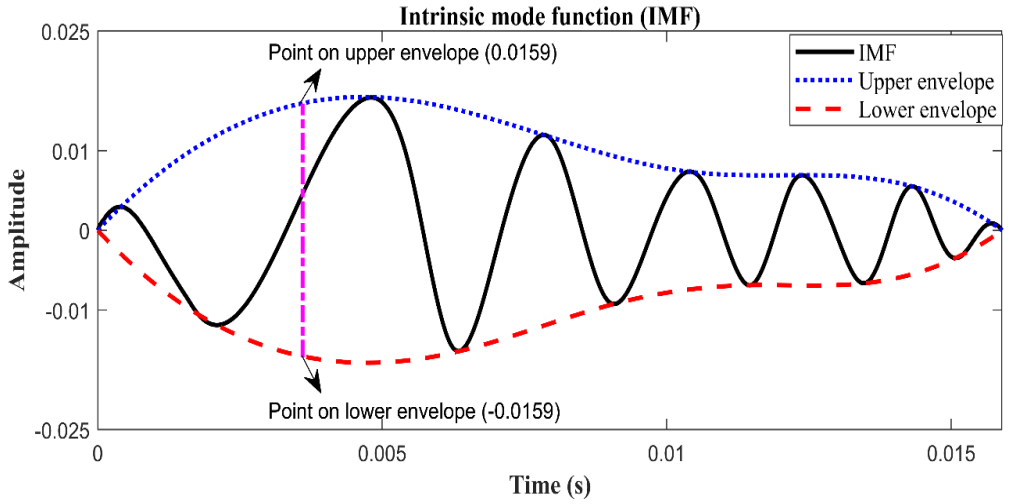
$$y(n) = \sum_{f=1}^{F} c_f(n) + r_F \quad (35)$$

Where F is the number of IMFs and r is the residue component.

The EMD method can decompose a signal into IMFs and a residual component, but it is not sufficient on its own for automatic separation of crackles from the normal breath sounds because the IMFs not only contain the crackle parts but also consist of a portion of the normal breath sounds, as shown in Figure 24, where an input lung sound signal is decomposed into 19 IMFs and a residual component. Figure 24 (a) shows an input lung sound signal with coarse crackles (marked with arrowheads). Figure 24 (b) and (c) show the extracted IMFs (19 IMFs) and a residual component, respectively.



(a)



(b)

Figure 22 IMF conditions: (a) IMF first condition (b) IMF second condition.

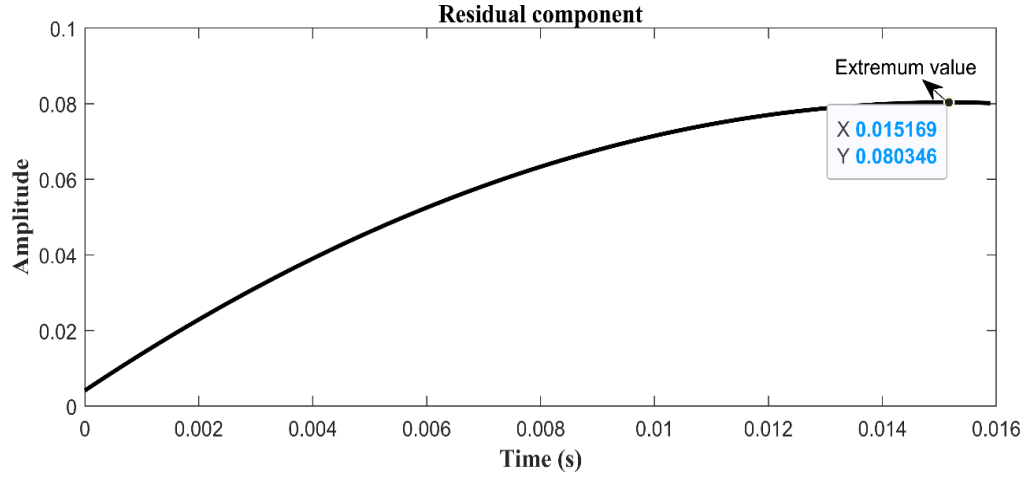


Figure 23 Residue component.

4.6.2 Empirical mode decomposition fractal dimension (EMD-FD) filter

Hadjileontiadis, (2007) presented a combination of the EMD method (Huang et al., 1998) with FD technique (Hadjileontiadis & Rekanos, 2003). The EMD-FD filter adaptively captures the crackles from the IMFs and successfully extracts them from the normal breath sounds. In this filtering method firstly, the input signal is decomposed using the EMD method (see section 4.6.1) and after decomposition, a number of IMFs, V, which are related to crackles are selected using an energy criterion. Now, the FD technique separates the crackles and normal breath sounds from the selected IMFs. Once the normal breath sounds is separated from the selected IMFs, the remaining parts are combined to estimate the crackles. The separated normal breath sounds, a combination of remaining IMFs and a residual component, provides the overall normal breath sounds in the input signal. The working process of the EMD-FD method is described in the following steps:-

Step (1):- First of all, the input signal is decomposed into different IMFs and a residual i.e. $c_f(n)$ and $r_F(n)$ using EMD (Huang et al., 1998).

Step (2):- Now, from the *F* estimated IMFs the first V IMFs are selected using the energy criterion.

$$V = \min\{f: |(\gamma_f')| > p \;\&|(\gamma_{f+1}')| \leq p \;\&\; \gamma_f'' > 0\}, f = 1,2,..,F \tag{36}$$

$$\gamma_f = 1 - \frac{\sum_{a_1=1}^{f}\overline{(c1_{a_1}^2(n))}}{\sum_{a_1=1}^{F}\overline{(c1_{a_1}^2(n))}} \tag{37}$$

where γ_f' and γ_f'' are a first and second derivatives of the γ_f, with respect to f. p is a small threshold value (p=0.05 (Hadjileontiadis, 2007)).

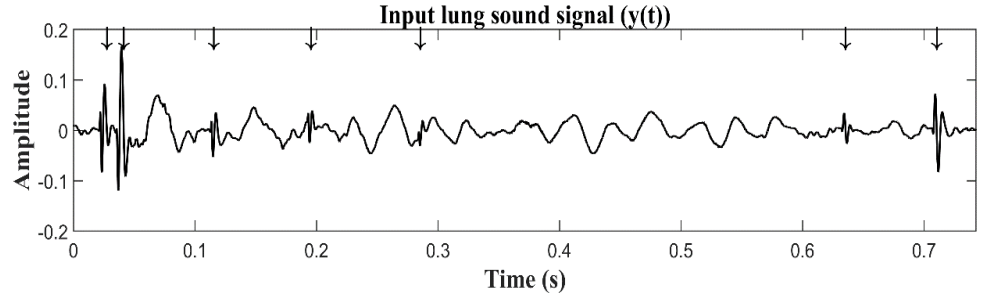
Step (3):- Now the FD technique (see section 4.3) is applied to all *V* selected IMFs and two binary thresholds i.e. $NBTH_v(n)$ and $SBTH_v(n)$ are calculated where $v = 1, \ldots, V$:

$$NBTH_v(n) = \begin{cases} 1 \; if \; F_{DPP_v}(n) \neq 1 \\ 0 \; if \; F_{DPP_v}(n) = 1 \end{cases} \tag{38}$$

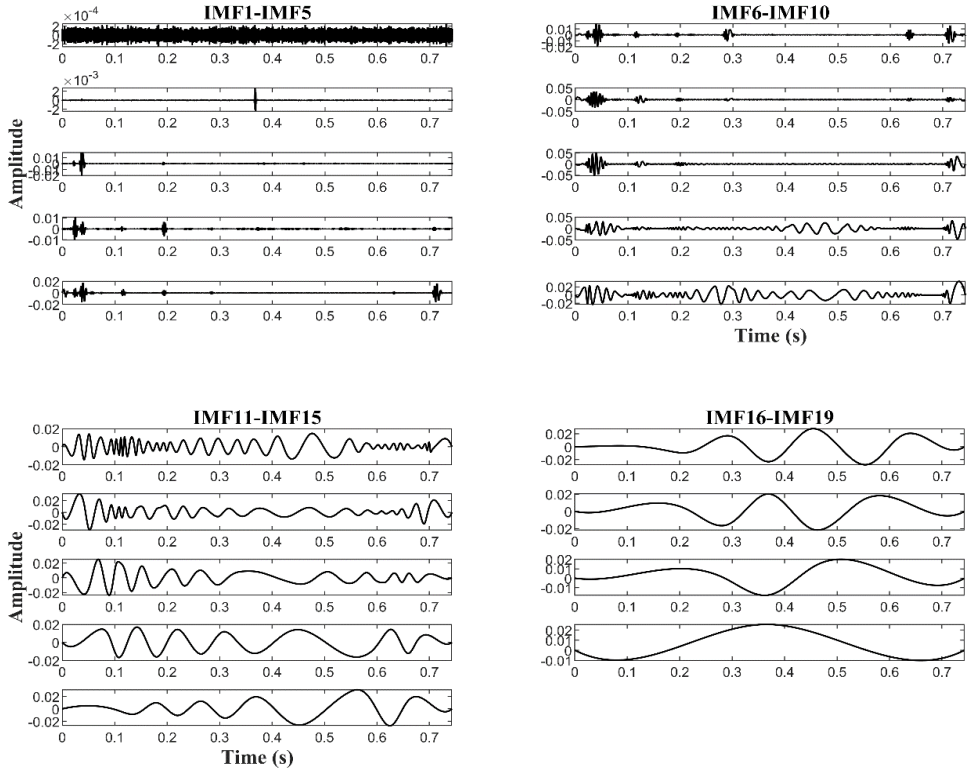
$$SBTH_v(n) = [1 - NBTH_v(n)] \tag{39}$$

where F_{DPP} is the FDPP algorithm which is used within the FD technique for estimating FD peaks which may correspond to crackles.

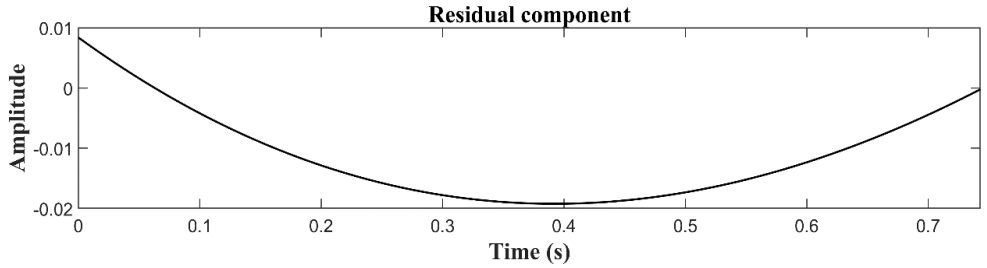
After that by multiplying the threshold $NBTH_v(n)$, with the corresponding IMFs, the part of the IMFs related to crackles $c_v^{Crackles}(n), v = 1, \ldots, V;\; n = 1, \ldots, N$ is obtained and in the same way by multiplying the threshold $SBTH_v(n)$ with the corresponding IMFs, the portions related to normal breath sounds $c_v^{NBS}(n), v = 1, \ldots, V;\; n = 1, \ldots, N,$ are kept.



(a)



(b)



(c)

Figure 24 Result of applying the EMD method (a) A time section of 0.743 s lung sound recorded from a BE patient (Chapter 3, Table 2 case RBCC), (b) Estimated IMFs (c) Estimated residual component.

Step (4):- Now the nonstationary and stationary parts of the signal are calculated using Eq. (40) and Eq. (41), respectively.

$$NSTS(n) = \sum_{v=1}^{V} c_v^{Crackles}(n), n = 1,2,\ldots,N \tag{40}$$

$$STS(n) = \sum_{v=1}^{V} c_v^{NBS}(n) + \sum_{u_1=V+1}^{F} c_{u_1}(n) + r_F(n) \tag{41}$$

If the crackles are not a good estimate of the expected crackle signal, the EMD-FD filter can be further extended to an iterative EMD-FD filter based on the stationary otput ($STS(n)$) (see Figure 25).

Step (5):- In the iterative process to stop the iteration a stopping criterion using Eq. (42) is calculated.

$$STC2_k = |\overline{(STS_{k-1}^2(n))} - \overline{(STS_k^2(n))}| \tag{42}$$

where k is the iteration index i.e. $k = 1,2,\ldots,K$ and $STS_0 = 0$.

Step (6):- In this step a computed stopping criterion i.e. *STC2* is compared with the constant value β_2.

If $STC2_k >= \beta_2$, input signal $y_{k+1} = STS_k(n)$ and repeat the step 1 to 5 else *k* = *K* and end the iterative loop.

where β_2 is the accuracy level whose value lie between 0 to 1. Here value of the β_2 is selected β_2=0.1.

Step (7):- In the final step, that nonstationary and stationary parts of the signal are calculated.

$$NS(n) = \sum_{k=1}^{K} NSTS_k(n) \tag{43}$$

$$ST(n) = STS_K(n), n = 1,2,\ldots,N \tag{44}$$

For better understanding of the EMD-FD filter a flow chart is presented in Figure 25 and a worked example of the EMD-FD filter is shown in Figure 26. Figure 26 (a) shows an input lung sound signal with coarse crackles (marked with arrowheads). The non-stationary and stationary outputs after applying the EMD-FD filter are shown in Figure 26 (b) and (c), respectively. Comparing these results with the input lung sound signal it can be observed that all the crackles are separated into

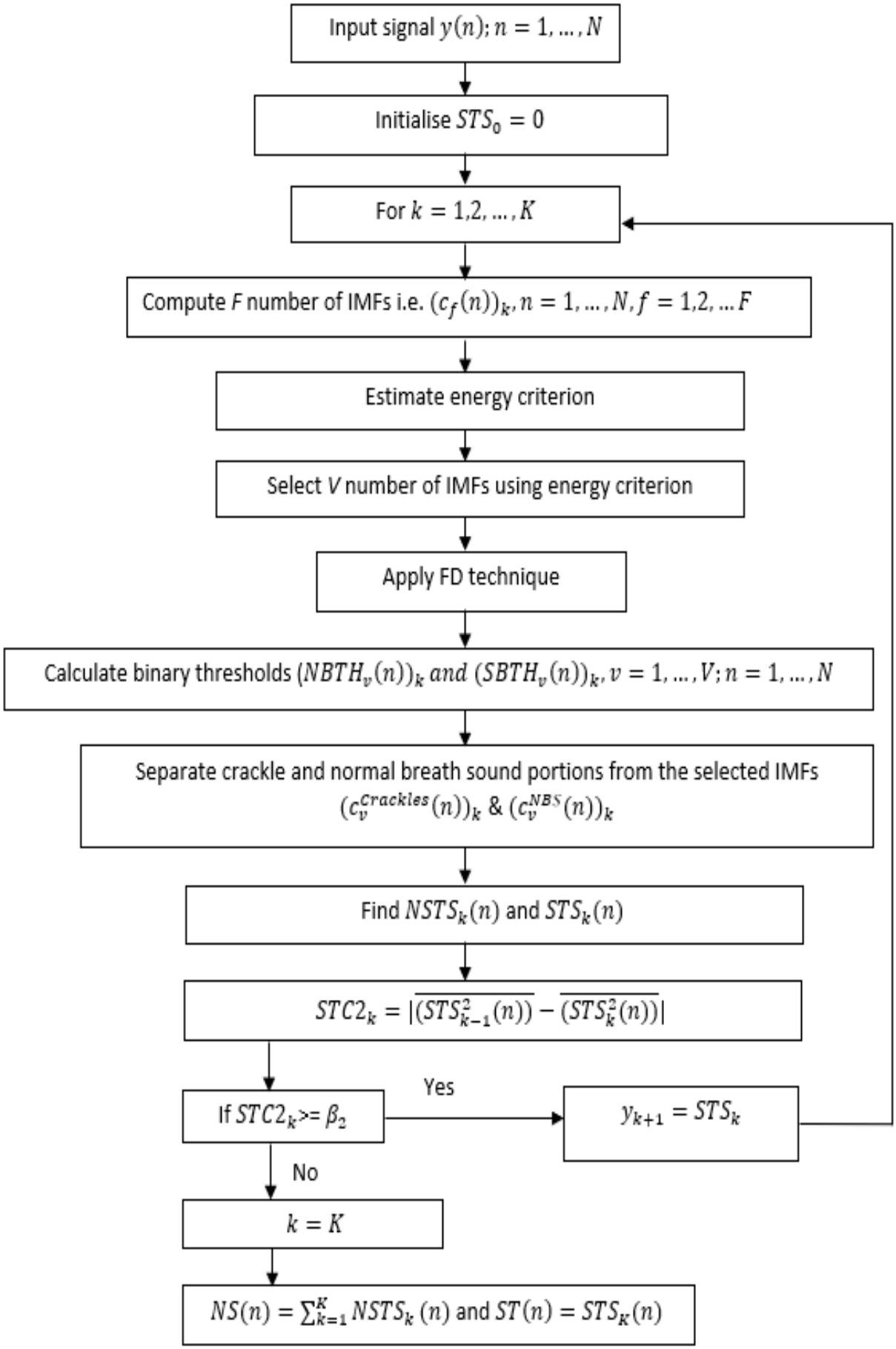
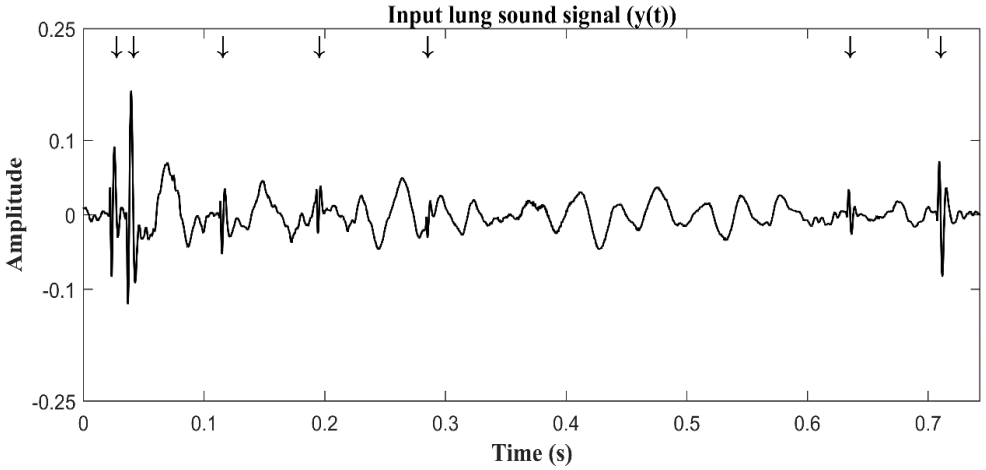
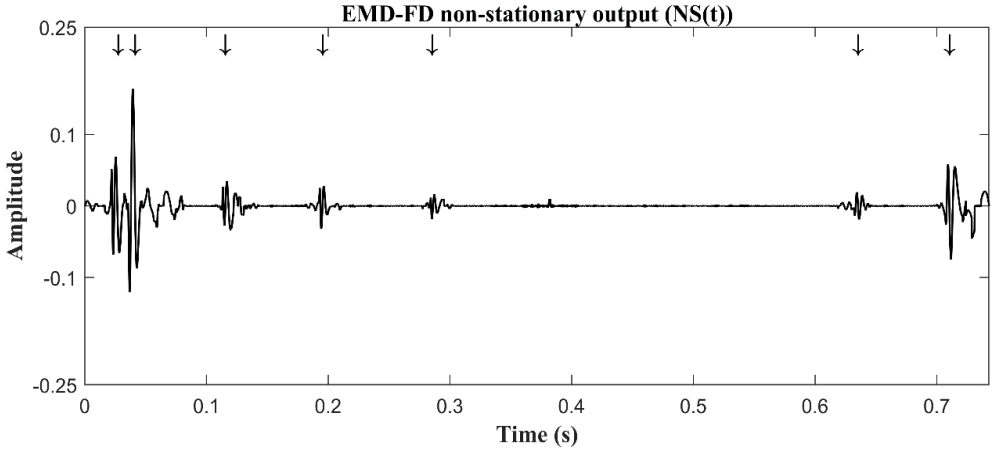


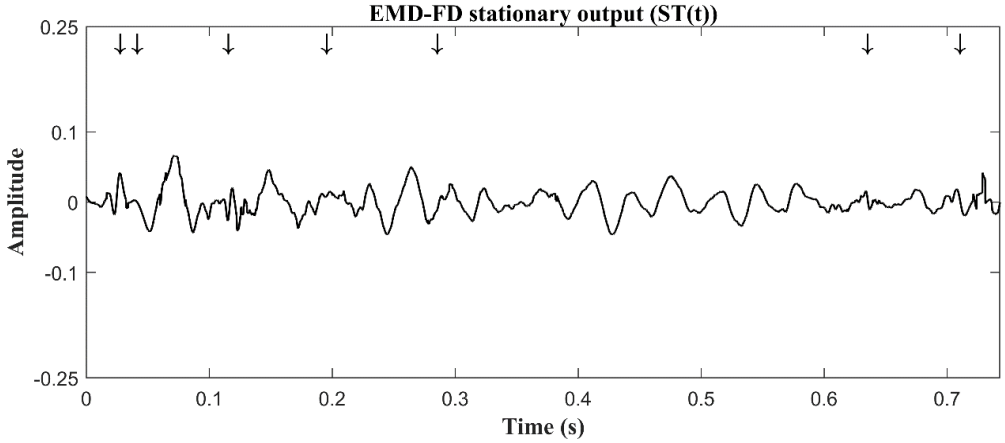
Figure 25 A schematic diagram of the EMD-FD filter.



(a)



(b)



(c)

Figure 26 Result of applying the EMD-FD filter (a) A time section of 0.743 s lung sound recorded from a BE patient (Chapter 2, Table 3 case RBCC), (b) EMD-FD filter non-stationary output (c) EMD-FD filter stationary output.

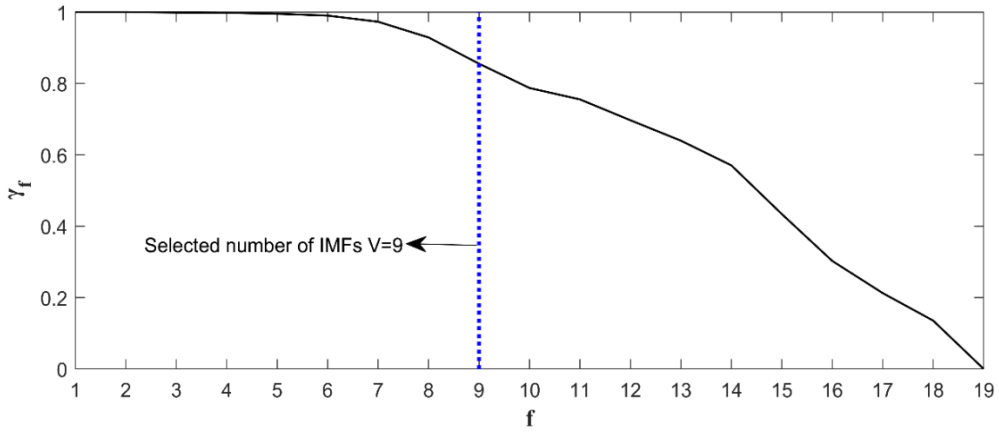


Figure 27 The estimated γ_f, parameter, corresponding to 19 IMFs.

the non-stationary output and normal breath sounds in the stationary output. However, in the EMD-FD filter the selection of how many independent mode functions contain crackle information is very challenging. A larger or smaller number of IMFs may lead to overestimation or underestimation, respectively. As can be observed in Figure 26 (b), the non-stationary output not only contains the crackles but also some portion of the normal breath sounds (overestimation), which is probably due to selecting to many IMFs using the criterion discribed in step 2. The estimated γ_f, parameter, corresponding to 19 IMFs is shown in Figure 27. The selected V IMFs using Eq. (36) is marked in Figure 27.

4.7 Summary

Chapter 4 provided the detailed working process of the FD technique and three different previously published crackle separation techniques: the WTST-NST filter, the WT-FD filter and the EMD-FD filter. Automatic separation of crackles from normal breath sounds can lead to better estimation of crackle features and hence can play an important role in early diagnosis or monitoring of different cardio pulmonary diseases. Separation can reveal not only the large amplitude crackles but also the small amplitude crackles, which are often significantly masked by the normal breath sounds. In the next chapter, a new method for automatic separation of crackles from normal breath sounds is presented.

Chapter 5 Proposed iterative envelope mean fractal dimension (IEM-FD) filter

5.1 Introduction

In the last chapter, the detailed working process of some previously published crackle separation techniques was presented. This chapter will provide the detailed working process of the proposed IEM-FD filter for automatically separating crackles from normal breath sounds which is the primary contribution of this study.

5.2 Iterative envelope mean fractal dimension (IEM-FD) filter

The IEM-FD filter comprises two techniques: (a) Iterative envelope mean (IEM) method and (b) FD technique (Hadjileontiadis & Rekanos, 2003). The FD technique is already used with the wavelet transform (WT-FD filter: Hadjileontiadis, 2005(I); Hadjileontiadis, 2005 (II)) and with empirical mode decomposition (EMD-FD filter: Hadjileontiadis, 2007) for crackle separation (Chapter 4). Here the use of the FD technique is paired with a new separation method: IEM. The IEM method separates the stationary and non-stationary parts of the lung sound signal and, the FD technique is then applied to the non-stationary output of the IEM method to refine the separation further. Lung sounds are a combination of normal breath sounds and aided sounds such as crackles. As with previously published techniques IEM relies on the relative stationarity of the short-duration crackles sounds as compared with the breath sounds.

5.2.1 Iterative envelope mean method

The IEM method is a new method for estimating stationary and non-stationary parts of the lung sound signal. In the IEM method, firstly we estimated the upper and lower envelopes of the smoothed lung sound signal. The upper and lower envelopes are estimated using the local maximum and local minimum locations respectively of the first derivative of the smoothed lung sound signal, next the envelope mean value is estimated using the average of the upper and lower envelopes of the smoothed lung sound signal. The idea behind the IEM method is to subtract the envelope mean value from the original lung sound signal and to use the resulting signal as the input for a subsequent iteration. After a number of iterations, *Q,* the IEM method will provide an estimate of the non-stationary part of the lung sound signal oscillating around the horizontal axis and the summation of the envelope mean at each iteration will give the stationary part of the lung sound signal.

The detailed working of the IEM method is described below.

In the first step, the lung sound signal is smoothed and its first and second derivatives are calculated using a filter from the SG family. The SG filter parameters are selected according to the guidelines proposed by Vannuccini et al., (1998): degree of fitting polynomial p_f = *4* and number of coefficients n_c equal to approximately one or two times the half-width of the shortest-duration feature of interest in the signal. In the case of crackles, the first deflection (IDW) is considered the shortest cycle of the crackle. The IDW of the crackle is generally less than 2 ms and in our data, where the sampling frequency of the lung sound signal is 44,100 Hz, the half width is less than 88 samples. The SG filter parameters used here are therefore p_f = 4, n_c = 89 and order of derivation d_o = 0, 1 and 2 for smoothing the lung sound signal, and estimating first and second derivatives of the smoothed lung sound signal, respectively.

Once the smoothed lung sound signal and its first and second derivatives are calculated using the SG filter, all the local maxima and minima of the first derivative ($y_s'(n)$) are identified and classified using sign changes of the second derivative ($y_s''(n)$) of the smoothed lung sound signal ($y_s(n)$).

The coordinates of the smoothed lung sound signal at the location of each of the first derivative local maxima and minima are then calculated. A cubic spline interpolation is used to connect the maxima in the smoothed lung sound signal to define the upper envelope ($UP_{env}(n)$), and correspondingly, the local minima are connected with each other to extract the lower envelope ($LW_{env}(n)$). The envelope mean value is then calculated using the estimated upper and lower envelopes of the smoothed lung sound signal:

$$m_q(n) = \frac{UP_{env_q}(n) + LW_{env_q}(n)}{2} \tag{45}$$

where *q* is the iteration index where *q*=1, 2,…, Q. The envelope mean value is then subtracted from the lung sound signal to get an estimate of the non-stationary signal $R_q(n)$:

$$R_q(n) = y_q(n) - m_q(n) \tag{46}$$

where $y_q(n)$ is the lung sound signal at iteration *q*. Note that the envelope mean value is calculated using the smoothed lung sound signal ($y_{s_q}(n)$), first derivative ($y_{s_q}'(n)$) and second derivative ($y_{s_q}''(n)$) and once the envelope mean value is calculated the envelope mean value is subtracted from the un-smoothed lung sound signal $y_q(n)$ at iteration *q*.

In order to end the IEM method iterative process, a stopping criterion is estimated:

$$STC1_q = |\overline{(R_{q-1}^2(n))} - \overline{(R_q^2(n))}| \quad (47)$$

where $R_0 = 0$.

The stopping criterion ($STC1$) is compared with accuracy level β_1: $\{0 < \beta_1 < 1\}$. In this study the value of the β_1 is empirically set equal to 0.01. Note that the value of the accuracy level too close to 0 may increase the total number of iterations *Q*, with the increasing number of iterations, the elements of normal breath sounds in the non-stationary output of the IEM method may reduce but at the cost of high computational complexity. On the other hand, the value of the accuracy level too close to 1 minimize the total number of iterations hence reduce the computational complexity of the method but may be at the cost of overestimation which means some portion of the normal breath sounds may still remain in the non-stationary output of the IEM method.

If $STC1_q \geq \beta_1$, a new input signal $y_{q+1}(n) = R_q(n)$ is defined and the process is repeated (usually one or two iterations are sufficient, see Chapter 7 Table 4).

When the stopping criterion is met, the estimates of the non-stationary and stationary parts of the lung sound signal are calculated:

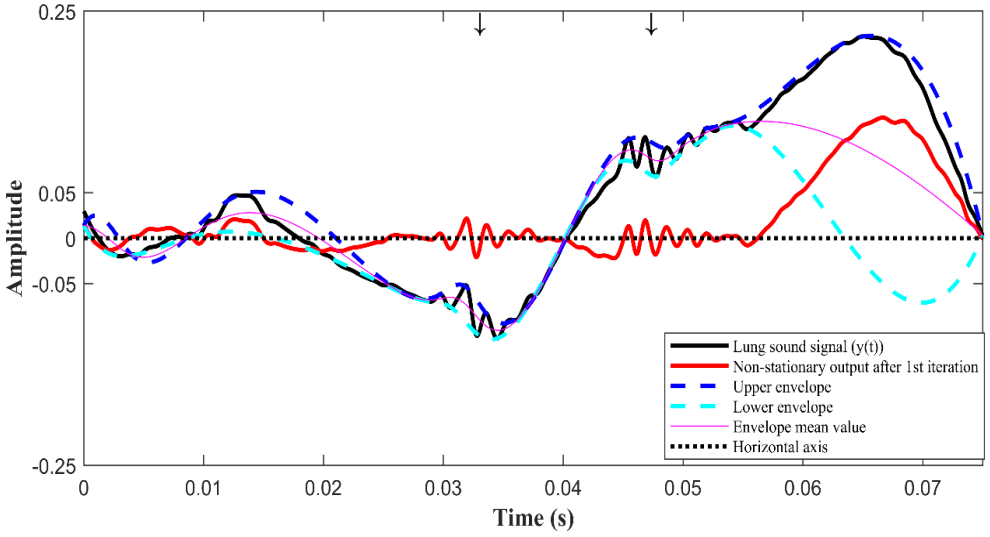
$$NSTS(n) = R_Q(n) \quad (48)$$

$$STS(n) = \sum_{q=1}^{Q} m_q(n) \quad (49)$$

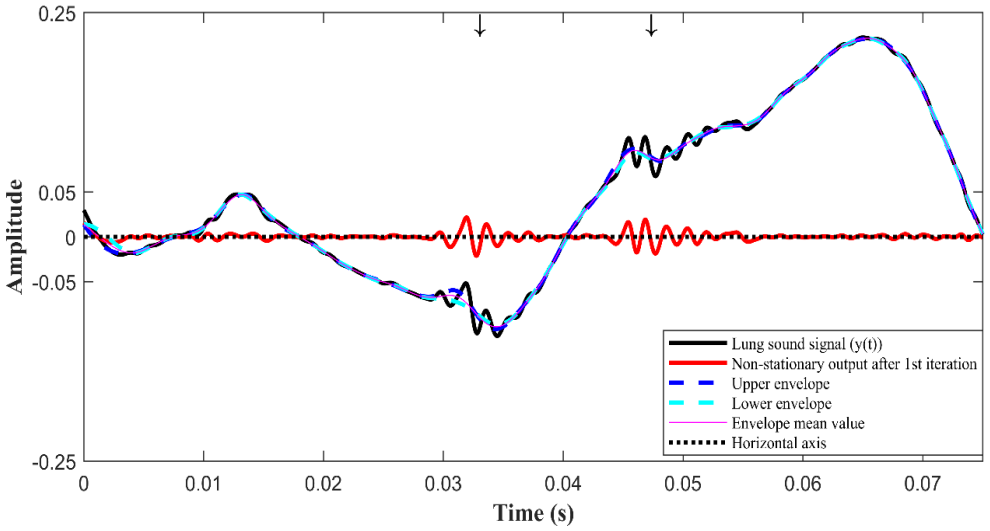
Note that in the IEM method the stopping criterion is defined in a same way it was used in (Hadjileontiadis & Panas, 1997), (Hadjileontiadis & Rekanos 2003), (Hadjileontiadis, 2005 (I)), (Hadjileontiadis, 2007) for stopping the iteration process.

As mentioned earlier, it is the first derivative local maxima and minima locations on the smoothed lung sound signal which are used for estimating the upper and lower envelopes if the upper and lower envelopes, using local maxima and minima of the smoothed lung sound signal itself are used, an inefficiency arises. The upper and lower envelopes can have large separation for regions with infrequently occurring extrema points (low frequency variation), which may require a large number of iterations for separating the lung sound signal. Using instead the upper and lower envelopes derived from the local maxima and minima of the first derivative reduces this inefficiency and consequently the number of iterations needed.

As an example, Figure 28 shows a section of duration 0.075 s of a lung sound signal recorded from a patient with IPF where the location of the crackles has been audio-visually identified by an experienced pulmonary acoustics researcher and marked with arrowheads. Figure 28 (a) displays



(a)



(b)

Figure 28 Illustration of the iterative envelope mean method applied to a section of 0.075 s of lung sound data recorded from a patient with IPF (Chapter 3, Table 2 Case RBFC); (a) estimation of the upper, lower and mean envelopes and the non-stationary signal estimate after one iteration using extrema points of the smoothed lung sound signal; (b) estimation of the upper, lower and mean envelopes and the non-stationary signal estimate after one iteration using extrema locations of the first derivative on the smoothed lung sound signal.

the non-stationary output of the IEM process after the 1st iteration, the upper and lower envelopes and the envelope mean value where upper and lower envelopes are estimated using directly the smoothed lung sound signal extrema points. It can be observed that between 0.055 s and 0.075 s, the separation between the upper and lower envelopes is large. As a result when the

envelope mean value is subtracted from the lung sound signal that region changes its shape but the non-stationary output not only contains the crackles but also consists of a large portion of normal breath sounds after the first iteration. On the other hand, in Figure 28 (b) where upper and lower envelopes are estimated using the first derivative local maxima and minima locations, we observe that the envelope mean is a closer fit to the lung sound signal and when it is subtracted, the amount of normal breath sounds in the non-stationary part is very much less, with crackles oscillating closely around the horizontal axis after only one iteration. Note that the lung sound signal is smoothed using SG filter prior to calculating the upper and lower envelopes to remove the high frequency peaks corresponding to the unwanted ripples in the signal without affecting the crackle waveform.

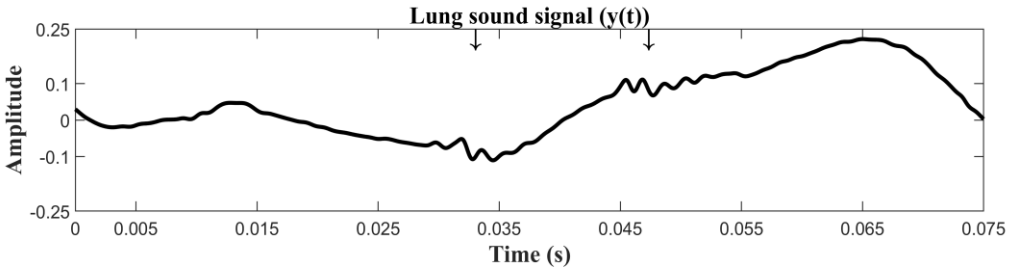
Huang (Huang et al., 1998) proposed the EMD technique for adaptively decomposing a signal into its IMFs and a residual component in decreasing order of frequency. Although the IEM method is superficially similar to the EMD technique in that processing of each iteration begins with an estimation of the upper and lower envelopes of a signal based on the local extrema which is then subtracted from the input signal, there are distinct differences between the two methods. In the EMD method, the input signal is not smoothed before calculating the upper and lower envelopes and the signal resulting from subtraction of the envelope is compared with a sifting criterion that checks whether the output meets the criteria for being an IMF (see section 4.6.1). On the other hand, in the IEM method the upper and lower envelopes are calculated from the smoothed lung sound signal and its derivatives. The envelope average is then subtracted from the original lung sound signal and the resulting signal is compared with a stopping criterion Eq. (47) to determine when to stop processing.

5.2.2 Fractal dimension technique

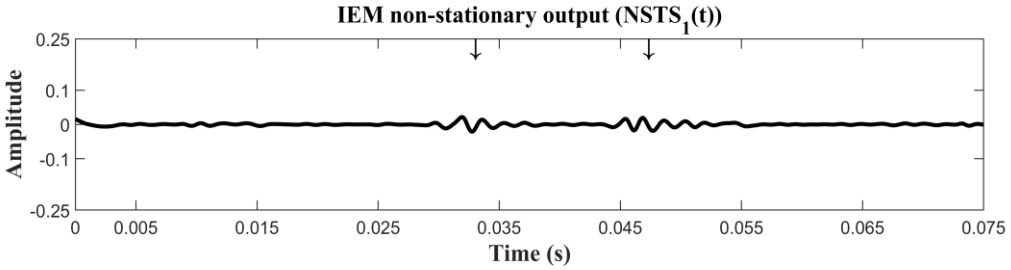
The IEM method makes a reasonable estimate of the stationary and non-stationary parts of the lung sound signal, however it is not usually sufficient by itself to extract only the crackles alone. To minimise the remaining elements of normal breath sounds in the non-stationary estimate ($NSTS(n)$) as shown in Figure 29 (b), the FD technique is applied (Hadjileontiadis & Rekanos, 2003). The detailed working process of the FD technique is discussed in section 4.3.

5.2.3 Iterative envelope mean fractal dimension filter

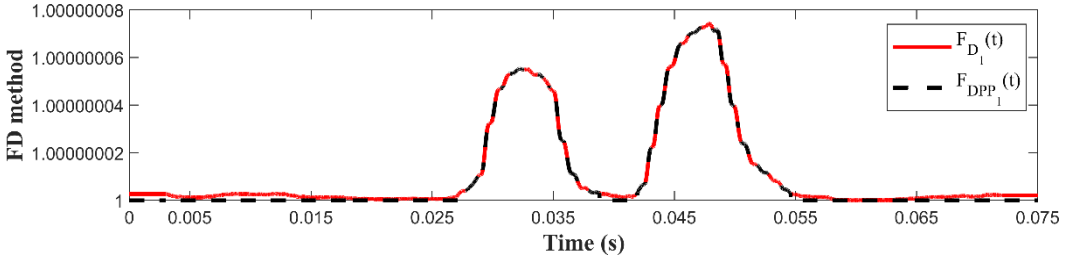
The IEM-FD filter is implemented using a combination of the IEM method and the FD technique (Hadjileontiadis & Rekanos, 2003). Two iteration loops are used in the IEM-FD filter: (i) an iteration loop related to IEM method $q = 1,2, \dots, Q$, and (ii) an iteration loop for the combination


Lung sound signal (y(t))
Amplitude
Time (s)

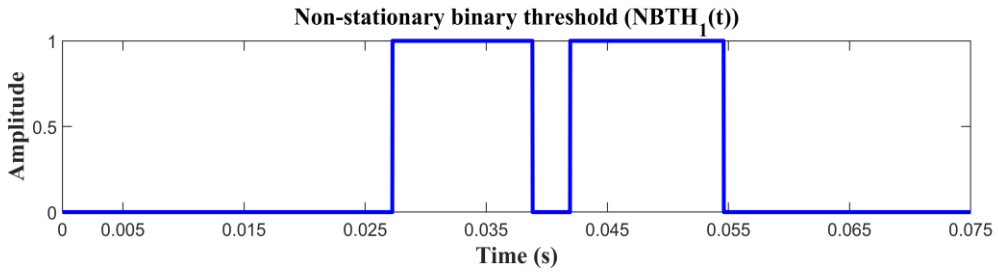
(a)


IEM non-stationary output (NSTS1(t))
Amplitude
Time (s)

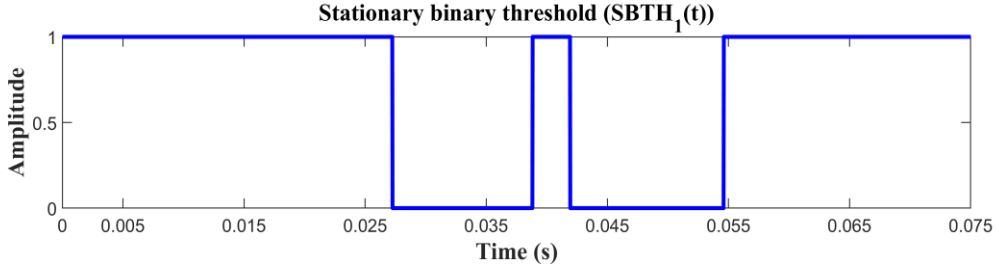
(b)


FD method
Time (s)
FD1(t)
FDPP1(t)

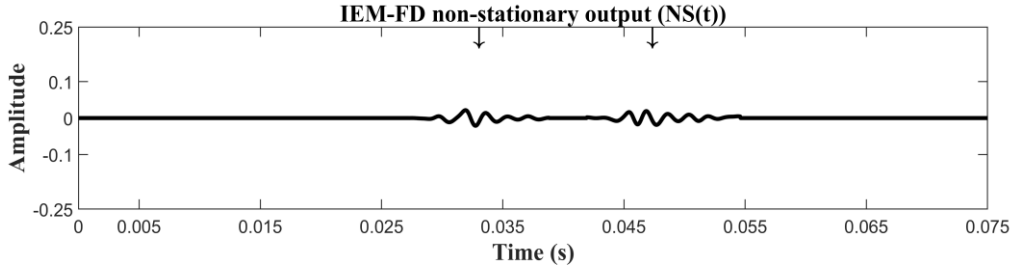
(c)


Non-stationary binary threshold (NBTH1(t))
Amplitude
Time (s)

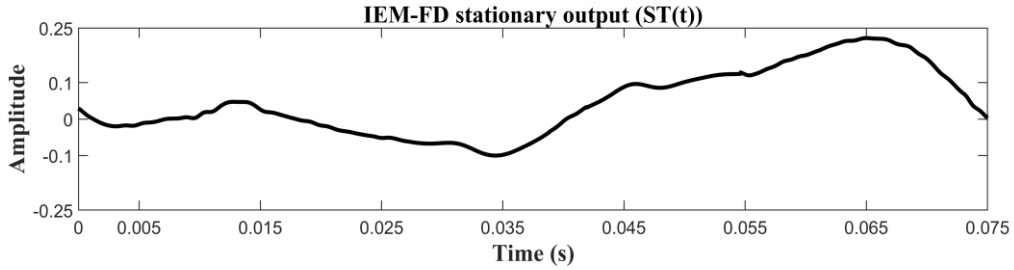
(d)


Stationary binary threshold (SBTH1(t))
Amplitude
Time (s)

(e)



(f)



(g)

Figure 29 Result of applying the proposed IEM-FD filter: (a) A time section of 0.075 s lung sound data recorded from a patient with IPF $\left(y(t)\right)$, where location of the crackles is marked with arrowheads; (b) Non-stationary output of the IEM method $(NSTS_1(t))$; (c) The FD of the IEM method non-stationary output $(F_{D_1}(t))$ and the FDPP algorithm for estimating FD valid peaks $(F_{DPP_1}(t))$; (d) The non-stationary binary threshold $(NBTH_1(t))$; (e) The stationary binary threshold $(SBTH_1(t))$; (f) The non-stationary output of the IEM-FD filter $(NS(t))$; (g) The stationary output of the IEM-FD filter $(ST(t))$.

of the IEM-FD filter $k = 1,2, \dots, K$. The IEM-FD filter working process is described below-

After the IEM method has estimated the non-stationary $(NSTS(n))$ and stationary $(STS(n))$ parts of the lung sound signal, point-to-point FD values of the estimated non-stationary output $(NSTS(n))$ are calculated. The FDPP algorithm (Hadjileontiadis & Rekanos, 2003) is now applied within the FD technique to automatically detect those peaks of the estimated $F_D(n)$ signal, as shown in Figure 29 (c), which may correspond to the crackles of the lung sound signal. Now using the estimated $F_{DPP}(n)$ sequence, two binary thresholds are calculated: the non-stationary binary threshold:

$$NBTH_k(n) = \begin{Bmatrix} 1 \; if \; F_{DPP_k}(n) \neq 1 \\ 0 \; if \; F_{DPP_k}(n) = 1 \end{Bmatrix} \quad (50)$$

and the stationary binary threshold

$$SBTH_k(n) = [1\text{-}NBTH_k(n)] \quad (51)$$

as displayed in Figure 29 (d) and Figure 29 (e), respectively.

The estimated non-stationary output of the IEM method ($NSTS(n)$) is multiplied by the non-stationary binary threshold ($NBTH(n)$) to get the refined non-stationary estimate $NST(n)$ and the non-stationary estimate of the IEM method ($NSTS(n)$) is multiplied by the stationary binary threshold ($SBTH(n)$) to obtain the remaining normal breath sounds $SSR(n)$ signal from the $NSTS(n)$:

$$NST_k(n) = NSTS_k(n)NBTH_k(n) \quad (52)$$

$$SSR_k(n) = NSTS_k(n)SBTH_k(n) \quad (53)$$

Equation (52) gives the estimate of the non-stationary output, $NST_k(n)$ and the summation of the Eq. (49) and Eq. (53) gives the estimate of the stationary output, $SSF_k(n)$ of the IEM-FD filter at iteration *k*.

$$SSF_k(n) = STS_k(n) + SSR_k(n) \quad (54)$$

To end the IEM-FD filter, a stopping criterion on the basis of stationary output can be calculated and compared with accuracy level (β_2).

The stopping criteria is calculated:

$$STC2_k = \left|\overline{(SSF_{k-1}^2(n))} - \overline{(SSF_k^2(n))}\right| \quad (55)$$

where $SSF_0 = 0$, and is compared with accuracy level β_2, where $0 < \beta_2 < 1$. If $STC2_k \geq \beta_2$, input signal $y_{k+1}(n) = SSF_k(n)$, otherwise $k = K$ and the iterative loop ends. Here the value of the β_2 is empirically set equal to 0.1 and, K represents the maximum iteration level. In the final step, the non-stationary and stationary parts of the signal are calculated, as shown in Figure 29 (f) and Figure 29 (g), respectively using Eq. (56) and Eq. (57) when *k=K.*

$$NS(n) = \sum_{k=1}^{K} NST_k\,(n) \quad (56)$$

$$ST(n) = SSF_K(n) \quad (57)$$

To provide better understanding a block diagram of the IEM-FD filter is shown in Figure 30 and a

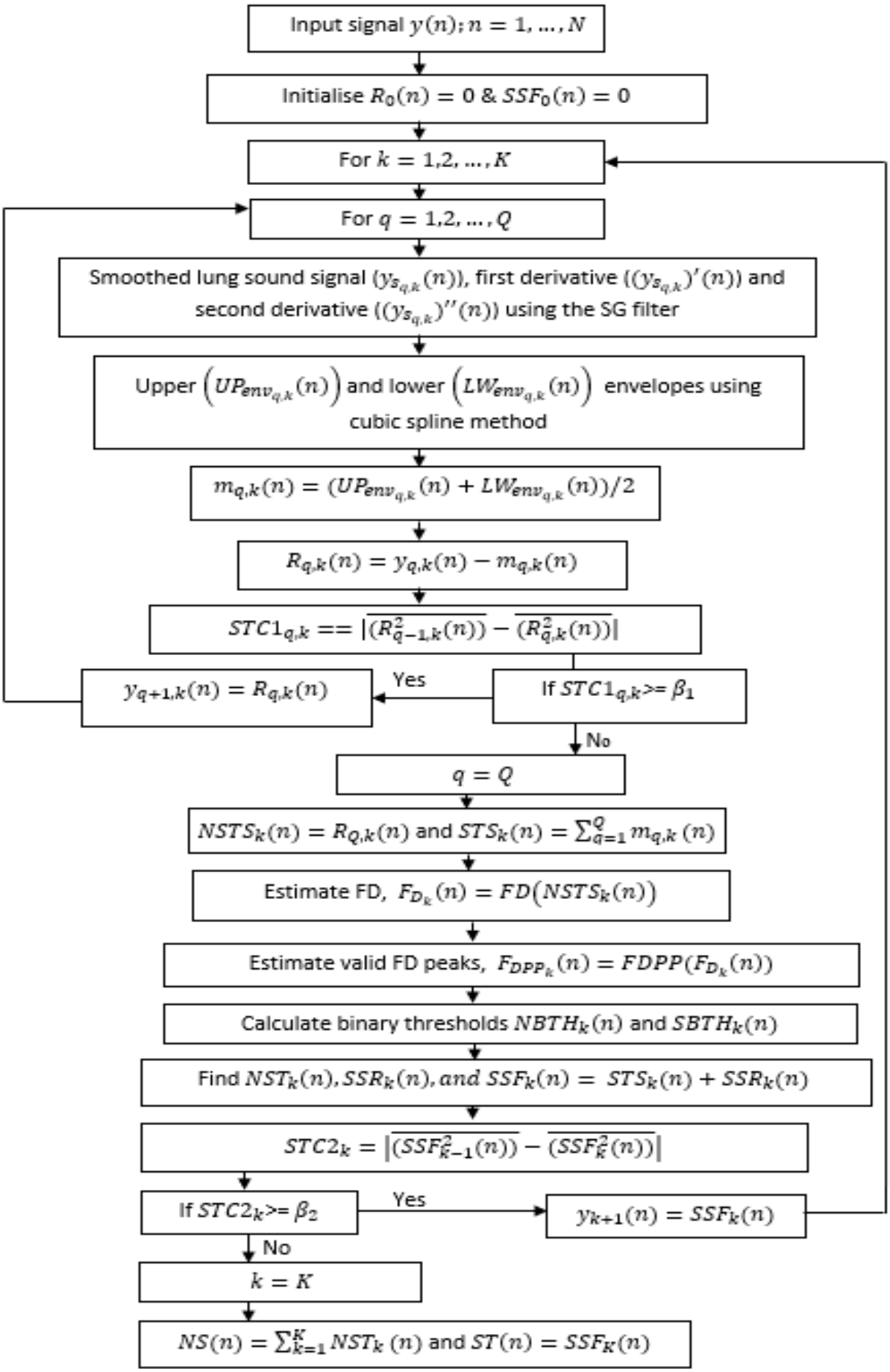
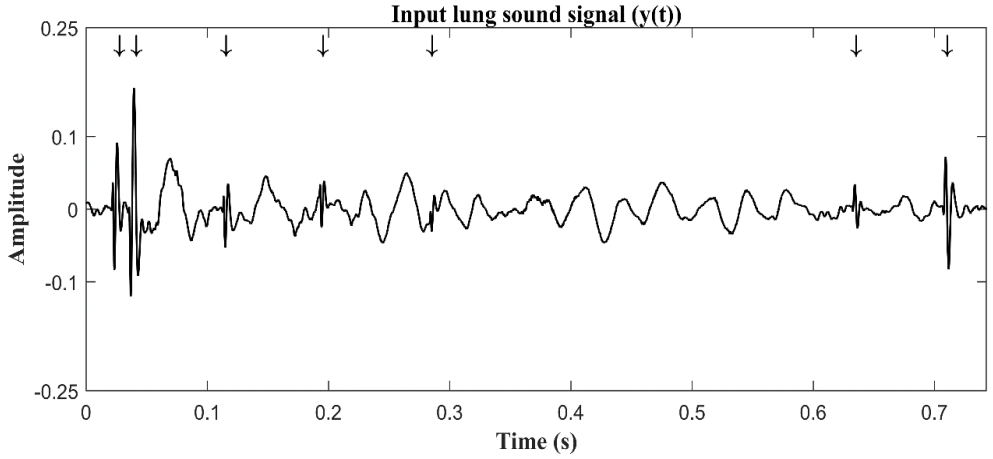
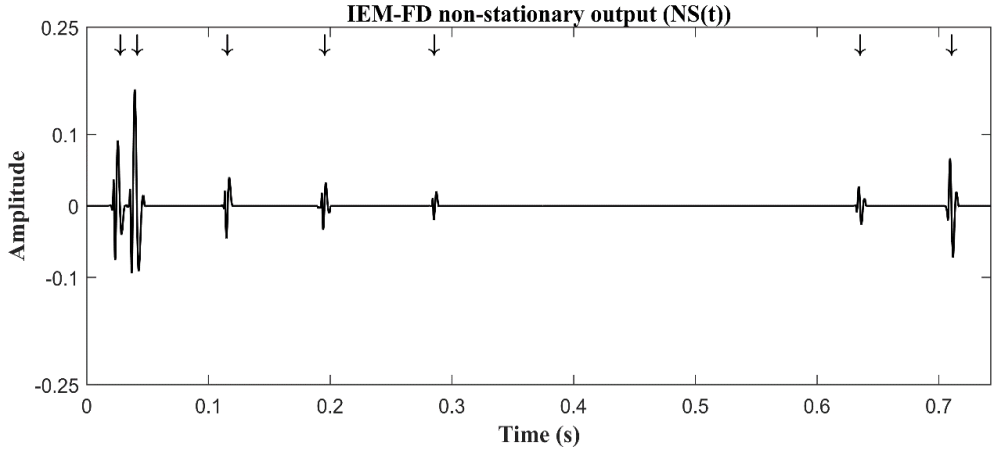


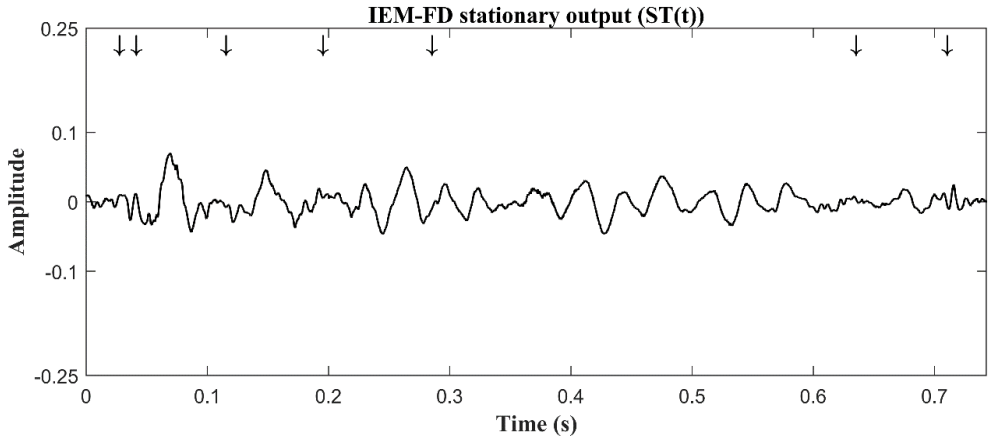
Figure 30 A schematic diagram of the IEM-FD filter.



(a)



(b)



(c)

Figure 31 Result of applying the IEM-FD filter (a) A time section of 0.743 s lung sound recorded from a BE patient (Chapter 3, Table 2 case RBCC), (b) IEM-FD filter non-stationary output (c) IEM-FD filter stationary output.

working example of the IEM-FD filter is shown in Figure 31. Figure 31 (a) shows an input lung sound signal with coarse crackles (marked with arrowheads). The non-stationary and stationary outputs after applying the IEM-FD filter are shown in Figure 31 (b) and (c), respectively. From the graphs it can be observed that not only are all the crackles separated into the non-stationary output with their time duration and morphology preserved, but also the normal breath sounds were retained with their proper shape and amplitude in the stationary output.

5.3 Summary

This chapter presented the detailed working process of the proposed IEM-FD filter. The crackle separation performance of the proposed IEM-FD filter will be compared with the selected previously published crackle separation techniques: the WTST-NST filter, WT-FD filter, and EMD-FD filter (see Chapter 4), using the dataset described in Chapter 3. The dataset, performance evaluators and filter parameters used for evaluating the crackle separation performance are discussed next.

Chapter 6 Dataset, quantitative evaluators and filter parameters

6.1 Introduction

In the last chapter we saw the detailed working process of the proposed IEM-FD filter. The dataset, different performance evaluators and filter parameters used for evaluating the crackle separation performance of the IEM-FD filter and of the selected previously published algorithms (WTST-NST filter, WT-FD filter and the EMD-FD filter) are discussed in this chapter.

6.2 Dataset and test samples

The dataset used for systematic testing of crackle separation techniques was described in Chapter 3. To explore the robustness of the separation process to noise, test samples were generated by burying 10 simulated or 10 real crackles within a simulated breath noise sample (BR_N). Average SNR was varied from -10 to 10 dB in steps of 1 dB. The local SNR for any given crackle in a test signal varies randomly, which may affect separation performance, therefore for each set of crackles and each average SNR, 501 test samples were generated, each with its own sample of BR_N. The use of 501 test samples at each SNR is justified in section 6.4.1. All test signals are sampled at 44,100 Hz.

6.3 Quantitative evaluators

Any successful crackle separation method must meet three criteria: extracting all the embedded crackles, minimizing the inclusion of non-crackle components and preserving crackle morphology after separation. As mentioned in Chapter 2, the failure to extract all crackles or loss of some portion of the crackle in the output signal is called under-estimation, and the inclusion of non-crackle components is over-estimation.

Using the test dataset (see Chapter 3, Table 2) the performance of the IEM-FD filter was evaluated against reference test signals and against the performance of the WTS-NST filter (Hadjileontiadis & Panas, 1997), WT-FD filter (Hadjileontiadis, 2005(I), Hadjileontiadis, 2005(II)) and EMD-FD filter (Hadjileontiadis, 2007). For the synthesized test signals, the time series of the crackles in the absence of breath noise was used as a reference signal. In the test samples measured in patients (RBFC and RBCC), the location of the crackles was audio visually marked by an experienced pulmonary acoustics researcher.

Separation performance was evaluated using several different metrics to assess its quality: To measure the similarity between the estimated non-stationary output of the each separation process and the crackle reference signal, the cross-correlation index (CCI) was used (except in the case of the EMD-FD filter, see section 7.2.1); Accuracy of crackle separation was assessed by two quantitative evaluators proposed by Hadjileontiadis & Panas, (1997): Rate of Detectability (D_R), and Total performance (TD_R); To quantify over- or under-estimation in the separation, the quality factors (QFs) proposed by Hadjileontiadis & Panas, (1997) were adapted to benefit from the existence, for our test data, of the reference signals; To evaluate the ability of the separation process to preserve crackle morphology, the 2CD percentage error (PE_{2CD}) was calculated.

The process for calculating each metric is as follows:

6.3.1 Cross Correlation Index (CCI)

The CCI indicates the associate between two signals. Therefore, CCI is used to show how well an algorithm separates crackles. The CCI was calculated using Pearson's correlation coefficient:

$$CCI = \frac{\sum_{n=1}^{N}(R_C(n) - \bar{R}_C)(NS(n) - \overline{NS})}{\sqrt{\sum_{n=1}^{N}(R_C(n) - \bar{R}_C)^2}\sqrt{\sum_{n=1}^{N}(NS(n) - \overline{NS})^2}} \times 100\ \% \quad (58)$$

where $R_C(n)$ is the crackle reference signal, $NS(n)$ is the non-stationary output of the separation method and $\bar{R}_C$ and $\overline{NS}$ are the average values of the crackle reference signal and non-stationary output, respectively.

6.3.2 Rate of Detectability (D_R)

The rate of detectability (D_R) measures the ability of an algorithm to separate crackles into their non-stationary outputs. D_R was calculated using Eq. (59).

$$D_R = \frac{N_E}{N_R} \times 100\ \% \quad (59)$$

where N_E is the number of crackles in the non-stationary output of a separation algorithm and N_R is the number of crackles in the input signal.

In the case of SFC, RFC, SCC and RCC where crackle reference signal is available, the number of separated crackles in the non-stationary output is calculated by correlating each crackle in the reference signal with the portion of a non-stationary output within the same time frame of the reference crackle. Note that each reference signal contains 10 crackles, see Chapter 3, Table 2. Therefore, this process will provide us 10 different values of the CCI corresponding to the each

reference crackle i.e. CCI_e where *e* is the crackle index, $e = 1,2, \ldots \ldots \ldots 10$. Note that corresponding to any reference crackle time frame if the complete portion of the non-stationary output is zero within that time frame, the CCI_e corresponding to that reference crackle will be undefined, in that case the value is set to zero. Now, out of 10 CCI values $CCI_e \geq 0.5$, where the cut-off CCI value was empirically selected, were counted as contributing to N_E.

For RBFC and RBCC where crackle reference signals do not exist, the number of detected crackles in the output was counted manually by comparing their location with the marked crackles in the input signal. D_R was calculated for each test sample of the all test signals. The mean value and standard deviation of the rate of detectability over all test samples of each test signal are reported (see Chapter 7, Table 4).

6.3.3 Total Performance

The total performance also measures accuracy in terms of separation of crackles in the non-stationary output by the proposed and previously selected methods. The total performance is calculated for each type of crackle (fine and coarse).

$$TD_R^{XX} = \overline{D}_R^{XX} \,\% \tag{60}$$

$$SD_{TD_R}^{XX} = \sigma(D_R^{XX}) \,\% \tag{61}$$

where σ is the standard deviation, D_R is the rate of detectability, and XX stands for FC for fine crackles and CC for coarse crackles.

In the case of fine crackles the total performance ($TD_R^{FC}, SD_{TD_R}^{FC}$) is calculated by taking the mean and standard deviation of the rate of detectability over all test samples of fine crackles: simulated (1,503 test samples, 501 test samples in each simulated case), real (501 test samples) and one test sample of RBFC (Chapter 7, Table 4, Cases SFC, RFC and RBFC). Correspondingly, the total performance in the case of coarse crackles ($TD_R^{CC}, SD_{TD_R}^{CC}$) is estimated by taking the mean and standard deviation of the rate of detectability over all test samples of coarse crackles: simulated (1,503 test samples, 501 test samples in each simulated case), real (501 test samples) and one test sample of RBCC (Chapter 7, Table 4, Cases SCC, RCC and RBCC).

6.3.4 Quality Factors (QFs)

Quality Factors measure over- and under-estimation in the non-stationary output signal. Hadjileontiadis & Panas, (1997) proposed QFs for assessing over-and under-estimation. However, the test data set here provides us with reference signals not available for Hadjileontiadis et al.'s

test signals, therefore in this study, four quality factors: a reference quality factor (R_{QF_U}) for under-estimation, an estimated quality factor (E_{QF_U}) for under-estimation, a reference quality factor (R_{QF_O}) for over-estimation and an estimated quality factor (E_{QF_O}) for over-estimation are defined. To calculate the QFs, firstly two thresholds are defined:

$$TH_1(n) = \begin{cases} 1 \; if R_C(n) \neq 0 \\ 0 \; if \; R_C(n) = 0 \end{cases} \quad (62)$$

$$TH_2(n) = [1 - TH_1(n)] \quad (63)$$

where $R_C(n)$ is the crackle reference signal and *n* is the sample index with $n = 1,2 \ldots \ldots N$. Secondly, the input signal $y(n)$, is multiplied by the threshold $TH_2(n)$ to calculate a background noise reference signal $R_{BN}(n)$.

$$R_{BN}(n) = y(n)TH_2(n) \quad (64)$$

Thirdly, the non-stationary output $NS(n)$ of a separation algorithm is divided into two parts: non-stationary signal with only crackle portion $(NS_C(n))$ and remaining non-stationary part $(NS_R(n))$, according to:

$$NS_C(n) = NS(n)\, TH_1(n) \quad (65)$$

$$NS_R(n) = NS(n)\, TH_2(n) \quad (66)$$

Now QFs for under-estimation are calculated using the area under the input signal $y(n)$, area under the crackle reference signal $R_C(n)$ and area under the crackle portion of the non-stationary signal $NS_C(n)$. Similarly, the two quality factors for over-estimation are evaluated using the area under the input signal $y(n)$, area under the background noise reference signal $R_{BN}(n)$ and area under the remaining non-stationary part $NS_R(n)$:

$$R_{QF_U} = \frac{\sum_1^N |y(n)|\,\Delta n - \sum_1^N |R_C(n)|\,\Delta n}{\sum_1^N |y(n)|\,\Delta n} \quad (67)$$

$$E_{QF_U} = \frac{\sum_1^N |y(n)|\,\Delta n - \sum_1^N |NS_C(n)|\,\Delta n}{\sum_1^N |y(n)|\,\Delta n} \quad (68)$$

$$R_{QF_O} = \frac{\sum_1^N |y(n)|\,\Delta n - \sum_1^N |R_{BN}(n)|\,\Delta n}{\sum_1^N |y(n)|\,\Delta n} \quad (69)$$

$$E_{QF_O} = \frac{\sum_1^N |y(n)|\,\Delta n - \sum_1^N |NS_R(n)|\,\Delta n}{\sum_1^N |y(n)|\,\Delta n} \quad (70)$$

where Δ*n* is the sample period in seconds. The maximum value of area under $|NS_C(n)|$ was set equal to the area under$|R_C(n)|$. E_{QF_U} in the range $R_{QF_U} < E_{QF_U} \leq 1$ represents under-estimation and a value close to 1 indicates high under-estimation. Similarly E_{QF_O}in the range $R_{QF_O} \leq E_{QF_O} < 1$ represents over-estimation and a value close to R_{QF_O} shows high over-estimation. Hence, a value of E_{QF_U} close to R_{QF_U} with a value of E_{QF_O} near to 1 represents good quality crackle separation without either high under- or high over-estimation.

6.3.5 2CD percentage error (PE_{2CD})

The 2CD percentage error assesses the ability of an algorithm to preserve crackle morphology after separation from background noise. The percentage error in 2CD following separation is calculated using:

$$PE_{2CD} = \left|\frac{AC_{2CD} - EC_{2CD}}{AC_{2CD}}\right| \times 100\ \% \tag{71}$$

where AC_{2CD} is the actual crackle 2CD calculated from the crackle reference signal $R_C(n)$ and EC_{2CD} is the estimated crackle 2CD calculated from the non-stationary filter output ($NS(n)$). The 2CD was calculated using first five zero crossings of the crackle.

For the RBFC and the RBCC signals, where a crackle reference signal does not exist, the crackle separation performance of the separation filters was not evaluated using the QFs or the PE_{2CD}.

6.4 Filter Parameters

The parameters used for the IEM-FD filter and the previously published WTST-NST filter (Hadjileontiadis & Panas, 1997), WT-FD filter (Hadjileontiadis, 2005(I); Hadjileontiadis, 2005(II)) and EMD-FD filter (Hadjileontiadis, 2007) are shown in Table 3.

6.4.1 Selection of number of test samples

As mentioned earlier (section 6.2), 501 test samples are generated for each simulated test signal to account for the effect of random variation of the local SNR around any given crackle. For each SNR, the same 10 crackles embedded in 501 unique noise signal samples are passed through an algorithm (except in the case of the EMD-FD filter) and the resulting CCIs are averaged to get one CCI value for each SNR value for the IEM-FD filter, WTST-NST filter and the WT-FD filter (as shown in Chapter 7, Figure 33).

To justify the choice to use 501 test samples, Figure 32 shows the average CCI for the IEM-FD filter

Table 3 Parameters used for different separating methods.

Parameters		**IEM-FD**	**WTST-NST**	**WT-FD**	**EMD-FD**
Number of samples (*N*)		32,768	32,768	32,768	32,768
Number of decomposition levels (*M*)		NA	$M = \log_2(N) = 15$	$M = 1$ or $m = 1$	NA
Type of wavelet		NA	*db4*	*db4*	NA
Sampling frequency (f_s)		44,100 Hz	44,100 Hz	44,100 Hz	44,100 Hz
Accuracy level (β)	β_1	0.01	NA	NA	NA
	β_2	0.1	0.1	0.1	0.1
	β_3	0.01	NA	0.01	0.01
Duration of the window in seconds		W = 0.006 s	NA	W = 0.006 s	W = 0.006 s
Threshold		NA	F_{adj} =3	NA	p= 0.05

NA: not applicable; *db4*: Daubechies quadrature mirror filters (QMFs) of eight coefficients; *M*: Number of WT decomposition levels; in the IEM-FD filter window length is taken from Hadjileontiadis & Rekanos, 2003; in the WTST-NST filter type of wavelet and threshold (F_{adj}) are taken from Hadjileontiadis & Panas, 1997; in the WT-FD filter number of decomposition levels, type of wavelet, accuracy level (β_3), and window length are taken from Hadjileontiadis, 2005(II); in the EMD-FD filter accuracy level (β_3), window length, and threshold (p) are taken from Hadjileontiadis, 2007; β_1: Accuracy level for the IEM method; β_2: Accuracy level for desired stationary and non-stationary outputs; β_3: Accuracy level for the FDPP algorithm.

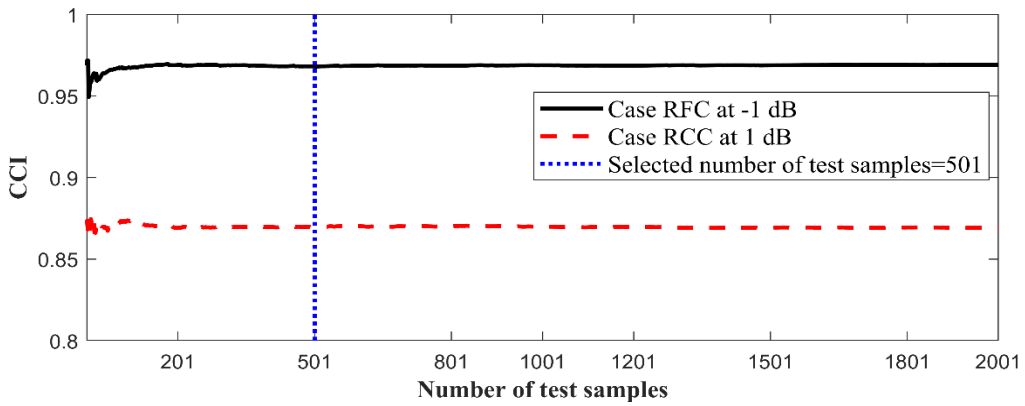


Figure 32 Selection of number of test samples to eliminate random variation on the IEM-FD filter crackle separation performance at local SNR using RFC case at -1 dB and RCC case at 1 dB.

in the case of RFC at an SNR of -1 dB and RCC at an SNR of 1 dB when the number of test samples is increased from 1 to 2001 in steps of 1. We note that for more than 501 samples, the increase in CCI is negligible in both cases. The selected number of 501 test samples is marked on Figure 32.

6.5 Summary

This chapter introduced the performance evaluators and filter parameters used for analyzing the crackle separation performance of the new and previously published crackle separation techniques. In the next chapter experimental results of the systematic comparison between the proposed and previously published crackle separation methods obtained using the test dataset (Chapter 3, Table 2) and above-mentioned quantitative evaluators are presented.

Chapter 7 Experimental results

7.1 Introduction

This chapter presents the results obtained using the IEM-FD filter and provides systematic comparison with the WTST-NST filter (Hadjileontiadis & Panas, 1997), the WT-FD filter (Hadjileontiadis, 2005(I), Hadjileontiadis, 2005(II)), and the EMD-FD filter (Hadjileontiadis, 2007). All the separation techniques are implemented using the Matlab (R2019a) programming language.

7.2 Experimental results

7.2.1 Performance of the IEM-FD filter

Figure 33 shows plots of CCI averaged over all 501 test signals against SNR using i) the IEM-FD filter, ii) the WTST-NST filter, and iii) the WT-FD filter for the separation. Plots labelled (a) show curves for SFC with three different values of IDW/2CD and RFC; plots labelled (b) show curves for SCC with three different values of IDW/2CD, and RCC.

Taking CCI >= 0.8 to indicate strong correlation between the separated signal and the test signal, strong correlation occurs for all fine crackle test signals with SNR greater than -1 dB except for SFC with 2CD = 6 ms for IEM-FD and WT-FD. For SFC with 2CD = 6 ms the CCI is just below 0.8 at SNR = - 1 dB but is above at SNR = 0 dB. For coarse crackles strong correlation occurs for SNR >= 1 dB except for WTST-NST filter in the case of RCC, SCC with 2CD=9.5 ms and SCC with 2CD=10 ms where performance is slightly low compared to the proposed IEM-FD filter and the WT-FD filter. For WTST-NST filter for RCC, SCC with 2CD=9.5 ms and SCC with 2CD=10 ms the CCI is low at SNR=1dB but gradually start to improve for SNR greater than 1dB. The plots therefore suggest two threshold SNRs above which good performance can be achieved by all three filters: SNR = - 1 dB for fine crackles and SNR = 1 dB for coarse crackles. From the results, we can see that the proposed IEM-FD filter performance is quite similar to, or better than, the WTST-NST filter and the WT-FD filter for SNRs greater than our selected thresholds but the WTST-NST filter has slightly better performance for SNRs less than our thresholds. Note that in this study the crackle separation performance of the EMD-FD filter using CCI plots was not tested in the range of -10 to 10 dB SNRs due to its tendency to fail to converge on some of the 501 test signals in each case.

Comparative evaluation of the IEM-FD filter and previously published methods was made using the synthesized signals at these threshold SNR values supplemented by the RBFC and RBCC using

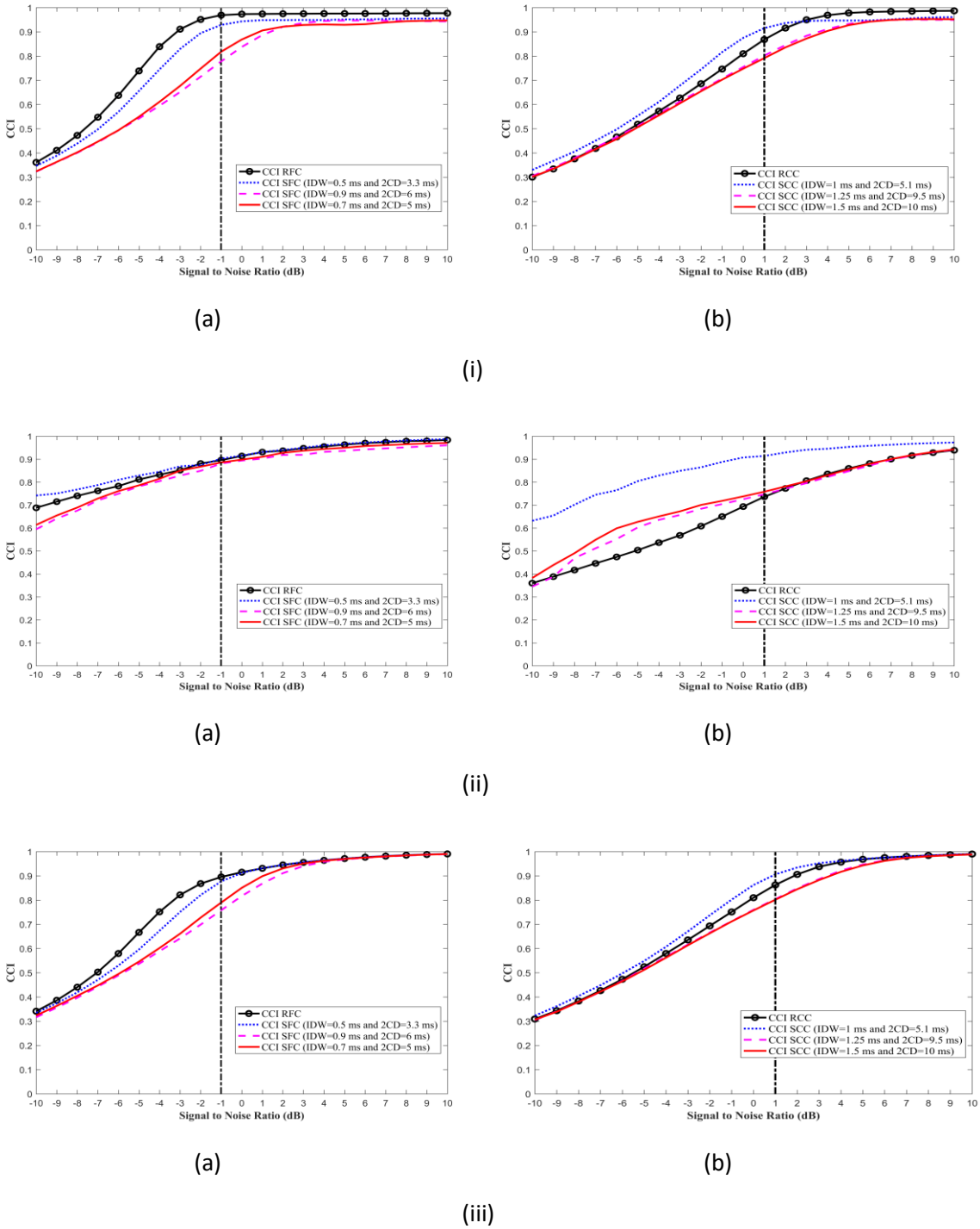


Figure 33 CCI plots for breath noise cases with a signal to noise ratio in the range of -10 to 10 dB (Chapter 3, Table 2). (i) The IEM-FD filter, (a) RFC and SFC cases; (b) RCC and SCC cases. (ii) The WTST-NST filter, (a) RFC and SFC cases; (b) RCC and SCC cases. (iii) The WT-FD filter, (a) RFC and SFC cases; (b) RCC and SCC cases.

the quantitative evaluators described in sections 6.3.2 to 6.3.5.

Separation by the IEM-FD filter into non-stationary and stationary estimates for test signals using two, randomly selected, real crackle cases at the threshold SNRs, one for fine and one for coarse

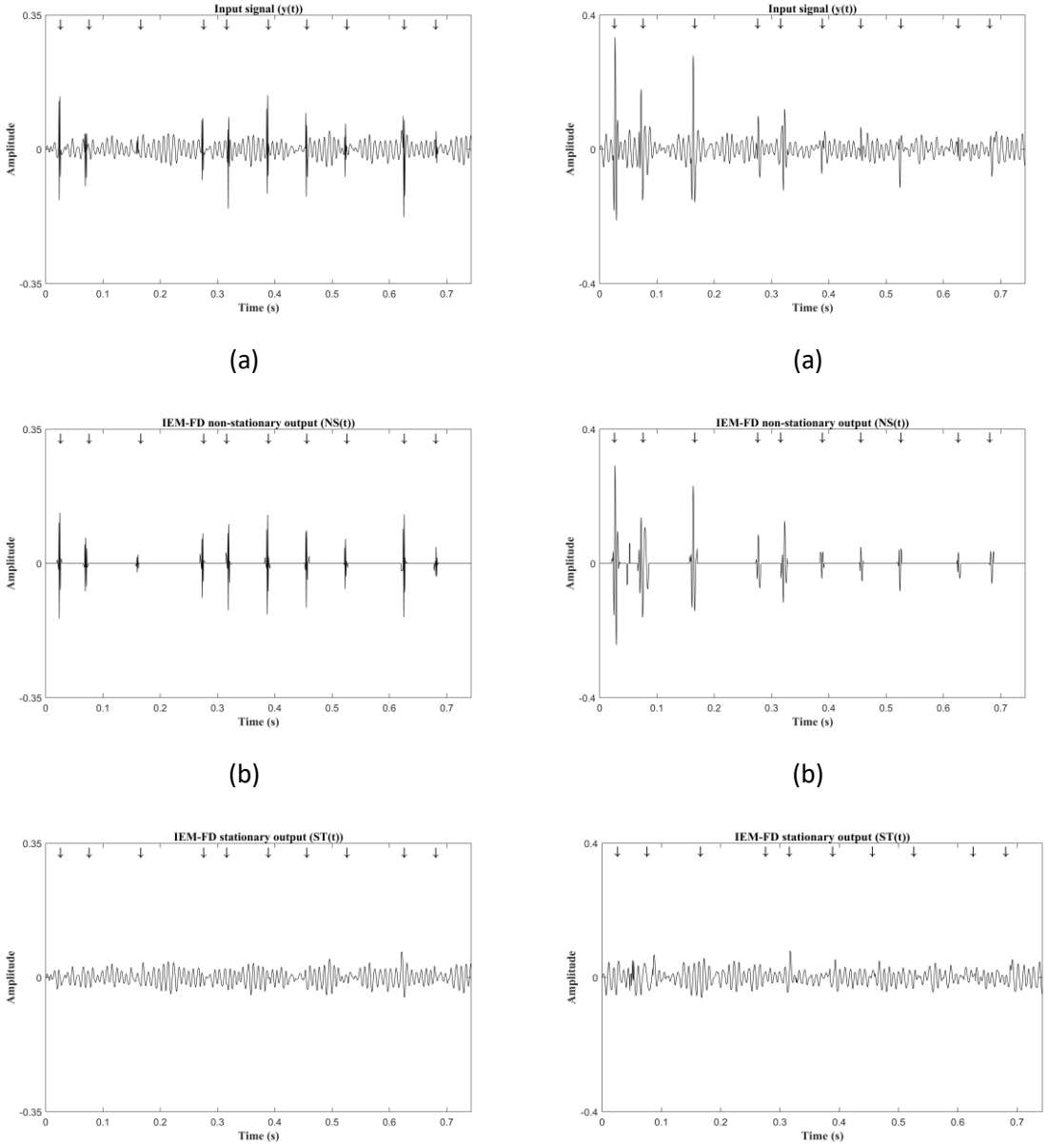


Figure 34 (i): (a) Input signal with RFC (Table 4, Case RFC); (b) IEM-FD filter non-stationary output; (c) IEM-FD filter stationary output. (ii): (a) Input signal with RCC (Table 4, Case RCC); (b) IEM-FD filter non-stationary output; (c) IEM-FD filter stationary output.

crackles (Table 4, cases RFC and RCC) are shown in Figure 34-i (a-c) and Figure 34-ii (a-c), respectively. The location of the crackles inserted into the background noise is marked with arrowheads. Figure 34 (a) displays the input signals. The non-stationary and stationary signal estimates after applying the IEM-FD filter are shown in Figure 34 (b) and (c), respectively. Comparing these with the input signal, we can clearly see that for both fine and coarse crackle

samples, all the fine and coarse crackles are separated from breath noise into the non-stationary signal estimate with their time duration and morphology preserved. For both fine and coarse crackles, the breath noise is retained in the stationary estimate with its proper shape and amplitude.

Now the comparison results of the proposed IEM-FD filter with the previously published methods is divided into two sections: the first part provides the comparison results of the IEM-FD filter with the WTST-NST filter and the WT-FD filter, and second part provides the comparison of the IEM-FD filter with the EMD-FD filter.

7.2.2 Comparison of the IEM-FD filter with the WTST-NST filter and the WT-FD filter

7.2.2.1 Rate of detectability and total performance

Table 4, Table 5, and Table 6 show the performance analysis of the IEM-FD filter, the WTST-NST filter and the WT-FD filter in terms of rate of detectability and total performance, respectively. The crackle separation performance of the IEM-FD filter, the WTST-NST filter and the WT-FD filter are evaluated using 501 test samples at SNR = -1 dB for real and simulated fine crackles, and 501 test samples at SNR = 1 dB for real and simulated coarse crackles, and on a RBFC and a RBCC. We note that the separation performance of all three methods in terms of rate of detectability is quite similar with slightly better performance by the WTST-NST filter in the case of RBFC. However, in terms of total performance the IEM-FD outperforms both WTST-NST and WT-FD in the case of coarse crackles (SCC, RCC and RBCC) and achieves the same performance as the WTST-NST and slightly better than the WT-FD in the case of fine crackles (SFC, RFC and RBFC). In the case of RBFC the performance of the IEM-FD filter and the WT-FD filter is slightly low due to the crackles remaining in the stationary output. However, the missing crackles of the IEM-FD filter and the WT-FD filter can be recovered either by changing the FDPP algorithm accuracy level (β_3) or by changing the accuracy level (β_2) for desired stationary and non- stationary outputs at the cost of increasing overestimation.

7.2.2.2 Quality of crackle separation (over or under estimation)

Tables 7 and 8 show the quality of crackle separation of the IEM-FD filter, the WTST-NST filter, and the WT-FD filter in terms of under- and over-estimation, respectively. Here we notice that (Table 7) in all three crackle separation techniques the average estimated under-estimation quality factor ($\bar{E}_{QF_U}$) is either very close or equal to the average reference under-estimation quality factor ($\bar{R}_{QF_U}$) indicating there is very little under-estimation. On the other hand, in terms of over-estimation (Table 8), we can see that the average estimated over-estimation quality factor

Table 4 Performance of proposed IEM-FD filter in the case of fine and coarse crackles.

IEM-FD										
Cases		**BN**	**D_g**	**SNR (dB)**	**N_R**	**NOTS**	**K (min-max)**	**Q (min-max)**	**N_E (min-max)**	**$\overline{D}_R(SD)$ (%)**
SFC	A_F	BR_N	NA	-1	10	501	1-1	2-2	10-10	100 (0)
	H_F	BR_N	NA	-1	10	501	1-1	1-2	10-10	100(0)
	C_F	BR_N	NA	-1	10	501	1-1	2-2	10-10	100 (0)
RFC		BR_N	IPF	-1	10	501	1-1	1-1	10-10	100 (0)
RBFC		NBS	IPF	ND	32	1	1	1	21	65.625
SCC	A_C	BR_N	NA	1	10	501	1-1	2-2	10-10	100 (0)
	H_C	BR_N	NA	1	10	501	1-1	1-2	10-10	100 (0)
	C_C	BR_N	NA	1	10	501	1-1	2-2	10-10	100 (0)
RCC		BR_N	BE	1	10	501	1-1	1-1	9-10	99.18 (2.74)
RBCC		NBS	BE	ND	6	1	1	1	6	100
Total performance (TD_R^{XX}, $SD_{TD_R}^{XX}$)			Fine crackles (SFC, RFC and RBFC)				$TD_R^{FC} = 99.98$ %, $SD_{TD_R}^{FC} = 0.77$ %			
			Coarse crackles (SCC, RCC and RBCC)				$TD_R^{CC} = 99.80$ %, $SD_{TD_R}^{CC} = 1.42$ %			

SFC: Simulated fine crackles; A_F: IDW=0.7 ms & 2CD=5 ms (Charbonneau et al., 2000); H_F: IDW= 0.5 ms & 2CD= 3.3 ms (Hoevers & Loudon, 1990); C_F : IDW=0.9 ms & 2CD=6 ms (Cohen, 1990); RFC: Real fine crackles; SCC: Simulated coarse crackles; A_C: IDW= 1.5 ms & 2CD=10 ms (Charbonneau et al., 2000); H_C: IDW=1 ms & 2CD= 5.1 ms (Hoevers & Loudon, 1990); C_C: IDW=1.25 ms & 2CD=9.5 ms (Cohen, 1990); RCC: Real coarse crackles; RBFC: Real breath sound with fine crackles; RBCC: Real breath sound with coarse crackles; BN: Background noise; BR_N: Breath noise; IPF: Idiopathic pulmonary fibrosis; BE: Bronchiectasis; D_g: Diagnosis; SNR: Signal to noise ratio; NOTS: number of test samples; $\overline{D}_R$: Mean of rate of detectability; SD: Standard deviation; NA: Not applicable; ND: Not defined; N_R: Real number of crackles; N_E: Separated crackles; TD_R^{XX}: Total performance; $SD_{TD_R}^{XX}$: Standard deviation; XX stands FC for fine crackles and CC for coarse crackles; K and Q : number of iterations; min: Minimum value; max: Maximum value; In all cases number of samples N=32,768.

Table 5 Performance of WTST-NST filter in the case of fine and coarse crackles.

WTST-NST									
Cases		**BN**	**D_g**	**SNR (dB)**	**N_R**	**NOTS**	**K (min-max)**	**N_E (min-max)**	**$\bar{D}_R(SD)$ (%)**
SFC	A_F	BR_N	NA	-1	10	501	1-1	10-10	100 (0)
	H_F	BR_N	NA	-1	10	501	1-1	10-10	100 (0)
	C_F	BR_N	NA	-1	10	501	1-1	10-10	100 (0)
RFC		BR_N	IPF	-1	10	501	1-1	9-10	99.92 (0.89)
RBFC		NBS	IPF	ND	32	1	1	32	100 (0)
SCC	A_C	BR_N	NA	1	10	501	1-1	10-10	100 (0)
	H_C	BR_N	NA	1	10	501	1-1	10-10	100 (0)
	C_C	BR_N	NA	1	10	501	1-1	10-10	100 (0)
RCC		BR_N	BE	1	10	501	1-1	8-10	96.13 (5.38)
RBCC		NBS	BE	ND	6	1	1	6	100 (0)
Total performance ($TD_R^{XX}, SD_{TD_R}^{XX}$)			Fine crackles (SFC, RFC and RBFC)				$TD_R^{FC} = 99.98\ \%$, $SD_{TD_R}^{FC} = 0.45\ \%$		
			Coarse crackles (SCC, RCC and RBCC)				$TD_R^{CC} = 99.32\ \%$, $SD_{TD_R}^{CC} = 3.17\ \%$		

See Table 4 footnote for the caption for Table 5.

($\bar{E}_{QF_O}$) of the IEM-FD filter is much closer to the '1' compared to the WTST-NST filter (Hadjileontiadis & Panas, 1997) and the WT-FD filter (Hadjileontiadis, 2005(I); Hadjileontiadis, 2005(II)) in all cases of fine and coarse crackles, which indicates less over-estimation in the IEM-FD filter compared to the WTST-NST filter and the WT-FD filter. Furthermore, the performance comparison between the IEM-FD filter, the WTST-NST filter and the WT-FD filter in the case of RBCC is shown in Figure 35. The location of the crackles was audio-visually identified by an experienced pulmonary acoustics researcher and marked with arrowheads, so that the

Table 6 Performance of WT-FD filter in the case of fine and coarse crackles.

WT-FD									
Cases		**BN**	**D_g**	**SNR (dB)**	**N_R**	**NOTS**	**K (min-max)**	**N_E (min-max)**	**$\bar{D}_R(SD)$ (%)**
SFC	A_F	BR_N	NA	-1	10	501	1-1	10-10	100 (0)
	H_F	BR_N	NA	-1	10	501	1-1	10-10	100 (0)
	C_F	BR_N	NA	-1	10	501	1-1	10-10	100 (00
RFC		BR_N	IPF	-1	10	501	1-1	9-10	99.48 (2.22)
RBFC		NBS	IPF	ND	32	1	1	20	62.500
SCC	A_C	BR_N	NA	1	10	501	1-1	10-10	100 (0)
	H_C	BR_N	NA	1	10	501	1-1	10-10	100 (0)
	C_C	BR_N	NA	1	10	501	1-1	10-10	100 (0)
RCC		BR_N	BE	1	10	501	1-1	8-10	96.73 (5.18)
RBCC		NBS	BE	ND	6	1	1	6	100
Total performance (TD_R^{XX}, $SD_{TD_R}^{XX}$)			Fine crackles (SFC, RFC and RBFC)				$TD_R^{FC} = 99.85\ \%$, $SD_{TD_R}^{FC} = 1.41\ \%$		
			Coarse crackles (SCC, RCC and RBCC)				$TD_R^{CC} = 99.18\ \%$, $SD_{TD_R}^{CC} = 2.95\ \%$		

See Table 4 footnote for the caption for Table 6.

automatically separated crackles in the non-stationary part can be compared with the input signal crackles.

Figure 35-i displays a 0.743-second section of samples of RBCC recorded from a patient with BE (Table 4, case RBCC). The non-stationary and stationary parts after applying the IEM-FD filter are shown in Figure 35-ii (a) and (b), respectively. Comparing these results with the input signal, we can clearly see that not only are all the coarse crackles correctly separated with their morphology

Table 7 Performance of IEM-FD filter, WTST-NST filter and WT-FD filter in terms of underestimation.

							IEM-FD	**WTST-NST**	**WT-FD**
Cases		**BN**	**D_g**	**SNR (dB)**	**NOTS**	$\bar{R}_{QF_U}$ (SD)	$\bar{E}_{QF_U}$ (SD)	$\bar{E}_{QF_U}$ (SD)	$\bar{E}_{QF_U}$ (SD)
SFC	A_F	BR_N	NA	-1	501	0.817 (0.004)	0.817 (0.004)	0.817 (0.004)	0.817 (0.004)
	H_F	BR_N	NA	-1	501	0.848 (0.003)	0.848 (0.003)	0.848 (0.003)	0.848 (0.003)
	C_F	BR_N	NA	-1	501	0.801 (0.004)	0.801 (0.004)	0.801 (0.004)	0.801 (0.004)
RFC		BR_N	IPF	-1	501	0.828 (0.003)	0.839 (0.003)	0.830 (0.004)	0.829 (0.003)
SCC	A_C	BR_N	NA	1	501	0.704 (0.006)	0.705 (0.006)	0.726 (0.024)	0.704 (0.006)
	H_C	BR_N	NA	1	501	0.777 (0.004)	0.779 (0.005)	0.776 (0.004)	0.776 (0.004)
	C_C	BR_N	NA	1	501	0.710 (0.005)	0.711 (0.006)	0.736 (0.026)	0.710 (0.005)
RCC		BR_N	BE	1	501	0.731 (0.006)	0.747 (0.008)	0.735 (0.018)	0.732 (0.006)

$\bar{R}_{QF_U}$: Mean of reference underestimation quality factor; $\bar{E}_{QF_U}$: Mean of estimated underestimation quality factor. See Table 4 footnote for the remaining caption for Table 7.

and time location preserved, but also the normal breath sound is retained in the stationary estimate with its proper shape and amplitude. The non-stationary and stationary outputs of the WTST-NST filter for the same input signal are displayed in Figure 35-iii (a) and (b), respectively. Here we can observe that the non-stationary output (Figure 35-iii (a)) of the WTST-NST filter not only contains the crackles but also consists of a large part of the normal breath sounds, which represents over-estimation. The non-stationary and stationary outputs of the WT-FD filter in the case of the same input are shown in Figure 35-iv (a) and (b), respectively. From the results we can see that the non-stationary output (Figure 35-iv (b)) not only contains crackles but also a portion of the normal breath sounds due to over-estimation. However, in the stationary output (Figure 35-iv (b)) it can be observed that at the location of crackles, normal breath sounds segments are missing. As is mentioned earlier (section 4.5) for any crackle separation technique it is not only important to separate normal breath sounds from non- crackle location, but it is also important to separate normal breath sounds from crackle locations, to reveal the actual morphology of the crackles.

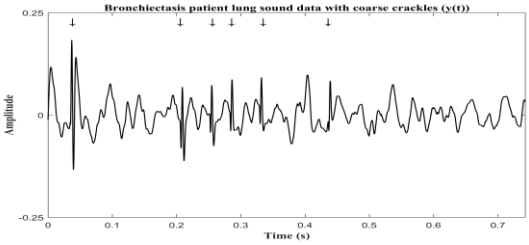
Table 8 Performance of IEM-FD filter, WTST-NST filter and WT-FD filter in terms of overestimation.

							IEM-FD	WTST-NST	WT-FD
Cases		**BN**	**D_g**	**SNR (dB)**	**NOTS**	$\bar{R}_{QF_O}$ (SD)	$\bar{E}_{QF_O}$ (SD)	$\bar{E}_{QF_O}$ (SD)	$\bar{E}_{QF_O}$ (SD)
SFC	A_F	BR_N	NA	-1	501	0.198 (0.005)	0.888 (0.032)	0.657 (0.156)	0.729 (0.036)
	H_F	BR_N	NA	-1	501	0.161 (0.004)	0.958 (0.015)	0.647 (0.172)	0.847 (0.032)
	C_F	BR_N	NA	-1	501	0.218 (0.006)	0.828 (0.036)	0.662 (0.159)	0.664 (0.033)
RFC		BR_N	IPF	-1	501	0.177 (0.005)	0.959 (0.008)	0.651 (0.161)	0.872 (0.014)
SCC	A_C	BR_N	NA	4	501	0.325 (0.011)	0.770 (0.036)	0.538 (0.199)	0.650 (0.038)
	H_C	BR_N	NA	4	501	0.236 (0.005)	0.932 (0.025)	0.671 (0.149)	0.854 (0.032)
	C_C	BR_N	NA	4	501	0.316 (0.010)	0.790 (0.033)	0.543 (0.203)	0.653 (0.036)
RCC		BR_N	BE	4	501	0.292 (0.010)	0.871 (0.033)	0.324 (0.079)	0.803 (0.041)

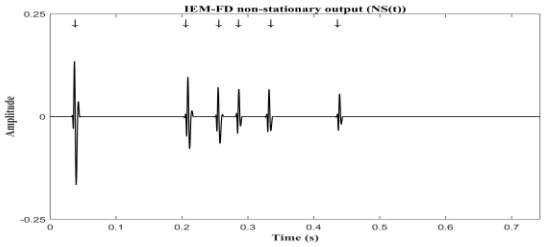
$\bar{R}_{QF_O}$: Mean of reference overestimation quality factor; $\bar{E}_{QF_O}$: Mean of estimated overestimation quality factor. See Table 4 footnote for the remaining caption for Table 8.

7.2.2.3 2CD percentage error

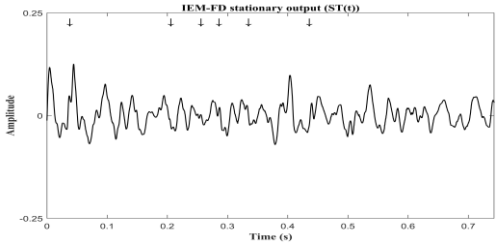
Table 9, Table 10, and Table 11 show the performance of the IEM-FD filter, WTST-NST filter and the WT-FD filter in terms of 2CD percentage error, respectively. In the IEM-FD filter (Table 9), the average 2CD percentage error is no more than 26.50 % for fine crackles (SFC and RFC) and less than 11 % for coarse crackles cases (SCC and RCC). On the other hand, in the WTST-NST filter (Table 10) the average 2CD percentage error is between 40 % to 122 % for fine crackles cases (SFC and RFC) and between 22 % to 54 % for coarse crackles (SCC and RCC) and in the WT-FD filter (Table 11) the average 2CD percentage error lies between 46 % and 88 % for fine crackles cases (SFC and RFC) and between 19 % and 59 % for coarse crackles (SCC and RCC). These results clearly indicate that the average 2CD percentage error is much less in the IEM-FD filter compared to the WTST-NST filter and the WT-FD filter in both fine and coarse crackles cases, which shows the IEM-FD filter better preserves the crackle morphology after separation compared to either of the other filters.



(i)

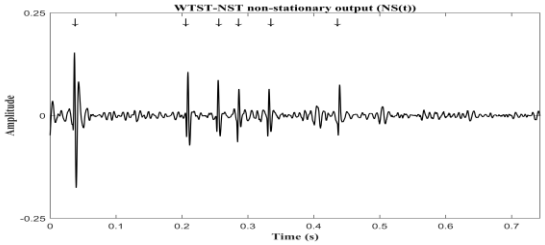


(a)

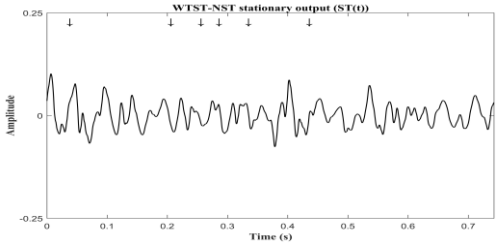


(b)

(ii)

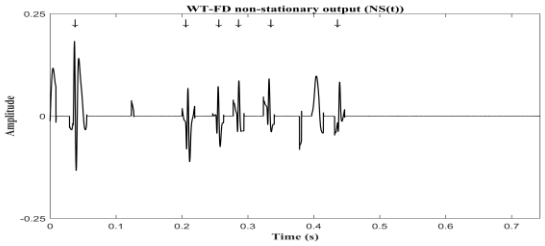


(a)

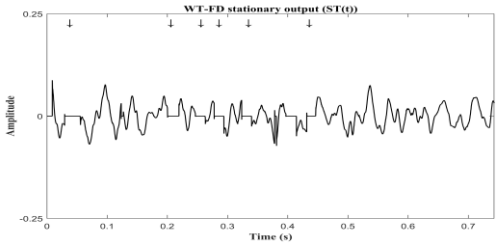


(b)

(iii)



(a)



(b)

(iv)

Figure 35 Comparison between crackle separation performance of the proposed IEM-FD filter, WTST-NST filter and the WT-FD filter; (i) Time section of 0.743 s of RBCC (Table 4, Case RBCC) recorded from a patient with BE. (ii) (a) IEM-FD filter non-stationary output; (b) IEM-FD filter stationary output. (iii): (a) WTST-NST filter non-stationary output; (b) WTST-NST filter stationary output. (iv): (a) WT-FD filter non-stationary output; (b) WT-FD filter stationary output.

Table 9 Performance of IEM-FD filter in terms of 2CD percentage error.

IEM-FD									
Cases		**BN**	**D_g**	**SNR (dB)**	**NOTS**	**NOC**	$\overline{AC}_{2CD}(SD)$ **(ms)**	$\overline{EC}_{2CD}(SD)$ **(ms)**	$\overline{PE}_{2CD}(SD)$ **(%)**
SFC	A_F	BR_N	NA	-1	501	5010	5 (0)	5.58 (0.51)	12 (10.06)
	H_F	BR_N	NA	-1	501	5010	3.3 (0)	4.16 (0.61)	26.50 (18.60)
	C_F	BR_N	NA	-1	501	5010	6 (0)	6.42 (0.40)	7.91 (5.94)
RFC		BR_N	IPF	-1	501	5010	3.52 (0.10)	3.45 (0.09)	2.37 (0.87)
SCC	A_C	BR_N	NA	1	501	5010	10 (0)	9.30 (0.51)	7.82 (3.31)
	H_C	BR_N	NA	1	501	5010	5.1 (0)	5.59 (0.51)	10.16 (10.05)
	C_C	BR_N	NA	1	501	5010	9.5 (0)	8.91 (0.39)	6.75 (2.83)
RCC		BR_N	BE	1	501	5010	8.41 (1.53)	8.32 (1.43)	10.07 (7.70)

NOC: Number of crackles (10 crackles in each test sample); $\overline{AC}_{2CD}$: Mean of actual crackles 2CD; $\overline{EC}_{2CD}$: Mean of estimated crackles 2CD; $\overline{PE}_{2CD}$: Mean of 2CD percentage error. See Table 4 footnote for the remaining caption for Table 9.

7.2.2.4 Computational cost

The computational cost for the FD technique is $2(N - W_{FD} + 1)[2(W_{FD} + L) + 1] + 4L + 1$ additions and $2(N - W_{FD} + 1)(W_{FD} + L + 2) + 8L$ multiplications (Hadjileontiadis & Rekanos, 2003), where, N is the number of samples in the input signal, W_{FD} is the fractal dimension window length and L is the maximum number of peeling levels in the FDPP algorithm. The MRD-MRR procedure requires $O(N\ log\ N)$ operations (Hadjileontiadis, 2005(II)): hence for the number of iterations *K* and the signal length of N, the MRD-MRR procedure will take $O(KN\ log\ N)$ operations in both the WTST-NST filter and the WT-FD filter. On the other hand, the IEM method requires only $O(QN)$ operations for the number of iterations *Q* and the signal length of N. This low computational cost of the IEM-FD filter can easily be achieved by an ordinary computer.

7.2.3 Comparison of the proposed IEM-FD filter with the EMD-FD filter

This section provides the performance comparison between the IEM-FD filter and the EMD-FD

Table 10 Performance of WTST-NST filter in terms of 2CD percentage error.

WTST-NST									
Cases		**BN**	**D_g**	**SNR (dB)**	**NOTS**	**NOC**	$\overline{AC}_{2CD}(SD)$ **(ms)**	$\overline{EC}_{2CD}(SD)$ **(ms)**	$\overline{PE}_{2CD}(SD)$ **(%)**
SFC	A_F	BR_N	NA	-1	501	5010	5 (0)	7.79 (5.45)	57.07(108.82)
	H_F	BR_N	NA	-1	501	5010	3.3 (0)	7.13 (8.93)	121.43(269.67)
	C_F	BR_N	NA	-1	501	5010	6 (0)	8.35 (6.38)	40.83 (106.06)
RFC		BR_N	IPF	-1	501	5010	3.52 (0.10)	5.14 (6.30)	48.85 (178.56)
SCC	A_C	BR_N	NA	1	501	5010	10 (0)	10.96 (4.07)	22.90 (35.14)
	H_C	BR_N	NA	1	501	5010	5.1 (0)	7.72(7.46)	53.89 (146.21)
	C_C	BR_N	NA	1	501	5010	9.5 (0)	10.19 (3.43)	23.50 (28.55)
RCC		BR_N	BE	1	501	5010	8.41 (1.53)	12.28 (4.50)	52.94 (51.72)

NOC: Number of crackles (10 crackles in each test sample). See Table 4 footnote and Table 9 footnote for the remaining caption for Table 10.

filter. Due to the EMD-FD filter's tendency to fail to converge on some of the 501 test samples in each case, one test sample was selected from 501 test samples at an SNR of -1 dB for RFC and each SFC case. Correspondingly, one test sample was selected from 501 test samples at an SNR of 1 dB for RCC and each SCC case. The crackle separation performance between the proposed IEM-FD filter and the EMD-FD filter was compared using these selected test samples supplemented by the RBCC and RBFC in terms of rate of detectability (D_R), total performance (TD_R^{XX}, $SD_{TD_R}^{XX}$), quality of crackle separation (over- or under-estimation), 2CD percentage error (PE_{2CD}) and separation time (S_T).

7.2.3.1 Rate of detectability and total performance

From the results in Table 12 it can be observed that the separation performance of the proposed IEM-FD filter and the EMD-FD filter (Hadjileontiadis, 2007) is quite similar in terms of rate of detectability except for a slightly lower performance of the IEM-FD filter in the case of RBFC. In terms of total performance, the proposed IEM-FD filter outperforms the EMD-FD filter in the case

Table 11 Performance of WT-FD filter in terms of 2CD percentage error.

WT-FD									
Cases		**BN**	**D_g**	**SNR (dB)**	**NOTS**	**NOC**	$\overline{AC}_{2CD}(SD)$ **(ms)**	$\overline{EC}_{2CD}(SD)$ **(ms)**	$\overline{PE}_{2CD}(SD)$ **(%)**
SFC	A_F	BR_N	NA	-1	501	5010	5 (0)	7.41 (2.25)	58.09 (31.97)
	H_F	BR_N	NA	-1	501	5010	3.3 (0)	5.87 (2.75)	87.68 (73.92)
	C_F	BR_N	NA	-1	501	5010	6 (0)	8.43 (1.88)	46.73 (21.60)
RFC		BR_N	IPF	-1	501	5010	3.52 (0.10)	5.64 (2.56)	63.85 (69.72)
SCC	A_C	BR_N	NA	1	501	5010	10 (0)	11.43 (2.60)	21.08 (21.12)
	H_C	BR_N	NA	1	501	5010	5.1 (0)	7.72 (2.10)	58.95 (30.45)
	C_C	BR_N	NA	1	501	5010	9.5 (0)	10.87 (2.35)	19.55 (21.15)
RCC		BR_N	BE	1	501	5010	8.41 (1.53)	11.11 (2.68)	36.71 (32.68)

NOC: Number of crackles (10 crackles in each test sample). See Table 4 footnote and Table 9 footnote for the remaining caption for Table 11.

of coarse crackles (SCC, RCC and RBCC) and achieves quite similar performance to the EMD-FD filter in the case of fine crackles (SFC, RFC and RBFC).

7.2.3.2 Quality of crackle separation (over or under estimation)

Table 13 shows the crackle separation comparison results of the proposed IEM-FD filter and the EMD-FD filter in terms of under-estimation and over-estimation, respectively. In terms of under-estimation, it can be noticed that the estimated under-estimation quality factor of the IEM-FD filter and the EMD-FD filter is either very close or equal to the reference under-estimation quality factor, which shows very low under-estimation. On the other hand, in terms of over-estimation, it can be noticed that the estimated over-estimation quality factor of the IEM-FD filter is very close to 1 compared to the EMD-FD filter, which indicates less over-estimation by the IEM-FD filter compared to the EMD-FD filter. Moreover, the performance comparison between the IEM-FD filter and the EMD-FD filter in the case of RFC (Table 12, case RFC) is shown in Figure 36. The location of the crackles before inserting into the background noise is marked with arrowheads, so that the automatically separated crackles in the non-stationary part can be compared with the

Table 12 Performance comparison of IEM-FD filter and EMD-FD filter in terms of rate of detectability and total performance.

							IEM-FD				EMD-FD		
Cases		**BN**	**D_g**	**SNR (dB)**	**N_R**	**NOTS**	**Q**	**K**	**N_E**	**D_R (%)**	**K**	**N_E**	**D_R (%)**
SFC	A_F	BR_N	NA	-1	10	1	2	1	10	100	1	10	100
	H_F	BR_N	NA	-1	10	1	2	1	10	100	1	10	100
	C_F	BR_N	NA	-1	10	1	2	1	10	100	1	10	100
RFC		BR_N	IPF	-1	10	1	1	1	10	100	1	10	100
RBFC		NBS	IPF	ND	32	1	1	1	21	65.6	1	29	90.6
SCC	A_C	BR_N	NA	1	10	1	2	1	10	100	1	10	100
	H_C	BR_N	NA	1	10	1	2	1	10	100	1	10	100
	C_C	BR_N	NA	1	10	1	2	1	10	100	1	10	100
RCC		BR_N	BE	1	10	1	1	1	10	100	1	8	80
RBCC		NBS	BE	ND	6	1	1	1	10	100	1	10	100
Total performance (TD_R^{XX}, $SD_{TD_R}^{XX}$)			Fine crackles (SFC, RFC and RBFC)				$TD_R^{FC} = 93.13\ \%$, $SD_{TD_R}^{FC} = 15.37\ \%$				$TD_R^{FC} = 98.12\%$, $SD_{TD_R}^{FC} = 4.20\%$		
			Coarse crackles (SCC, RCC and RBCC)				$TD_R^{CC} = 100\%$, $SD_{TD_R}^{CC} = 0$				$TD_R^{CC} = 96\ \%$, $SD_{TD_R}^{CC} = 8.94\ \%$		

See Table 4 footnote for the caption for Table 12.

input signal crackles. Figure 36-i displays the input signal with embedded RFC in breath noise. The non-stationary and stationary parts after applying the IEM-FD filter are shown in Figure 36-ii (a) and (b), respectively. Comparing these with the input signal, we can clearly see that not only are all the fine crackles separated from breath noise into the non-stationary signal estimate with their time duration and morphology preserved, but also the breath noise is retained in the stationary estimate with its proper shape and amplitude. On the other hand, the non-stationary and

Table 13 Performance comparison of IEM-FD filter and EMD-FD filter in terms of underestimation and overestimation.

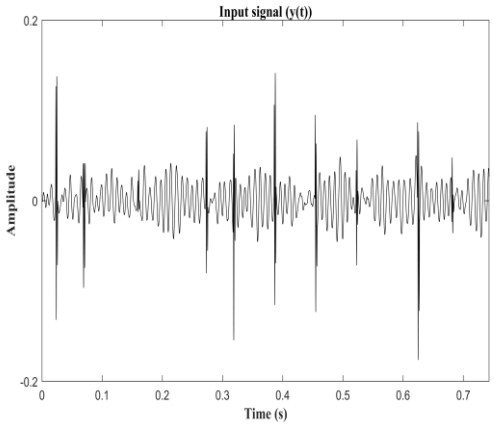
								IEM-FD		EMD-FD	
Cases		**BN**	**D_g**	**SNR (dB)**	**NOTS**	R_{QF_U}	R_{QF_O}	E_{QF_U}	E_{QF_O}	E_{QF_U}	E_{QF_O}
SFC	A_F	BR_N	NA	-1	1	0.808	0.205	0.808	0.832	0.808	0.620
	H_F	BR_N	NA	-1	1	0.849	0.159	0.849	0.967	0.849	0.322
	C_F	BR_N	NA	-1	1	0.804	0.206	0.804	0.851	0.842	0.689
RFC		BR_N	IPF	-1	1	0.857	0.147	0.862	0.970	0.858	0.379
SCC	A_C	BR_N	NA	1	1	0.716	0.326	0.716	0.771	0.716	0.439
	H_C	BR_N	NA	1	1	0.782	0.233	0.785	0.944	0.786	0.400
	C_C	BR_N	NA	1	1	0.710	0.308	0.710	0.761	0.751	0.436
RCC		BR_N	BE	1	1	0.725	0.300	0.743	0.943	0.822	0.738

R_{QF_U}: Reference underestimation quality factor; R_{QF_O}: Reference overestimation quality factor; E_{QF_U}: Estimated underestimation quality factor; E_{QF_O}: Estimated overestimation quality factor. See Table 4 footnote for the remaining caption for Table 13.

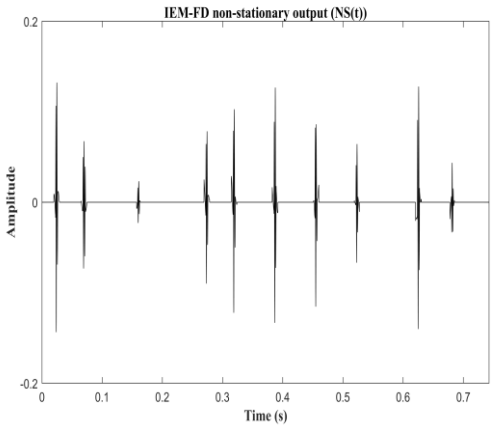
stationary outputs of the EMD-FD filter are displayed in Figure 36-iii (a) and (b), respectively. Here we can observe that the non-stationary output (Fig. 36-iii (a)) of the EMD-FD filter not only contains the crackles but also consists of a large portion of the breath noise due to over-estimation. These results clearly indicate that the IEM-FD filter provides better quality of crackle separation (less over-estimation) compared to the EMD-FD filter.

7.2.3.3 2CD percentage error

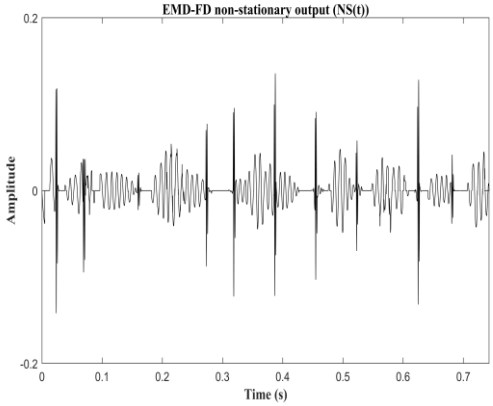
Table 14 and Table 15 show the performance comparison of the IEM-FD filter and the EMD-FD filter (Hadjileontiadis, 2007) in terms of 2CD percentage error, respectively. From the results it can be observed that in the IEM-FD filter the average 2CD percentage error is very much less than the average 2CD percentage error in the EMD-FD filter for both fine (SFC and RFC) and coarse (SCC and RCC) crackles. These results clearly indicate that the IEM-FD filter can better preserve the crackle morphology after separation compared to the EMD-FD filter.



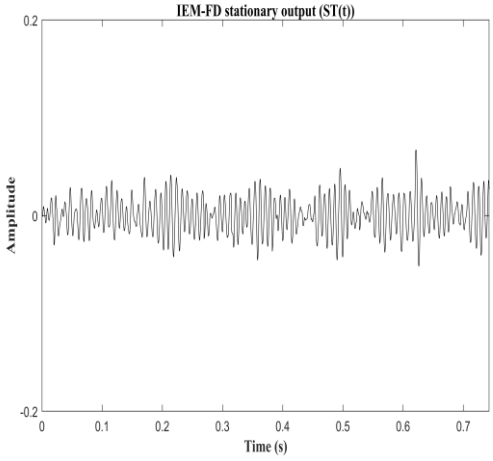
(i)



(a)

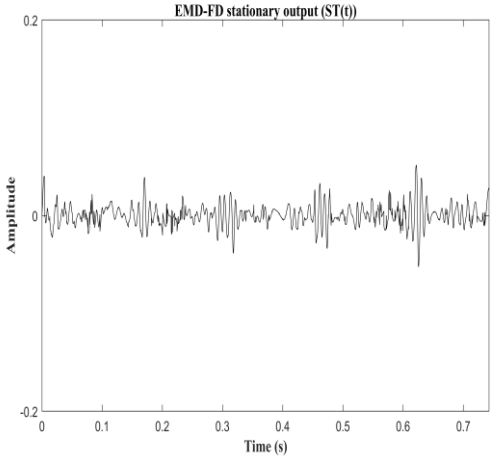


(a)



(b)

(ii)



(b)

(iii)

Figure 36 Comparison between crackle separation performance of the IEM-FD filter and the EMD-FD filter; (i) RFC (Table 12, Case RFC); (ii) (a) IEM-FD filter non-stationary output; (b) IEM-FD filter stationary output. (iii): (a) EMD-FD filter non-stationary output; (b) EMD-FD filter stationary output.

Table 14 Performance of IEM-FD filter in terms of 2CD percentage error.

								IEM-FD	
Cases		**BN**	**D_g**	**SNR (dB)**	**NOTS**	**NOC**	$\overline{AC}_{2CD}(SD)$ **(ms)**	$\overline{EC}_{2CD}(SD)$ **(ms)**	$\overline{PE}_{2CD}(SD)$ **(%)**
SFC	A_F	BR_N	NA	-1	1	10	5 (0)	5.76 (0.55)	15.82 (10.62)
	H_F	BR_N	NA	-1	1	10	3.3 (0)	3.79 (0.40)	15.31 (12.00)
	C_F	BR_N	NA	-1	1	10	6 (0)	6.54 (0.42)	9.74 (6.37)
RFC		BR_N	IPF	-1	1	10	3.52 (0.10)	3.44 (0.09)	2.57 (0.65)
SCC	A_C	BR_N	NA	1	1	10	10 (0)	9.34 (0.32)	6.61 (2.64)
	H_C	BR_N	NA	1	1	10	5.1 (0)	5.59 (0.53)	10.05 (10.38)
	C_C	BR_N	NA	1	1	10	9.5 (0)	8.80 (0.15)	7.13 (1.53)
RCC		BR_N	BE	1	1	10	8.41 (1.53)	8.31 (1.34)	12.24 (9.48)

Table 15 Performance of EMD-FD filter in terms of 2CD percentage error.

								EMD-FD	
Cases		**BN**	**D_g**	**SNR (dB)**	**NOTS**	**NOC**	$\overline{AC}_{2CD}(SD)$ **(ms)**	$\overline{EC}_{2CD}(SD)$ **(ms)**	$\overline{PE}_{2CD}(SD)$ **(%)**
SFC	A_F	BR_N	NA	-1	1	10	5 (0)	5.51 (1.05)	17.86 (14.59)
	H_F	BR_N	NA	-1	1	10	3.3 (0)	5.44 (2.85)	76.07 (76.31)
	C_F	BR_N	NA	-1	1	10	6 (0)	5.43 (0.88)	16.44 (3.41)
RFC		BR_N	IPF	-1	1	10	3.52 (0.10)	3.84 (0.59)	10.10 (15.78)
SCC	A_C	BR_N	NA	1	1	10	10 (0)	12.10(2.30)	25.41 (17.85)
	H_C	BR_N	NA	1	1	10	5.1 (0)	5.94 (1.47)	23.88 (22.91)
	C_C	BR_N	NA	1	1	10	9.5 (0)	9.01 (1.61)	14.57 (8.92)
RCC		BR_N	BE	1	1	10	8.41 (1.53)	7.65 (2.74)	21.04 (18.21)

NOC: Number of crackles (10 crackles in each test sample). See Table 4 footnote and Table 9 footnote for the remaining caption for Table 14 and Table 15.

Table 16 Performance comparison of IEM-FD filter and EMD-FD filter in terms of separation time.

						IEM-FD	**EMD-FD**
Cases		**BN**	**D_g**	**SNR (dB)**	**NOTS**	**S_T (s)**	**S_T (s)**
SFC	A_F	BR_N	NA	-1	1	0.57	215.42
	H_F	BR_N	NA	-1	1	0.56	189.63
	C_F	BR_N	NA	-1	1	0.57	201.27
RFC		BR_N	IPF	-1	1	0.49	42.38
RBFC		NBS	IPF	ND	1	0.43	45.29
SCC	A_C	BR_N	NA	1	1	0.58	149.98
	H_C	BR_N	NA	1	1	0.61	276.76
	C_C	BR_N	NA	1	1	0.55	261.95
RCC		BR_N	BE	1	1	0.55	38.44
RBCC		NBS	BE	ND	1	0.48	39.33
Average (S_T (s))						0.54	146.05

RBFC: Real breath sound with fine crackles; RBCC: Real breath sound with coarse crackles; S_T: Separation time; s: Second; NBS: Normal breath sound. See Table 4 footnote for the remaining caption for Table 16.

7.2.3.4 Separation time

As mentioned earlier (section 7.2.2.4), the low computational cost of the IEM-FD filter can easily be achieved. In this section, the crackle separation performance of the IEM-FD filter was tested in terms of separation time and compared with the EMD-FD filter. From Table 16 it can be noticed that in all the cases of fine and coarse crackles, the separation time of the IEM-FD filter is less than 1 second. By comparing the average crackle separation time of the IEM-FD filter (0.54 s) with the EMD-FD filter (146.05 s) it is clear that the IEM-FD filter is much faster than the EMD-FD filter.

7.3 Summary

Chapter 7 has presented test results for the proposed IEM-FD filter and a systematic comparison

with three previously published crackle separation techniques (WTST-NST filter, WT-FD filter and EMD-FD filter). It is clear that the IEM-FD filter can achieve the high rate of crackle identification accuracy with low computational cost. The comparative results indicate that the IEM-FD filter can provide a better quality of crackle separation (less over-estimation) and better preserved crackle morphology after separation compared to the previously published crackle separation techniques. Next the IEM-FD filter is applied to real data from two different datasets to explore its potential to for use on clinical sound recordings. Chapter 8 presents the detailed analysis of two different datasets using the IEM-FD filter.

Chapter 8 Two case studies

8.1 Introduction

This chapter will present the findings of two different case studies: one based on a year-long longitudinal study of patients with a diagnosis of IPF and the other a cross-sectional study of subjects who were referred for a HRCT scan of the chest for various clinical indications including suspected IPF. Both datasets were analysed using the IEM-FD filter. This chapter is divided into three main sections. Section 8.2 presents a brief introduction of IPF and the importance of Velcro crackles in IPF diagnosis or monitoring. The longitudinal study and the cross-sectional study are presented as case studies in section 8.3 and 8.5, respectively.

8.2 Idiopathic pulmonary fibrosis

IPF is the most common and the most life threatening idiopathic interstitial pneumonia (Bois, 2012), with a median survival time of only 3-5 years from the time of diagnosis (Oldham & Noth, 2014). The risk of IPF increases with age and is most common in the age group of 55-75 years with median diagnosis age of 66 years (King Jr et al., 2011). The main symptoms of IPF are chronic dry cough and exercise induced breathlessness (Meltzer & Noble, 2008). As mentioned in Lamas et al., (2011) the common median delay from symptom onset to diagnosis is 2.2 years. IPF is a non-curable disease (Wuyts et al., 2019); the only curative treatment available is lung transplantation (Purokivi et al., 2017), and the delay in IPF diagnosis may increase the rate of mortality or lower the survival rate regardless of disease severity (Lamas et al., 2011). Therefore, the early detection of IPF is more important especially with the more recent availability of the disease progression delaying drugs i.e. pirfenidone and nintedanib (Richeldi, 2016).

Velcro type crackles are considered as an early sign of IPF (Sellares et al., 2016). They are high pitched, discontinuous adventitious sounds, which can be heard in IPF patients in lung auscultation during slow and deep breathing. They consist of bursts of, sometimes overlapping, fine crackles which are non-musical and transient in nature and considered to be generated due to sudden opening of abnormally closed small airways (Cottin & Cordier, 2012). The sound of Velcro crackles is similar to the sound produced when strip of Velcro is gently detached from a blood pressure cuff (Cordier & Cottin, 2013; Aiello et al., 2017). Velcro crackles mostly appear in the mid to late inspiratory phase and are sometimes associated with the presence expiratory crackles (Sellares et al., 2016). They appear in the basal lung in the early stages of IPF and as the

disease progresses they start also to be generated in higher zones of the lung (Cottin & Cordier, 2012).

Chest HRCT is the most common clinical test for initial diagnosis of IPF, and surgical lung biopsy is used to confirm IPF diagnosis in patients with suspected IPF where HRCT presentation is not very promising for diagnosis (Cottin & Richeldi, 2014). However, in an ageing population lung biopsy is difficult to perform and without a biopsy the patient is left with suspected IPF (Cottin & Richeldi, 2014). Cottin & Cordier, (2012) suggested that on lung auscultation the assessment of Velcro crackles may provide a practical way to improve the earlier diagnosis of IPF. Although IPF is the most common interstitial lung disease and it is classed as a rare disease (defined as less than 1 case per 2000 people in the general population) so there can be a need for accessible indicators to use in diagnosis by GPs to ensure early referral.

Physicians can identify Velcro crackles in lung auscultation by listening to the lung sounds using a conventional stethoscope, but this method is very challenging especially in poor listening conditions (Chen et al., 2014). Moreover, the overlap of Velcro crackles with normal breath sounds makes it difficult to separate them from background normal breath sounds just by listening. Therefore, the accurate detection of Velcro crackles is highly dependent on a physician's expertise and hearing ability, which can lead to inter observer disagreement on crackle detection or to misdiagnosis (Chen et al., 2014). On the other hand, separation of crackles from normal breath sounds using an automatic crackle separation method may provide a better way of estimating number of crackles and their time domain features such as 2CD (as discussed in section 1.3). Therefore, in this chapter two different datasets are analysed using the IEM-FD filter and results obtained from both case studies are described below.

8.3 Longitudinal dataset analysis

Recently, Sgalla et al., (2019) showed that a set of acoustic features extracted from lung sounds recorded in IPF patients were a reproducible and valid metric for assessing disease severity and progression in patients with IPF. In this study, the lung sounds dataset recorded from 19 patients diagnosed with IPF was used for the analysis. For each patient, the lung sounds were recorded over a total of 7 visits (every 2 months) over a 1 year period. At each visit, the lung sounds were recorded at 10 chest locations (shown in Figure 37), identified using the guidelines of the computerized respiratory sound analysis (CORSA) study (Sovijarvi et al., 2000c), using a digital stethoscope. For each lung sound signal, a set of 481 acoustic features was estimated.

The original acoustic signal was preprocessed using a high pass filter with a cut-off frequency of 75 Hz. After preprocessing, the acoustic features were calculated using the MIR toolbox (Lartillot

& Toiviainen, 2007) from the filtered lung sound signal as a whole, from different IMFs, from the EMD 'crackle component' (first 3 or first 4 IMFs), and from the EMD 'respiratory component' (the sum of remaining IMFs). The IMFs were calculated using the EMD technique (described in section 4.6.1). Since for each sound file the number of IMFs was different, only the first 10 IMFs were used for the analysis (Sgalla, 2017). The features were generated from the energy content, from the statistical properties of the signal and its components (such as skewness and kurtosis), and from isolated frequency bands (such as cepstral features). Out of 481 acoustic features, a set 19 reproducible acoustic features were identified using intra subject reliability analysis. The intra subject reliability analysis was applied to three repeated recordings obtained during visit 1 from 4 IPF patients, and acoustic features with intra rater correlation coefficient (ICC) values greater than 0.5 were considered to have acceptable repeatability. Note that 6 of the 19 features were calculated from the sum of the first 3 or 4 IMFs estimated using the EMD technique, which the reference identifies as being associated with the crackle component of a lung sound when crackles are known to be present. The remaining 13 features were calculated from original acoustic signal which would include any crackle sounds present.

Out of 19 reproducible acoustic features, 6 features were found to significantly change between baseline and end of study measurements. These 6 features were then correlated with several clinically established parameters of disease progression such as forced vital capacity (FVC), and visual and computer-based evaluation of HRCT scans. Using multivariate regression analysis, it was found that the set of 6 features was significantly associated with the clinically established parameters of disease progression. Sgalla et al., (2019) concluded that acoustic features generated from lung sounds are a reproducible and valid measure for assessing disease severity in patients with IPF and suggested that quantitative analysis of lung sounds may provide a noninvasive clinical biomarker in IPF.

Although, Sgalla et al., (2019) looked at changes in the acoustic characteristics of the recorded lung sound, the acoustic features they selected were global features of the whole recorded signal, or of summed subsets of IMFs, which comprised both breath sound and any added sounds. They were not able to make a direct connection between the features they selected and the presence or characteristics of Velcro crackles. In our study, we extended the Sgalla et al., (2019) work to explore whether the reproducible acoustic features they identified were associate with a change in the NOC/BC over consecutive visits. Sovijarvi et al., (2000a) note that in interstitial lung diseases, the NOC/BC is associated with severity of the disease. Therefore, we investigated the relationship between the global acoustic features identified by Sgalla, (2017) and the NOC/BC extracted from the same set of lung sound data.

In the EMD technique, the selection of number of IMFs for crackle component and for respiratory component is very challenging because for every lung sound signal data dependent decision making is needed. Moreover, the EMD technique is a slow process and, in many cases, fails to converge so that the IMFs cannot be extracted, hence it is not ideal for a clinical setting where fast processing is desirable for decision making. Due to these limitations, we have only investigated the relationship between the 13 reproducible acoustic features generated from the whole signal after preprocessing and the NOC/BC. The process of calculating the NOC/BC is discussed below and the reproducible acoustic features are presented in section 8.3.7.

The aim of our study was to investigate the relationship between the NOC/BC with the reproducible acoustic features directly generated from the original lung sound signal after preprocessing.

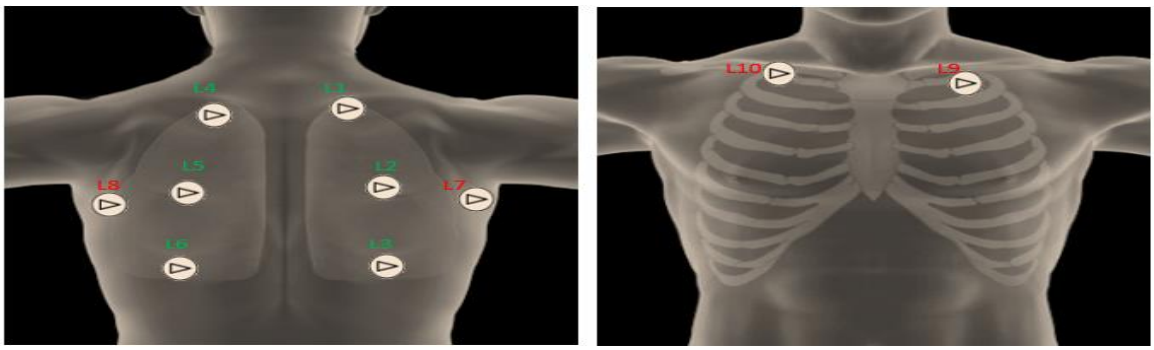


Figure 37 10 lung sounds recording sites (L1-L10), in this study lung sounds recorded from 6 posterior locations (L1-L6) are selected and these 6 posterior locations are shown in green.

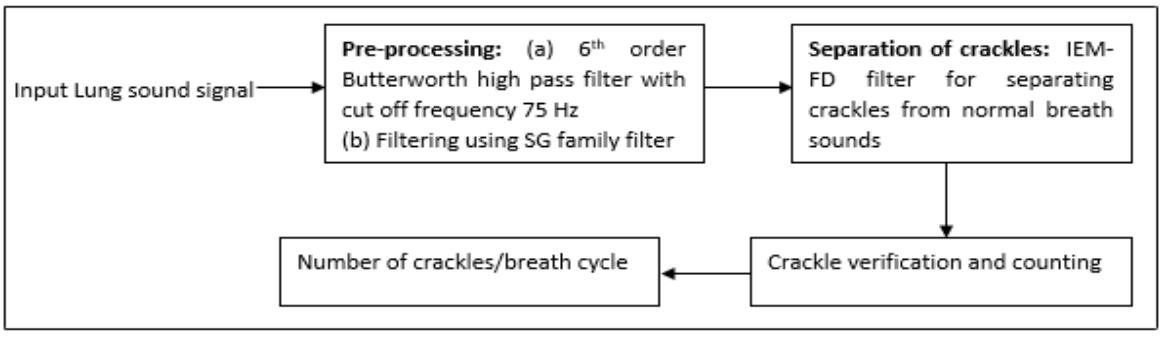


Figure 38 Steps used for analysing longitudinal dataset.

8.3.1 Audio-visual marking of breathing cycles

In each lung sound file the number of breathing cycles was audio-visually marked by the Author with the help of open access Audacity software. Only full breathing cycles were taken for the analysis as shown in Figure 39. Only lung sounds recorded from the 6 posterior locations: (as

shown in Figure 37 in green: L1-L6) were selected for further analysis. As mentioned in Cottin & Cordier, (2012) in the early course of IPF crackles appears in the basal areas of the lungs and as the disease progresses, they start to generate in upper zones therefore all 6 posterior locations were used for the analysis. Three patients died during the observation period and two patients withdrew from the study due to poor health. One patient completed the total 7 visits but missed one lung sound recording at location L1 (see Figure 37) in one of the 7 visits. Therefore, in total 689 lung sound files available from 19 patients were marked.

8.3.2 Pre-processing

Figure. 38 shows the number of steps used for analysing the IPF dataset. Each lung sound signal was pre-processed using a 6th order Butterworth high pass filter with cut off frequency 75 Hz (as used by Sgalla et al., 2019) and filtered using a SG filter to eliminating the high frequency peaks corresponding to the unwanted ripples in the signal. The SG filter parameters are: degree of fitting polynomial p_f = 4, number of coefficients $n_c = 89$, and order of derivation d_o = 0. Input lung sound signal and the pre-processed lung sound signal are shown in Figure 39 (a) and (b), respectively.

8.3.3 Crackle separation

After completing the pre-processing step, the IEM-FD filter is applied to each lung sound signal for separating crackles from normal breath sounds. Figure 39 (c) and (d) show an example of the stationary and non-stationary output of the IEM-FD filter (Chapter 5), respectively.

8.3.4 Crackle verification and counting

Once the crackles are separated from normal breath sounds, the absolute value of the non-stationary output is calculated (as shown in Figure 39 (e)). Now, a moving window of variable length containing 6 peaks of the absolute non-stationary (ANST) output (to guarantee the calculation of crackle 2CD, 5 zero crossings) is applied to verify and count crackles. Note that the process starts with the first 6 peaks of the ANST output. Inside a moving window, if all the 5 valleys between the 6 peaks represent the corresponding zero crossing of the non-stationary output of the IEM-FD filter, that window is considered as a potential crackle window (PCW) with location of the first peak as the beginning and location of the sixth peak as the ending of the PCW. The valley after the first peak of the PCW is considered the starting point of a potential crackle (PC). An example is shown in Figure 40 where all the valleys inside the 6 peaks (P1-P6) represent the corresponding zero crossing of the non-stationary output of the IEM-FD filter, hence, that

window is considered as PCW. On the other hand, inside a moving window, if all the 5 valleys between the 6 peaks do not represent the corresponding zero crossing of the non-stationary output of the IEM-FD filter, a new set of 6 peaks of the ANST output is calculated in the forward direction from the previous second peak of the moving window and the distance between new estimated first peak and sixth peak represents the new length of a moving window containing a new set of estimated peaks of the ANST output which will used to look for a PCW. Note that the presence of high frequency peaks corresponding to the unwanted ripples in the input lung sound signal may generate false peaks and valleys therefore in the pre-processing step the high frequency peaks corresponding to the unwanted ripples was minimised in the input lung sound signal using the SG filter.

After estimating a potential crackle window, two windows are calculated: before window (BW) and after window (AW). As shown in Figure 41 before and after windows are calculated, using 5 valleys just before and after the PCW and including first and last valley of the PCW, respectively. Note that in the case when before the potential crackle the number of valleys was less than 5, the value of the BW mean was considered equal to zero as shown in Figure. 41, where before window (BW1) mean for the first crackle (C1) is zero. Similarly, after the potential crackle when number of valleys was less than 5, the value of the after-window mean was considered zero. The reason for selecting only 5 valleys of the ANST output for the BW and AW is to make the length of BW and AW as close as possible to the PCW, which also contains 5 valleys of the ANST output. The idea behind this is to minimize the chances of identifying false crackles without losing true crackles. For a longer BW and AW, we may lose true crackles and for shorter BW and AW we may increase false crackle detection. Now, each potential crackle is verified using the criteria defined by Murphy et al., (1989) and Pinho et al., (2015), and additional conditions empirically generated by the Author:

Pre-existed criteria:

(1) The beginning of the event had a sharp deflection in either a negative or a positive direction (Murphy et al., 1989; Kaisla et al., 1991).

(2) After the IDW, the baseline crossings of all other peaks have to be progressively wider (Murphy et al., 1989; Kaisla et al., 1991). As mentioned by Pinho et al., (2015), not all crackles follow the standard rules, therefore ± 50 % deviation of width of the peaks is empirically selected to verify this condition.

(3) The IDW of the potential crackle must be 1/8 times greater than the LDW (Pinho et al., (2015).

Empirically established conditions generated by the author:

(4) The amplitude of the LDW peak of the crackle has to be greater than all other peaks of a crackle.

(5) The PCW mean has to be greater than 1.2 times the BW mean.

(6) The PCW mean has to be greater than the AW mean. Note that the condition (4) and (5) not only help to count temporally overlapping crackles but also make sure that any crackle with more than 5 zero crossings is not considered for more than once.

(7) The amplitude of the IDW peak and LDW peak of the potential crackle must be greater than one peak before the IDW peak (amplitude of the first peak of the PCW).

(8) The 2CD of the potential crackle has to be less than 20 ms.

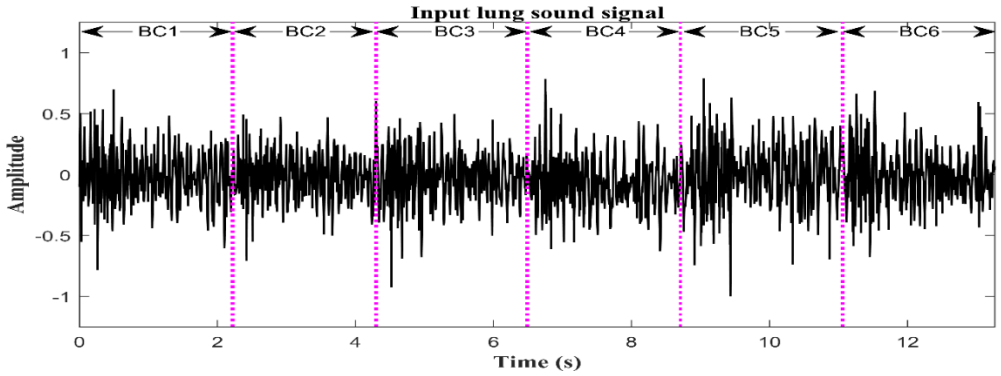
(9) The IDW of the potential crackle has to be less than 3 ms.

Note that if crackle failed to follow any of the above conditions the PC was not counted as a true crackle. A new set of 6 peaks of the ANST output is identified in the forward direction from the second peak of the failed PCW. The distance between the new estimated first peak and sixth peak represent the new length of a moving window containing a new set of estimated 6 peaks of the ANST output which are used to look for a new PCW. On the other hand, if PC meets all the conditions the PC is considered as a true crackle. The process is then repeated starting in the forward direction from the first peak after the old PCW. Note that the length of the moving window varies according to the location of the new estimated 6 peaks of the ANST output.

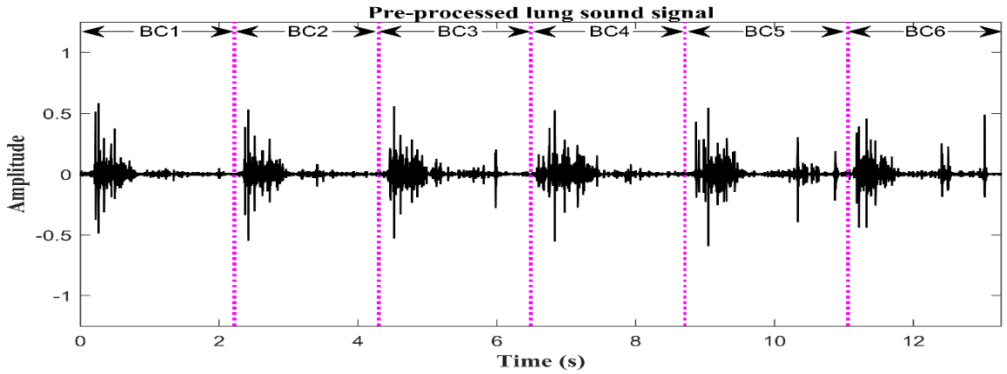
8.3.5 Number of crackles/breath cycle

Figure 42 (a) shows a lung sound signal with identified true crackles. The non-stationary output of the IEM-FD filter and its absolute value are shown in Figure 42 (b) and 42 (c), respectively. To get a better visualization of the automatically detected crackles, the fourth breath cycle (BC4) in Figure 42 is shown separately in Figure 43.

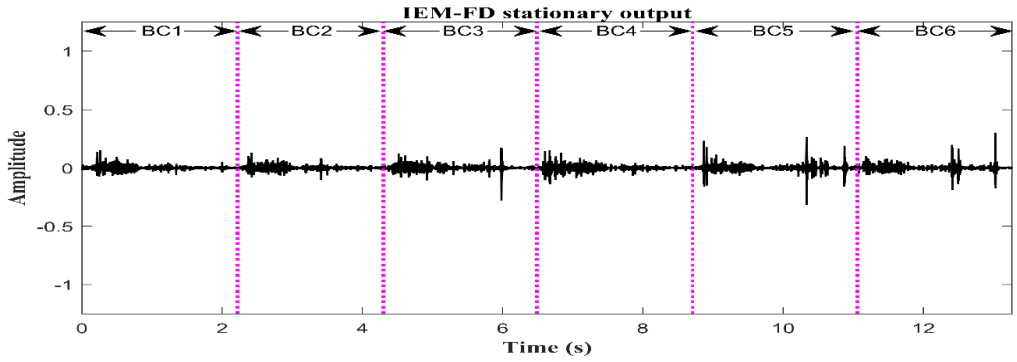
Note that in this study the crackle detection method is not validated in terms of sensitivity or positive predictive value because of the unavailability of a reference signal to compare with. Therefore, it may be possible that some of the true crackles are missed and that some false crackles are detected. This may be due to true crackles occurring which do not meet established crackle verification conditions or because the IEM-FD filter fails to separate all the crackles from the input lung sound signal. However, it is important to notice that: (a) the crackle detection


Input lung sound signal
BC1
BC2
BC3
BC4
BC5
BC6
Amplitude
Time (s)

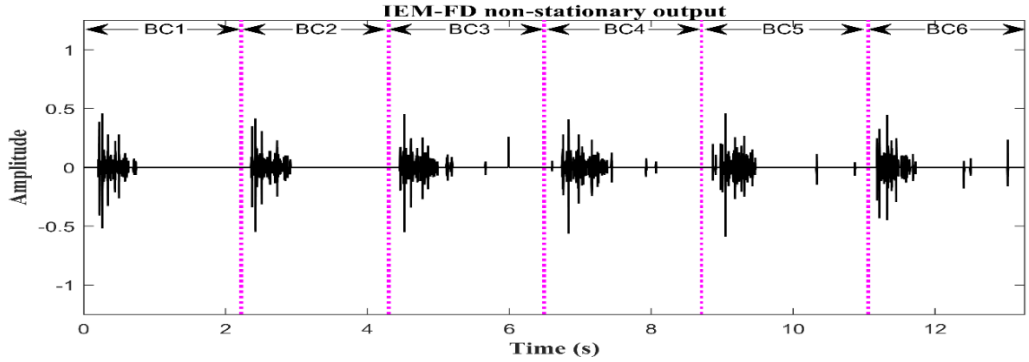
(a)


Pre-processed lung sound signal
BC1
BC2
BC3
BC4
BC5
BC6
Amplitude
Time (s)

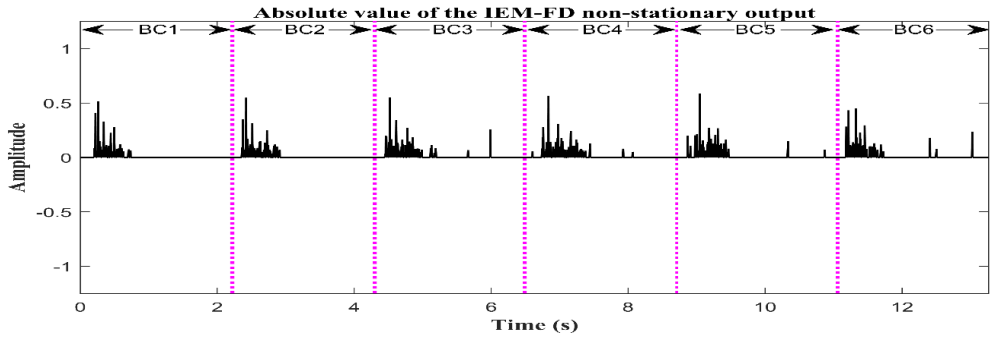
(b)


IEM-FD stationary output
BC1
BC2
BC3
BC4
BC5
BC6
Amplitude
Time (s)

(c)


IEM-FD non-stationary output
BC1
BC2
BC3
BC4
BC5
BC6
Amplitude
Time (s)

(d)



(e)

Figure 39 (a) Input lung sound signal; (b) Pre-processed input lung sound signal; (c) IEM-FD filter stationary output; (d) IEM-FD filter non-stationary output; (e) Absolute value of the IEM-FD filter non-stationary output. Note that here BC: Breath cycle.

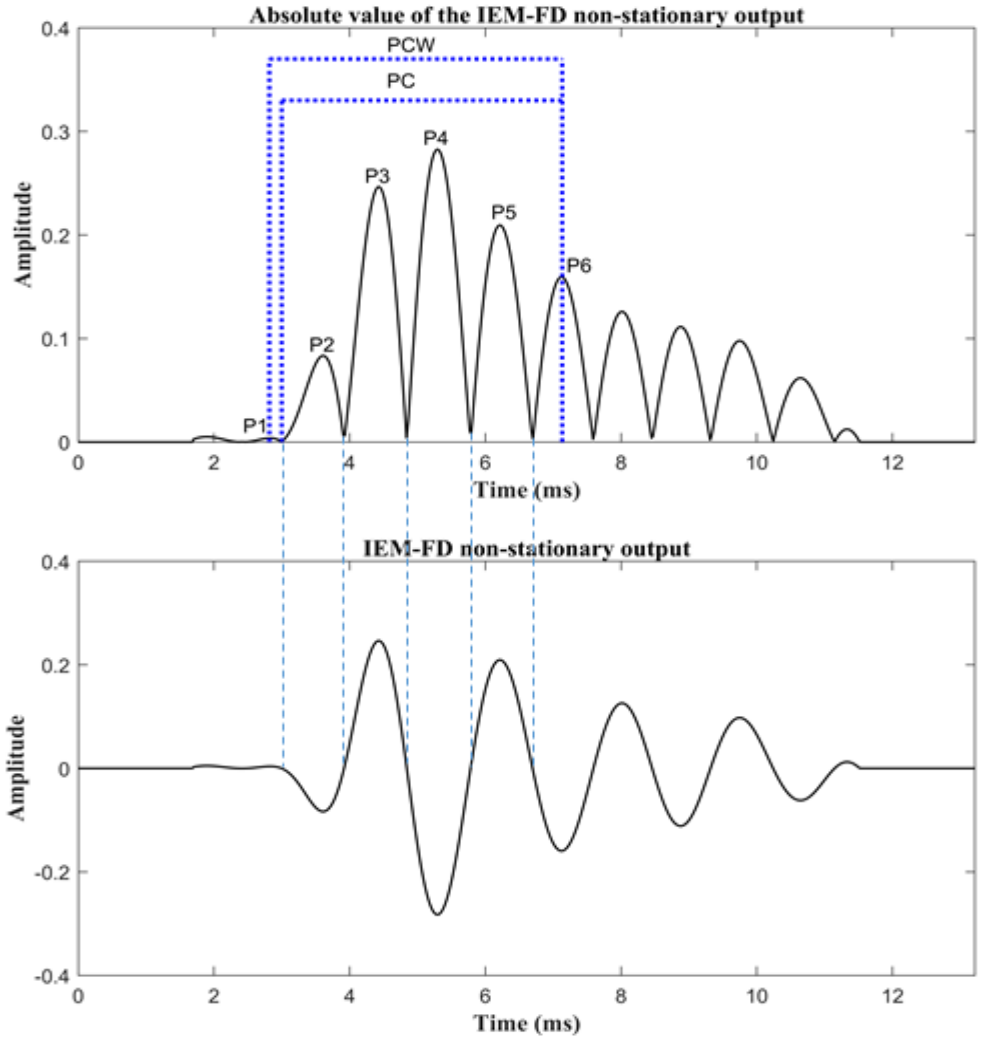


Figure 40 A PCW with PC and six estimated peaks (P1-P6), where it is shown that all the valleys inside the six peaks of the potential crackle window are corresponding to the zero crossing of the IEM-FD filter non-stationary output.

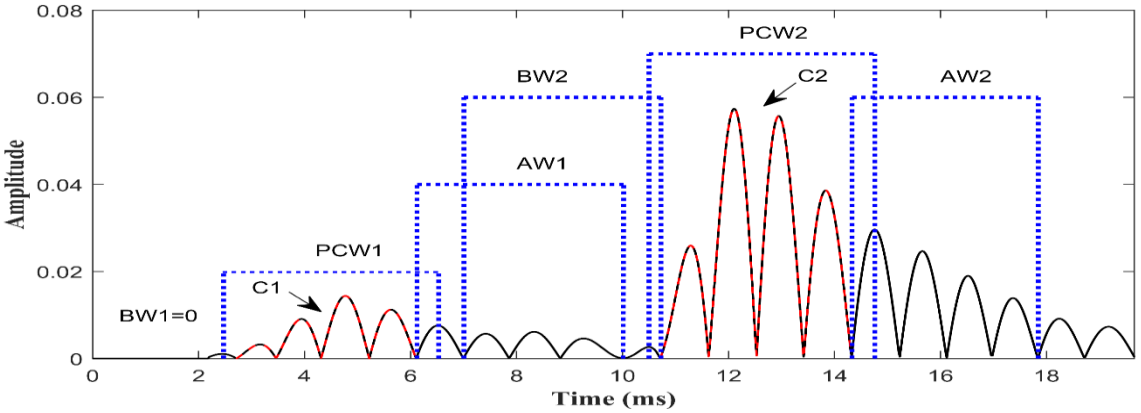


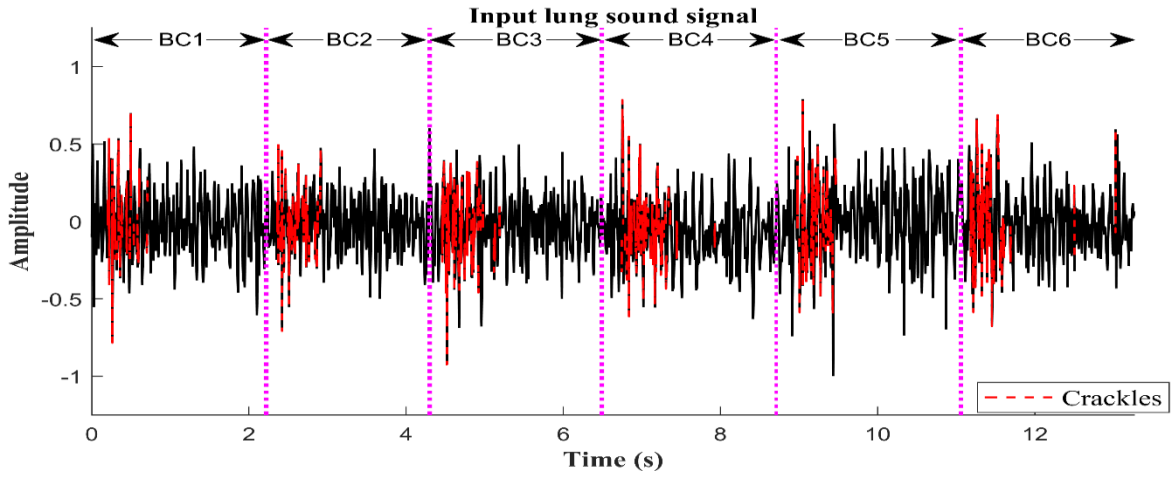
Figure 41 Absolute value of the IEM-FD non-stationary output with PCW, BW and AW. Note that here BW1: Before window of the first crackle (C1); PCW1: Potential crackle window of the first crackle (C1); AW1: After window of the first crackle (C1); C1: First crackle; BW2: Before window of the second crackle (C2); PCW2: Potential crackle window of the second crackle (C2); AW2: After window of the second crackle (C2); and C2: Second crackle. All the valleys shown in the absolute value of the IEM-FD non-stationary output were corresponding to the zero crossing of the IEM-FD filter non-stationary output.

method is not directly applied to the input lung sound signal rather it is applied to the absolute non-stationary output of the IEM-FD filter after separating background normal breath sounds from it, which reduces the chances of detecting false crackles; and (b) most of the crackle verification conditions used here were already tested in previous studies (Kaisla et al., 1991; Pinho et al. 2015; Murphy et al., 1989).

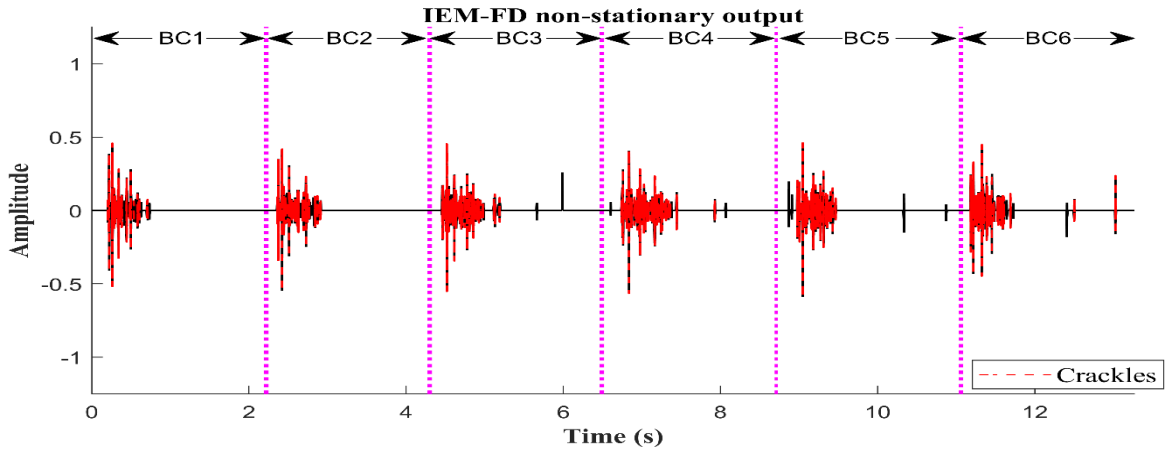
After automatically extracting crackles from the IPF dataset, in each lung sound signal the NOC/BC was calculated. Next the change in extracted NOC/BC over the 7 visits (12 months follow up, visit every 2 months) was analysed. Finally, the estimated NOC/BC was correlated with the 13 reproducible acoustic features directly estimated from the original signal after pre-processing. All the analyses were made using SPSS (IBM SPSS Statistics 26) and MATLAB (R2019a). For all the statistical analysis, statistical significance was set at $p < 0.05$.

8.3.6 Number of crackles/breath cycle

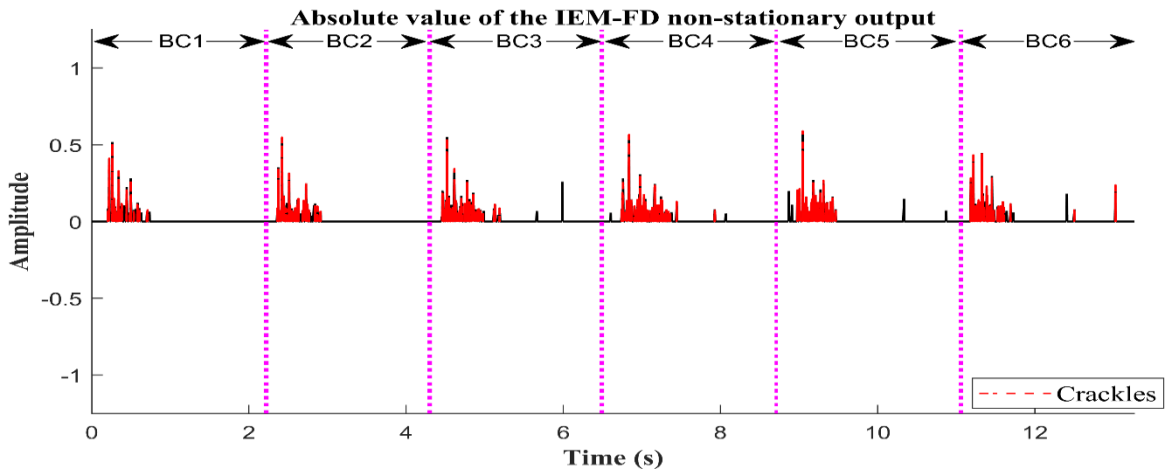
At each visit the number of observations was different because of patient drop-out. Therefore, the analysis was performed using 13 available observations across all 7 visits. The estimated mean NOC/BC averaged over all 13 available subjects and all locations for each of the 7 visits to clinic is reported in Table 17. Note that at each visit the NOC/BC calculated from all six posterior locations was used for the analysis.



(a)



(b)



(c)

Figure 42 (a) Input lung sound signal with detected crackles; (b) IEM-FD filter non-stationary output with detected crackle; (c) Absolute value of the IEM-FD filter non-stationary output with detected crackles. Here BC: Breath cycle.

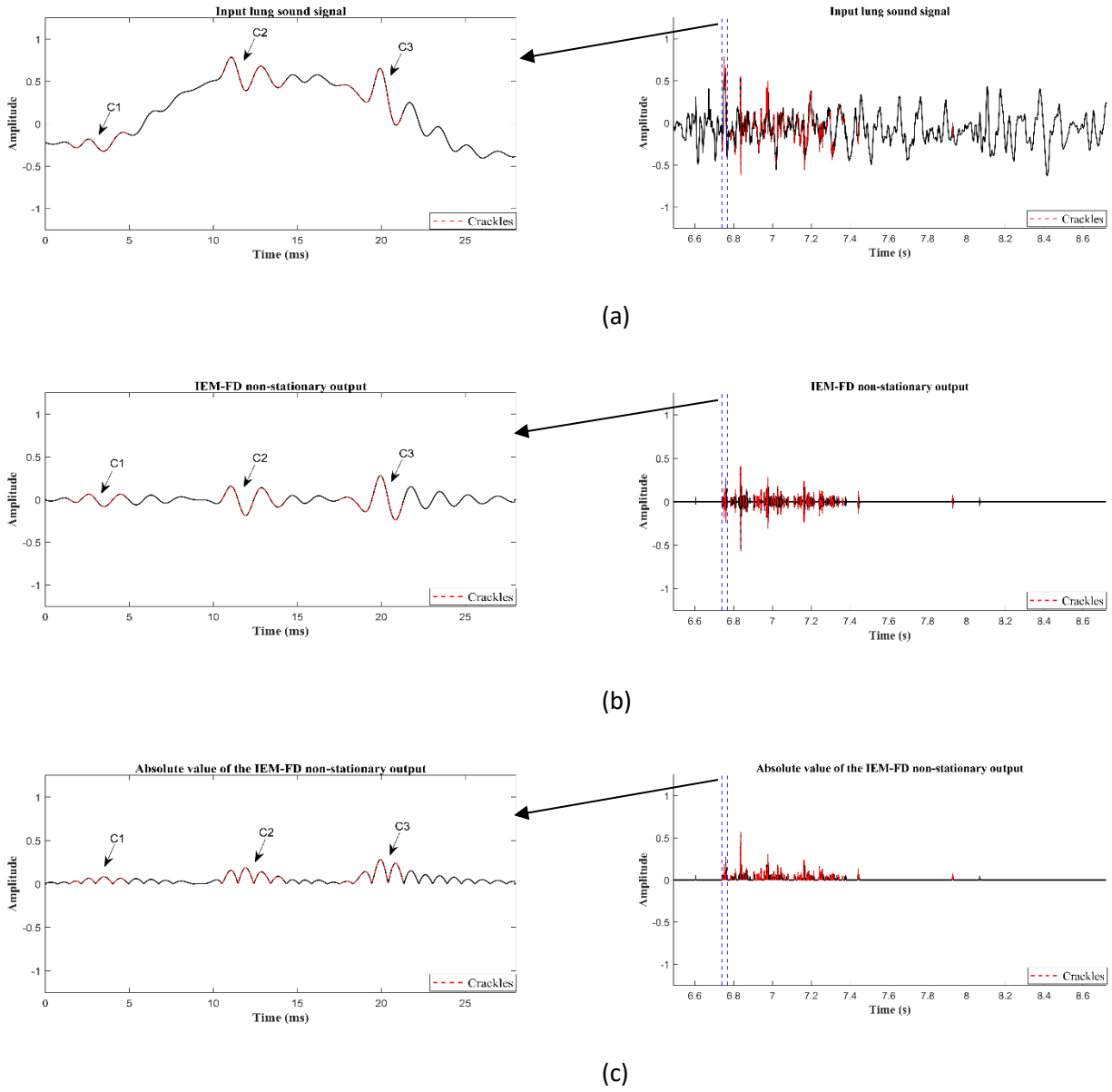


Figure 43 One breathing cycle of the lung sound signal (a) Input lung sound signal with detected crackles; (b) IEM-FD filter non-stationary output with detected crackle; (c) Absolute value of the IEM-FD filter non-stationary output with detected crackles.

The normality of the data over 7 visits was tested using Shapiro-Wilk test; the complete results obtained are shown in Table 18 and for visual analysis of the data samples distribution the histogram plots are reported in Appendix A (A1). Due to the non-normal distribution of the data samples in all the visits, non-parametric tests were selected here for the statistical analysis. The Friedman test was used to test whether the mean NOC/BC for different visits are all the same (within statistical limits) or not and the Wilcoxon signed-rank test was used to compare the difference between pairs of visits; a Bonferroni correction was used for correcting the statistical significance level where multiple statistical tests were made.

Table 17 Estimated NOC/BC using 13 available observations across 7 visits in the IPF dataset. Data presented as mean, standard deviation, maximum and minimum values.

Visits	Parameter	Subjects	Minimum	Maximum	Mean	Std. Deviation
Visit 1	NOC/BC	13	1.67	76.50	29.19	18.86
Visit 2	NOC/BC	13	1.33	75.33	26.56	18.57
Visit 3	NOC/BC	13	0.83	67.00	24.80	15.60
Visit 4	NOC/BC	13	0.14	86.50	28.38	20.83
Visit 5	NOC/BC	13	0.57	65.75	22.27	15.24
Visit 6	NOC/BC	13	0.60	71.67	20.62	15.56
Visit 7	NOC/BC	13	0.75	84.00	24.09	17.09

The change in the mean NOC/BC over the complete 12 months follow up study was significant (*Chi-Square*=19.060, *p*=0.004) under the Friedman test. The pairwise comparison between visit 1 and visit 6 was found to be significant (*Z*=3.481, *p*=0.021) under the Wilcoxon signed-rank test. The full pairwise comparison over 7 visits using the Wilcoxon signed-rank test is reported in Appendix A (A2).

Table 18 Shapiro-Wilk test for normality of the distribution of the NOC/BC across 7 visits using 13 available observations in the IPF group.

Visits	Parameter	Subjects	Statistic	*p* value
Visit 1	NOC/BC	13	0.950	0.004
Visit 2	NOC/BC	13	0.939	0.001
Visit 3	NOC/BC	13	0.953	0.006
Visit 4	NOC/BC	13	0.935	0.001
Visit 5	NOC/BC	13	0.944	0.002
Visit 6	NOC/BC	13	0.924	0.000
Visit 7	NOC/BC	13	0.924	0.000

8.3.7 Reproducible acoustic features in IPF dataset

Table 19 shows the 19 reproducible acoustic features (Sgalla, 2017; Sgalla et al., 2019). In these 19 features, 13 features were directly estimated from the original acoustic signal after pre-processing and the remaining 6 features were calculated from first 3 or 4 IMFs. The data in terms of mean, standard deviation, maximum value, and minimum values of all 19 reproducible acoustic features from sound files recorded at 6 posterior locations over complete study of 1 year in the IPF dataset (Sgalla, 2017) is reported in Appendix A (A3).

Table 19 A set of 19 reproducible acoustic features (Sgalla, 2017).

A set 19 reproducible acoustic features	**Description**
C3 EW_200_500 Hz	Crackle component (number of IMFs=3) energy weight (EW) in the frequency range of 200-500 Hz
C4 EW_75_200 Hz	Crackle component (number of IMFs=4) energy weight (EW) in the frequency range of 75-200 Hz
C4 EW_200_500 Hz	Crackle component (number of IMFs=4) energy weight (EW) in the frequency range of 200-500 Hz
sig_zerocross	Zero-cross of the original signal
sig_mfcc02	Mel Frequency Cepstral Coefficient (MFCC) of the original signal
sig_75_200 Hz_zerocross	Zero-cross of the original signal in the frequency range of 75-200 Hz
sig_75_200 Hz_centroid	Centroid of the original signal in the frequency range of 75-200 Hz
sig_200_500 Hz_rms	RMS of the original signal in the frequency range of 200-500 Hz
sig_200_500 Hz_lowenergy	Low Energy of the original signal in the frequency range of 200-500 Hz
sig_200_500 Hz_lowenergyASR	Average silence ratio of the original signal in the frequency range of 200-500 Hz

sig_200_500 Hz_zerocross	Zero-cross of the original signal in the frequency range of 200-500 Hz
sig_200_500 Hz_std_meanframes	Mean of the frames of the original signal in the frequency range of 200-500 Hz
sig_200_500 Hz_std_medianframes	Median of the frames of the original signal in the frequency range of 200-500 Hz
sig_500_1000 Hz_zerocross	Zero-cross of the original signal in the frequency range of 500-1000 Hz
sig_500_1000 Hz_rolloff85	Roll-off (threshold=85 %) of the original signal in the frequency range of 500-1000 Hz
sig_500_1000 Hz_centroid	Centroid of the original signal in the frequency range of 500-1000 Hz
C3_mfcc02	Mel Frequency Cepstral Coefficient (MFCC) of the crackle component (number of IMFs=3)
C4_zerocross	Zero-cross of the crackle component (number of IMFs=4) in the specified frequency range
C4_mfcc02	Mel Frequency Cepstral Coefficient (MFCC) of the crackle component (number of IMFs=4)

8.3.8 Correlation of NOC/BC and 13 reproducible acoustic features generated from original signal

In this section, the estimated NOC/BC was correlated with the 13 reproducible acoustic features estimated directly from the original signal after pre-processing. Note that in this study the NOC/BC and the acoustic features measured at 6 posterior recording locations over 7 visits in the IPF dataset were used for the analysis. The NOC/BC showed significant correlation with 11 acoustic features out of 13 in the univariate analysis as shown in Table 20.

After the univariate analysis, simple linear regression was also performed for all the reproducible acoustic features that showed significant correlation with NOC/BC. The R^2 values of the linear regression model are reported in Table 21. The R^2 values show the amount of variation in the dependent acoustic features that can be explained by the NOC/BC. In the regression model, the

Table 20 Univariate correlation analysis between NOC/BC and acoustic features measured at 6 posterior locations over 7 visits in the IPF dataset.

	Number of crackles per breath cycle	
13 reproducible Acoustic features generated from original signal	**Pearson correlation coefficient**	***p* value**
sig_zerocross	0.628**	0.000
sig_mfcc02	0.185**	0.000
sig_75_200 Hz_zerocross	0.214**	0.000
sig_75_200 Hz_centroid	0.151**	0.000
sig_200_500 Hz_rms	0.300**	0.000
sig_200_500 Hz_lowenergy	-0.373**	0.000
sig_200_500 Hz_lowenergyASR	-0.430**	0.000
sig_200_500 Hz_zerocross	0.590**	0.000
sig_200_500 Hz_std_meanframes	0.242**	0.000
sig_200_500 Hz_std_medianframes	0.347**	0.000
sig_500_1000 Hz_zerocross	-0.103**	0.007
sig_500_1000 Hz_rolloff85	0.026	0.487
sig_500_1000 Hz_centroid	-0.050	0.194

** = correlation was significant at the 0.01 level.

strongest relationship of the NOC/BC were found with sig_zerocross and sig_200_500 Hz_zerocross, which account for 39.4 % and 34.8 % variability of the features, respectively.

8.4 Discussion

In clinical practice, FVC via spirometry is used as a feasible and reliable tool for assessing

Table 21 R^2 values of linear regression models for NOC/BC showing significant correlation with the acoustic features generated from original signal.

Independent variable	**Reproducible Acoustic features generated from original signal (Dependent variables)**	**R^2 value**
NOC/BC	sig_zerocross	0.394
	sig_mfcc02	0.034
	sig_75_200 Hz_zerocross	0.046
	sig_75_200 Hz_centroid	0.023
	sig_200_500 Hz_rms	0.090
	sig_200_500 Hz_lowenergy	0.140
	sig_200_500 Hz_lowenergyASR	0.185
	sig_200_500 Hz_zerocross	0.348
	sig_200_500 Hz_std_meanframes	0.059
	sig_200_500 Hz_std_medianframes	0.121
	sig_500_1000 Hz_zerocross	0.011

deterioration in IPF patients (Sgalla et al., 2019). Sgalla, (2017) in his study presented longitudinal measurements of % predicted FVC to determine whether the sample population experienced disease progression over the duration of study. Sgalla, (2017) found statistically significant evidence of disease progression, which showed at decline in % predicted FVC over the study period. However, the mean change in % predicted FVC was not statistically significant in any pairwise comparison between visits. Furthermore, Sgalla, (2017) considered the longitudinal change in reproducible acoustic features in IPF to see which reproducible features might be useful for assessing progression of the disease. He found that out of 19 reproducible acoustic features 15 reproducible features underwent the significant change over the period of the 12-month study. Moreover, out of these 15 reproducible acoustic features 6 showed significant change between baseline and end of study, 3 features showed significant change between baseline and visit 6, and one feature showed significant change between baseline and visit 5. Additionally, when these 15 reproducible acoustic features were correlated with the % predicted FVC; 9 reproducible acoustic features were significantly correlated with the % predicted FVC, 6 at the 0.01 significant level and

3 at the 0.05 significant level. However, the correlation was negligible (Pearson correlation coefficient: 0 to 0.3 or 0 to -0.3), indicating that the reproducible acoustic features were poorly responsive to change in % predicted FVC.

In our study, we extended the Sgalla et al., 2019 work. Firstly, we calculated the NOC/BC from each lung sound file in the longitudinal dataset (see sections 8.3.1-8.3.5). After calculating NOC/BC, the change in NOC/BC over the 12 months of study period was examined because, as mentioned in Sovijarvi et al. (2000a) the NOC/BC is associated with disease severity in patients with interstitial lung diseases. A statistically significant change in mean NOC/BC was found over the study period. This change may be due to disease progression, however, the pairwise comparisons between visits were not found to have a statistically significant differences. This is may be because two-thirds of the IPF patients were taking disease-modifying treatments (either pirfenidone or nintedanib) over the course of the study. Nevertheless, a significant mean change in NOC/BC was observed between baseline and visit 6.

Secondly, this study found a relationship between reproducible acoustic features generated from original signal and NOC/BC. Out of 13 reproducible acoustic features generated from original signal, 11-showed significant correlation with the NOC/BC at p = 0.01 significance level. However, most of these features showed either negligible (Pearson correlation coefficient: 0 to 0.3 or 0 to -0.3) or low (Pearson correlation coefficient: 0.3 to 0.5 or -0.3 to -0.5) correlation with the NOC/BC. Nevertheless, 2 features: zero-cross of the original signal (sig_zerocross) and zero-cross of the original signal in the frequency range of 200-500 Hz (sig_200_500 Hz_zerocross) were moderately correlated (Pearson correlation coefficient: 0.5 to 0.7 or -0.5 to -0.7) with the NOC/BC indicating the properties of these two acoustic features were more responsive to the change in NOC/BC compared to the other acoustic features considered. These findings were further supported by linear regression analysis, which showed a high predictive power for these two acoustic features (sig_zerocross and sig_200_500 Hz_zerocross) towards change in NOC/BC compared to the other reproducible acoustic features. Interestingly, these two acoustic features also underwent a significant change over the 12 months of study duration and showed significant correlation with the % predicted FVC at the 0.01 significant level (Sgalla et al., 2019). Note that the Pearson correlation coefficient was interpreted using the guidelines provided by Mukaka, (2012).

The purpose of this study was to investigate whether the acoustic features generated from the original signal were related to the NOC/BC. This study shown that two specific reproducible acoustic features directly generated from the original lung sound signal correlate with NOC/BC.

In clinics, lung sounds can easily be recorded using an electronic stethoscope and the estimation of these acoustic features after preprocessing of the recorded lung sounds may provide an

efficient tool for assessing disease severity in IPF patients and, together with other established measures of disease severity, they can help in clinical decision making. The results of this analysis support the Sgalla et al., (2019) finding that the reproducible acoustic features can be used for assessing disease severity in patients with IPF.

It is important notice that, in our study we only focused on reproducible acoustic features generated from original signal and did not consider the sub-signals generated by combining IMFs, for the reasons given in section 8.3. In addition, this analysis was made using a single centre dataset collected from a small population, therefore future research should explore the link between the NOC/BC and the global acoustic features of the signal on larger multicenter IPF dataset.

8.5 Cross sectional dataset analysis

The aim of this study was to develop an automatic system, which can be able to differentiate IPF patients from patients with other types of lung pathology. Sgalla et al., (2018) showed that the presence of Velcro crackles is directly correlated with the extent of distinct radiologic features of pulmonary fibrosis on HRCT and suggested that auscultation for Velcro crackles may help in early detection of fibrotic lung disease. In this study, 254 subjects who were referred for HRCT scan of the chest for various clinical indications were considered. Just before the participants underwent HRCT, lung sounds were recorded using an electronic stethoscope (3M Littmann 3200) at 6 posterior chest locations (as shown in Figure 44) based on the guidelines of CORSA (Sovijarvi et al., 2000c). Two radiologists reviewed the imaging data for the presence and the extent of fibrotic abnormalities in the lungs. 76 participants were identified as having fibrosis on the HRCT and from the remaining participants 72 age and sex matched subjects indicating no signs of fibrosis on the HRCT were selected. Therefore, in the final study population 148 participants were enrolled. The recorded lung sounds were assessed by two expert interstitial lung disease physicians for the presence of Velcro crackles and in the case of disagreement, the lung sound file was marked as having Velcro crackles present. The assessment of the Velcro crackles was then correlated with different radiologic features of pulmonary fibrosis at HRCT scans. Sgalla et al., (2018) concluded that Velcro crackles not only predict the presence of fibrotic interstitial lung disease patterns at HRCT but also closely correlate with the extent of different radiological features of pulmonary fibrosis.

The above-mentioned study showed that the Velcro crackles in the recorded lung sounds may facilitate the early detection of fibrotic lung disease (Sgalla et al., (2018)). Furthermore, as mentioned in Chapter 1, computer based lung sound analysis may provide an objective way of

analysing recorded lung sounds. Therefore, in our study, we developed an automatic system, which can be used for differentiating IPF patients from patients with other types of pathology based on the average NOC/BC. The automatic system was tested on the cross sectional dataset used in Sgalla et al.'s, (2018) study and compared with the assessment of the expert physicians. Note that in our study, the dataset recorded from 55 available participants was used for the analysis and only lung sounds recorded from the 4 lung bases: (as shown in Figure 44 in green: L2, L3, L5 and L6) were used for further analysis. As mentioned in (Cottin & Cordier, 2012) at the early stage of IPF crackles appear in the lower lung bases. The dataset consists 28 patients diagnosed with IPF and 27 patients with non-IPF. The full HRCT images marked for the presence or absence of pulmonary by the two expert radiologists from the (Sgalla et al., 2018) study were used as the ground truth for our study.

The main aim of the study was to investigate whether automatic system performs as well as expert assessment for differentiating IPF patients from patients with other types of pulmonary pathology.

As for the longitudinal study, in this analysis the author audio-visually marked the breath cycles in each lung sound file with the help of open access Audacity software. Only full breathing cycles were taken for the analysis. Due to inaudibility of breath cycles, 4 patients from the IPF group and 3 patients from the non-IPF group was excluded from further analysis. Since the Author was not able to mark number of breath cycles in these recordings. Therefore, in total 185 available lung sounds recorded from 24 fibrosis and 24 non-fibrosis groups were analysed. Table 22 shows the age and sex of the subjects in the fibrosis and the non-fibrosis groups. In the fibrosis group the mean age of the patients was 69.83 with 15 males and 9 females. In the non-fibrosis group, the mean age of the patients was 69.79 with 15 males and 9 females. Males were predominant in both fibrosis and non-fibrosis groups and the fibrosis and non-fibrosis groups were age and sex matched.

Table 22 Characteristics of the fibrosis and non-fibrosis groups. Data are presented as the mean (standard deviation) and counts (%).

		Fibrosis (No. of patients=24)	**Non fibrosis (No. of patients=24)**
Age, years		69.83 (9.24)	69.79 (10.54)
Sex (%)	Male	15 (62.5%)	15 (62.5%)
	Female	9 (37.5%)	9 (37.5%)

8.5.1 Estimation of average number of crackles per breath cycle

In this study, the same steps were used for calculating the NOC/BC from each lung sound signal (see Figure 38) as in longitudinal study (section 8.3). Using the NOC/BC for all available lung sound files, the average NOC/BC was estimated for each patient by taking the mean of NOC/BC over all 4 lower lung base recording locations.

Note that if one or more of the 4 lower lung base recordings was not present for any patient, the average NOC/BC was calculated by taking the mean of NOC/BC over those lung sound recordings which were available for that patient. Box and Whisker plots of the estimated average NOC/BC in the fibrosis and non-fibrosis groups are displayed in Figure 45. The average NOC/BC for each patient in the fibrosis and non-fibrosis groups with their age and sex is reported in Appendix A (A4).

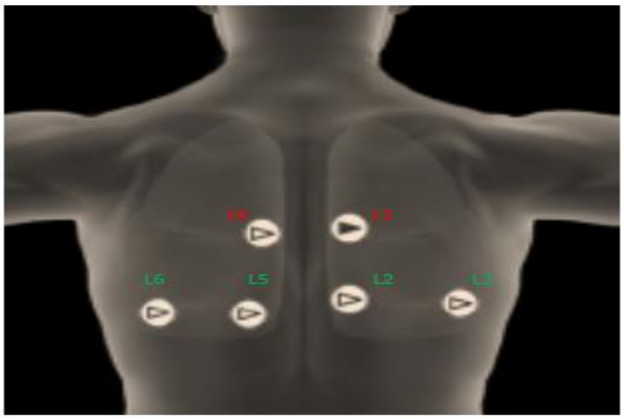


Figure 44 6 lung sounds recording sites (L1-L6), in this study lung sounds recorded from 4 posterior base locations (L2, L3, L5 and L6) are selected and these 4 posterior base locations are shown in green.

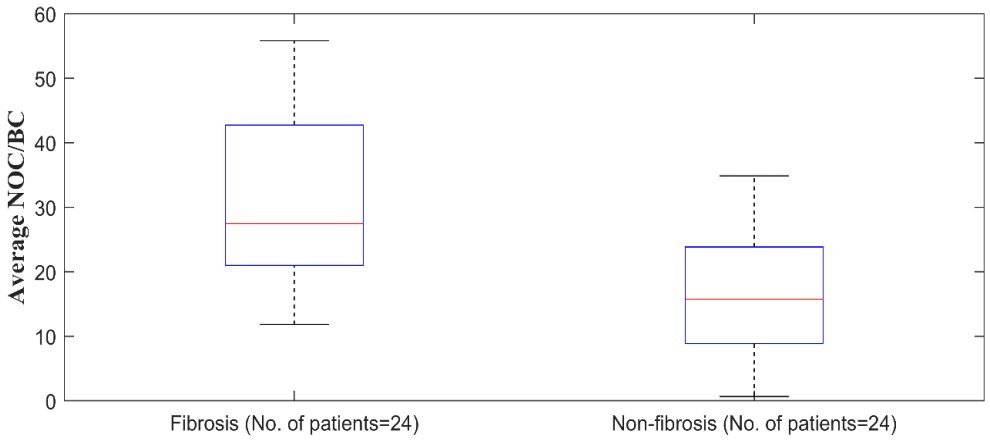


Figure 45 Estimated Box and Whisker plots of average NOC/BC in fibrosis and non-fibrosis group.

The Shapiro-Wilk test was used for analysing the distribution of average NOC/BC for each group and for visual analysis the histogram plots of the distribution of average NOC/BC are reported in Appendix A (A5). The normality test was non-significant for the fibrosis group (p =0.236) and for the non- fibrosis group (p = 0.320) hence the parametric, independent sample t-test was used for comparing the difference of average NOC/BC between the two groups. The difference in average NOC/BC between two groups was significantly different (t=4.94, p<0.001) at the independent sample t-test, as shown in Table 23.

Table 23 Independent sample t-test for average NOC/BC at fibrosis and non-fibrosis groups.

		t-test for Equality of Means						
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Average NOC/BC	Equal variances not assumed	4.938	43.202	0.000	15.047	3.047	8.903	21.192

8.5.2 Receiver operating characteristic curve

A receiver operating characteristic (ROC) curve is a plot of a test's sensitivity (plotted on the vertical axis), versus its false positive rate or 1-specificity (plotted on the horizontal axis) (Obuchowski, 2005). Sensitivity evaluates the proportion of positives correctly classified; specificity calculates the proportion of negatives correctly classified, and false positive rate is one minus specificity (Flach, 2010). In our study, to see the potential of the average NOC/BC to differentiate two groups (fibrosis or non-fibrosis) a ROC curve was generated, as shown in Figure 46. The estimated ROC curve is reported in Table 24. The cut-off value for average NOC/BC, 18.62 (AUC=0.845; sensitivity=91.7; specificity=59.3) was taken to differentiate two groups. The crackles in IPF patients are usually generated in the inspiratory phase and Flietstra et al., (2011) shown that the mean (standard deviation) number of crackles in the inspiratory phase in IPF patients was 18 (14), which supports the selection of 18.62 as a cut off average NOC/BC.

8.5.3 Velcro crackles assessment

In Sgalla, (2017) study, two physicians independently listened to each recorded lung sound file using open access Audacity software and over-ear-headphones (Sennheiser HD201 closed

Table 24 ROC curve for the average NOC/BC. Data presented as area under the curve with 95% confidence interval and p-value.

		95% CI	
Area under the curve (AUC)	**p-value**	**Lower bound**	**Upper bound**
0.845	0.000	0.739	0.952

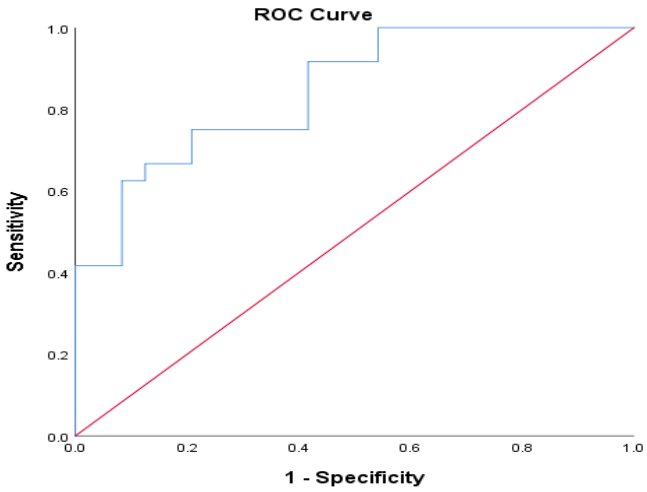


Figure 46 ROC curve for the average NOC/BC.

dynamic stereo) and marked for Velcro crackles ("1= Velcro crackles", "0=no Velcro crackles"). In case of disagreement, a file was marked as Velcro crackles. Where at least one physician identified Velcro crackles as present in at least one recording site for a given lung base (left or right) that lobe was identified as having Velcro crackles present. If at least one lung base had Velcro crackles, the patient was classified as having unilateral Velcro crackles. Patients with no Velcro crackles identified in either lobe were placed in the no unilateral Velcro crackles category. The physician assessment was then compared to the evidence from the HRCT scans.

Out of 48 patients, physician 1 heard unilateral Velcro crackles at lung sounds recorded from 30

Table 25 Patients with unilateral Velcro crackles assessed by two physicians. Data presented in counts and percentages (%).

	Physician 1	Physician 2
No unilateral Velcro crackles	18 (37.5%)	17 (35.42%)
Unilateral Velcro crackles	30 (62.5%)	31 (64.58%)

Table 26 Cross tabulation for inter rater agreement of evaluation of unilateral Velcro crackles at recorded lung sounds for individual patients. Data are expressed as counts.

		Physician 2		
		No unilateral Velcro crackles	Unilateral Velcro crackles	Total
Physician 1	No unilateral Velcro crackles	12	6	18
	Unilateral Velcro crackles	5	25	30
Total		17	31	48

patients and physician 2 reported unilateral Velcro crackles in the lung sounds recorded from 31 patients, as presented in Table 25. The inter-rater reliability of assessment of unilateral Velcro crackles between two physicians was moderate (Cohen's kappa= 0.446, p=0.001), as shown in cross tabulation in Table 26.

8.6 Results and Discussion

The performance of physician 1 and physician 2 in the differentiation of patients with fibrosis and non-fibrosis based on their evaluation of the presence of unilateral Velcro crackles is reported using cross tabulation in Table 27 and Table 28, respectively. The results in terms of sensitivity and specificity are shown in Table 29. Both physicians had equal sensitivity (83.3 %), which showed their ability to identify the presence of Velcro crackles matched well to the diagnosis of IPF via the gold standard of HRCT. On the other hand, specificity was 58.3 % and 54.2 % for physician 1 and physician 2, respectively. When we compare the average or individual performance of each physicians with the average NOC/BC at the selected 18.65 cut-off value it can be noticed that the average NOC/BC achieved higher sensitivity (91.7 %) and specificity (59.3 %) compared to average for either physician individually or for both on average, as shown in Table 29.

The aim of the study was to develop an automatic system, which can be able to differentiate IPF patients from patients with other types of pathology. The results indicate that the automatic system presented here can perform as well as the assessment of expert physicians in terms of differentiating patients with IPF from patients with other types of pathology based on the presence of Velcro crackles. Moreover, the developed system has shown strong potential for diagnostic support, especially for assisting non-expert clinicians in their auscultatory assessment of lung sounds.

Table 27 Cross tabulation for physician 1 in the recognition of fibrosis based on unilateral Velcro crackles assessment. Data are expressed as counts.

		No fibrosis	Fibrosis	Total
Physician 1	No unilateral Velcro crackles	14	4	18
	Unilateral Velcro crackles	10	20	30
	Total	24	24	48

Table 28 Cross tabulation for physician 2 in the recognition of fibrosis based on unilateral Velcro crackles assessment. Data are expressed as counts.

		No fibrosis	Fibrosis	Total
Physician 2	No unilateral Velcro crackles	13	4	17
	Unilateral Velcro crackles	11	20	31
	Total	24	24	48

Table 29 Performance of two physicians in the identification of fibrosis at HRCT scan using assessment of unilateral Velcro crackles in recorded lung sounds and their comparison with selected average NOC/BC cut-off value using sensitivity and specificity. Data presented as percentages (%).

Method	Sensitivity (%)	Specificity (%)
Physician 1	83.3	58.3
Physician 2	83.3	54.2
Average	83.3	56.25
Average NOC/BC (18.65)	91.7	59.3

8.7 Summary

In this chapter, two case studies: a longitudinal study and a cross sectional study of patients with lung pathologies are presented. Both studies were analyzed using an automated crackle detection process with separation by the IEM-FD filter as a first step.

In the longitudinal study, the extracted NOC/BC were correlated with global acoustic features

directly extracted from the original acoustic signal and it was found that some (but not all) of these features were directly associated with NOC/BC, hence might be useful for monitoring progression of IPF.

In the cross-sectional study, the IEM-FD filter formed the basis of an automatic system based on the average NOC/BC calculated from posterior base locations of a set of patients. Average NOC/BC so detected was shown to match the specificity and exceed the sensitivity of the individual and average performance of two expert physicians in terms of differentiating IPF patients from non-IPF patients when compared the gold standard HRCT. Hence, the automated system may assist physician's in evaluating their auscultatory findings of lung sounds in clinics.

In the next Chapter, the detailed working process of a new automatic breath cycle detection algorithm is presented.

Chapter 9 An automatic breath cycle detection method based on the estimation of the breathing phases

9.1 Introduction

In the previous chapter we saw the results of two case studies analysed using the IEM-FD filter. In both studies the breath cycles were audio-visually marked by the author in each lung sound file. The manual marking of breath cycles is highly time consuming, therefore an automatic breath cycle detection method was developed based on the estimation of the breathing phases. Each breath cycle contains two phases: inspiratory phase and expiratory phase. First the algorithm estimates the location of the breathing phases and then the estimated breathing phases are used for calculating the number of breath cycles. The rest of the chapter is organized as follows: Section 9.2 discuss why there is a need of a new algorithm for estimating the breathing phases. Section 9.3 presents the detailed working process of the new automatic method. The dataset and the quantitative evaluators used for performance analysis are discussed in section 9.4 and 9.5, respectively. The experimental results are shown in section 9.6 and evaluation of the method is discussed in section 9.7.

9.2 Why there is a need of a new algorithm for estimating breathing phases

The detection of breathing phases or breathing cycles is essential for the analysis of lung sounds in relation to cardiorespiratory diseases. As mentioned in Piirila & Sovijarvi, (1995) and Sovijarvi et al., (2000a) the timing of crackles in the breathing cycle (early/mid/late inspiratory or expiratory) may have clinical significance for the differential diagnosis of cardiorespiratory disorders. Moreover, in patients with interstitial lung disorders the NOC/BC is associated with the severity of the disease (Sovijarvi et al., 2000a; Epler et al., 1978). Generally, direct airflow measurement using a pneumotachograph is used for estimating breathing phases (Tarrant et al., 1997). The simultaneously measured airflow may be combined with counted crackles on recorded lung sounds for estimating timing of crackles in breathing phases. However, this approach demands a complex set up, and hence is not suitable for use in clinical practice (Jacome et al., 2019). On the other hand, the audio-visual marking of breath cycles on each recorded lung sound file (as was done for the two case studies described in Chapter 8) is very time consuming and subjective. Therefore, automatic detection of breathing phases (inspiration and expiration) may provide a

more efficient way of estimating breathing phases which may in turn help in counting the number of breath cycles.

In recent years, researchers have proposed several methods to detect breathing phases directly from recorded lung sounds such as: automatic breathing phase detection using acoustical means (Chuah & Moussavi, 2000), trachea breath sounds based acoustic breath phase detection method (Huq & Moussavi, 2012), and automatic breath phase classification using smartphones (Reyes et al., 2016). The breath phase detection accuracy of these methods lies in the range of 93 % to 100 %. These methods had high accuracy for breath phase detection but were very depended on tracheal sound. At the trachea, sounds are easily heard during both inspiratory and expiratory phases (Bohadana et al., 2014). On the other hand, breath sounds measured on the chest wall are more audible in the inspiratory phase and almost silent in the expiratory phase. However, tracheal auscultation is not typically performed to monitor cardiorespiratory diseases. It is mainly used to display the status of the upper airways (Bohadana et al., 2014). Furthermore, these methods were validated using lung sounds recorded from healthy subjects and the breathing phase detection accuracy of these methods was not tested on lung sounds recorded from patients with respiratory diseases. It is known that the breathing pattern and lung sounds change in the presence of respiratory diseases (Dellweg et al., 2008; Todd et al., 2018). Therefore, it is possible that these methods may not perform as well as reported in the literature when applied to lung sounds recorded from patients with respiratory diseases (Jacome et al., 2019).

Recently, Jacome et al., (2019) proposed a convolutional neural network with a spectrogram-based method for breathing phase detection in lung sounds to address some of the limitations mentioned above. This method was tested on lung sounds recorded from subjects both with and without respiratory disease. The lung sounds were recorded at six posterior chest locations. This method showed average sensitivity of 97 % and an average specificity of 84 % in identifying the breathing phases. However, the three subsets (two subsets for training the algorithm and one subset for evaluation) used in this study came from the same dataset- the Tromso 7 lung sound dataset. A common problem with machine learning methods is that they often work well with a generated dataset with samples recorded in an identical manner for both training and testing sets but may not perform as well when applied to new unseen datasets. Training an algorithm with large datasets recorded in different conditions may overcome this problem, but it is always challenging to get a large amount of data, which can limit overall generalization of a machine learning algorithm.

The estimated breathing phases in a recorded lung sound can be used for calculating the number of breath cycles. As mentioned earlier, a breath cycle consists of two phases one inspiratory phase

and one expiratory phase. Therefore, a new automatic breath cycle detection algorithm based on estimating the location of the breathing phases is proposed. The algorithm is validated using the longitudinal dataset, cross-sectional dataset, and the small dataset recorded from healthy subjects. The description of the datasets is provided later in this Chapter. A detailed working process of the algorithm is provided in the next section.

9.3 An automatic algorithm for breath cycle detection based on the estimation of the breathing phases

The algorithm is based on several steps:

9.3.1 Estimation of second derivative

In the first step, the second derivative of the input signal is calculated using SG filter (Savitzky & Golay, 1964). The SG filter parameters are degree of fitting polynomial p_f = 4; number of

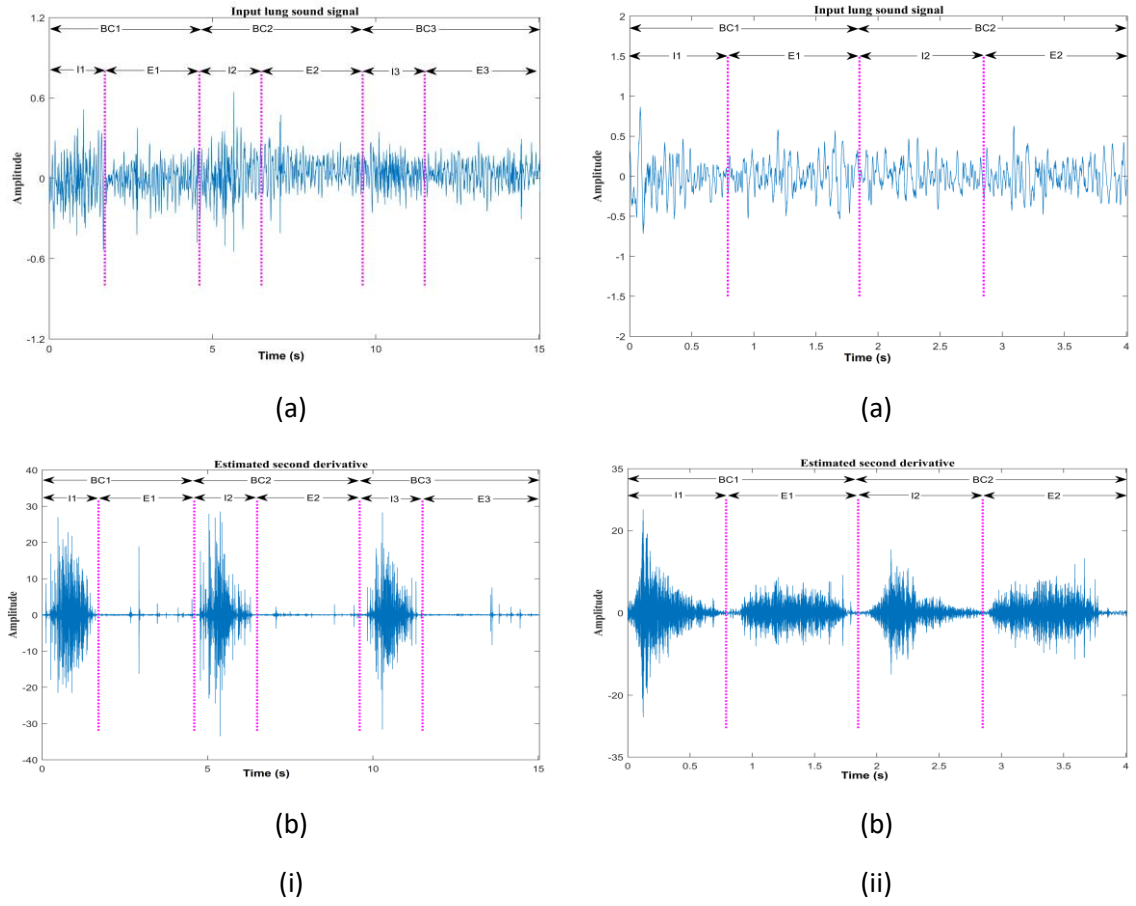


Figure 47 (i) :- (a) Input lung sound signal (b) Estimated second derivative of the input lung sound signal. (ii) :- (a) Input lung sound signal (b) Estimated second derivative of the input lung sound signal. I: Inspiratory phase; E: Expiratory phase; BC: Breath cycle.

coefficients $n_c = 89$; order of derivation d_o = 0, 1 and 2 for smoothing the lung sound signal, and estimating first and second derivative of the smoothed lung sound signal respectively. The SG filter is used calculating the second derivative because generates the second derivative from a smoothed version of the input signal. The reason for estimating the second derivative is that the coefficients corresponding to a higher intensity part of an input lung sound signal will remain in its second derivative with large amplitude but only the residue of coefficients related to quieter parts will remain, which may help to reveal the breathing phases. As an example: two cases are shown in Figure 47 (i) and Figure 47 (ii) in which inspiratory phases, expiratory phases, and breath cycles are audio visually marked by the author using open access Audacity software. Figure 47 (i-a) shows an input lung sound signal in which inspiratory phases are of high intensity and expiratory phases are quieter. From its second derivative as shown in Figure 47 (i-b) it can be seen that the coefficients related to high intensity inspiratory phases still have large amplitude. On the other hand, the coefficients corresponding to quieter, expiratory phases are smaller in amplitude. In another example, Figure 47 (ii-a) shows an input lung sound signal in which both inspiratory and expiratory phases are of high intensity. From its second derivative as shown in Figure 47 (ii-b), it is clear that when both inspiratory and expiratory phases are of high intensity the coefficients related to both phases have large amplitude in the second derivative signal.

The second derivative of an input lung sound signal estimated using SG filter may more clearly reveal inspiratory and expiratory phases but there is a still need for automatically identifying inspiratory and expiratory phases or automatic counting of the number of breath cycles from the estimated second derivative. Therefore, several further steps are used which are described below.

9.3.2 Estimation of absolute value of second derivative and normalised absolute value of second derivative

After estimating the second derivative of the input lung sound signal, its absolute value is calculated. Input lung sound signal, its second derivative, and absolute value of the second derivative are shown in Figure 48 (a), Figure 48 (b), and Figure 48 (c), respectively, with the inspiratory phases, expiratory phases, and breath cycles were marked by arrows. The absolute value of the second derivative may vary significantly, therefore, its value is normalized to an amplitude range of 0 to 1 using Eq. 72.

$$ASD_{norm}(n) = \frac{ASD(n) - ASD_{min}}{ASD_{max} - ASD_{min}} \quad (72)$$

where $ASD(n)$ is the absolute second derivative, n is the sample index, ASD_{min} is the minimum

value of the absolute second derivative, ASD_{max} is it's the maximum value and $ASD_{norm}(n)$ is the normalized absolute second derivative. The normalized absolute second derivative is shown in Figure 48 (d).

9.3.3 Clipping large amplitude peaks from the normalized absolute second derivative

The large amplitude peaks present (due to adventitious sounds such as crackles, movement artefact, heart sounds etc.) in the input lung sound signal may confuse the automated detection about the breathing phases. Therefore, once the normalized absolute second derivative is calculated, the large amplitude peaks are clipped using a threshold value estimated from the frequency histogram plot of the normalized absolute second derivative. The threshold is taken as the value having, on its left, 80 % of total area. The cut off of 80 % is selected based on Vannuccini et al., (1998). The frequency histogram plot of the normalized absolute second derivative is shown in Figure 48 (e). Figure 48 (f) shows the normalized absolute second derivative after clipping the large amplitude peaks.

9.3.4 Low pass filter

Next the clipped normalized absolute second derivative is passed through the 3rd order Butterworth low pass filter with cut off frequency 1.2 Hz or 1 Hz. Usually, a patient with cardiorespiratory disease breathes faster than a healthy subject. Therefore, the higher cut off frequency: 1.2 Hz, is selected for the longitudinal and cross sectional IPF datasets and the lower cut-off frequency: 1 Hz, is selected for healthy subjects. Here, the cut off frequencies 1.2 Hz and 1 Hz are empirically selected. Figure 48 (g) shows the output of the low pass filter.

9.3.5 Estimation of potential breathing phases and their onsets

In this step, peaks and valleys of the low pass filter output are used for calculating the potential breathing phases and breath onsets, respectively. Figure 48 (h) shows all possible peaks and valleys of the low pass filter output, where peaks are shown using black stars and valleys are displayed using green stars. This step is divided into several sub steps:

(i) Firstly, peaks of the low pass filter output which are 0.5 standard deviations above the mean of the low pass filter output are identified (see Figure 48 (i)).

(ii) Secondly, valleys of the low pass filter output which are less than the mean of the low pass filter output are identified, as shown in Figure 48 (j).

(iii) Now, from the valleys in sub step (ii) only valleys which are just before and after the

estimated peaks are considered. This may include the first sample in the signal (if there is no valley before the first estimated peak) and the last point of the signal (see Figure 48 (k)).

(iv) As mentioned by Chuah & Moussavi, (2000) typically the average duration of a breathing phase is approximately 1 s. Therefore, the distance between an estimated peak and the potential onset of its associated breathing phase is expected to be about 500 ms and the distance between any two potential onsets is expected to be of the order of 1s. However, breath cycle duration may change from breath to breath, hence any potential onsets (or valleys) which are closer than 200 ms to any estimated peaks are discarded.

(v) Additionally, if any potential onset is less than 500 ms from a previous onset it is excluded from the analysis. As an example, see Figure 48 (l), where the potential onset just after 2 sec in Figure 48 (k) is excluded. Note that once any potential onset is excluded then it is not used for comparing with other onsets.

(vi) Now, if more than one estimated peak exists between two potential onsets, only the maximum peak amongst them is considered, as shown in Figure 48 (m).

(vii) In the next step, average and standard deviation of the estimated peaks remaining after sub step (vi) are calculated.

(viii) Now, a peak threshold is estimated using the average and standard deviation calculated in the previous sub step:

$$Peak_{TH} = \frac{Peak_{avg}}{2} + Peak_{std} \tag{73}$$

where $Peak_{avg}$ and $Peak_{std}$ are the average and standard deviation of the remaining estimated peaks after sub step (vi), respectively.

(ix) Finally, estimated potential breath onsets are connected with each other using alternating positive and negative rectangles (starting with a positive rectangle) where positive rectangles show potential inspiratory phases and negative rectangles represent potential expiratory phases, as shown in Figure 48 (n).

9.3.6 **Estimating breathing phases**

Usually, in lung sounds recorded from posterior chest locations, the intensity of the inspiratory phases is expected to be high compared to the expiratory phases (Kompis et al., 1997; Moussavi

et al., 1998). Therefore, as a last step, potential inspiratory phases are verified using the calculated peak threshold (Eq. 73).

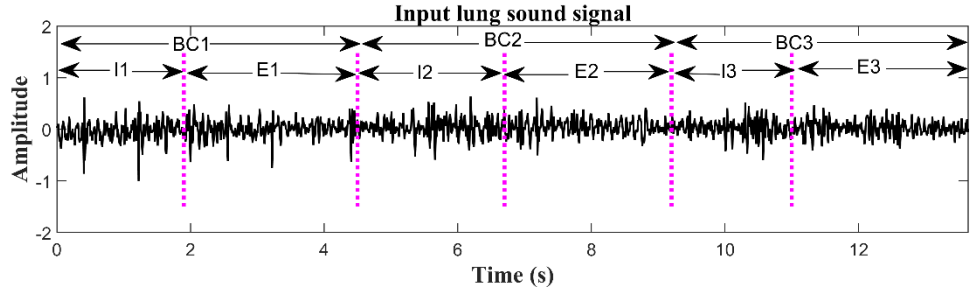
The process starts with the first potential inspiratory phase:

If the first potential inspiratory phase consists of an estimated peak of greater than the peak threshold, the potential inspiratory phase is considered as a true inspiratory phase and the next potential inspiratory phase is verified. During the verification process if any potential inspiratory phase (including the first potential inspiratory phase) without or with an estimated peak of amplitude less than or equal to the peak threshold occurs, the first potential onset point of that potential inspiratory phase is eliminated from the process. Now, the remaining potential onsets are reconnected with each other using alternating positive and negative rectangles (starting with positive rectangle) where any positive rectangles before the eliminated potential onset point show the inspiratory phases (because they are already verified) and after the eliminated onset point indicate potential inspiratory phases which still need to verify. Moreover, any negative rectangle between the two verified inspiratory phases represents an expiratory phase and any negative rectangle before or after an unverified potential inspiratory phase shows a potential expiratory phase.

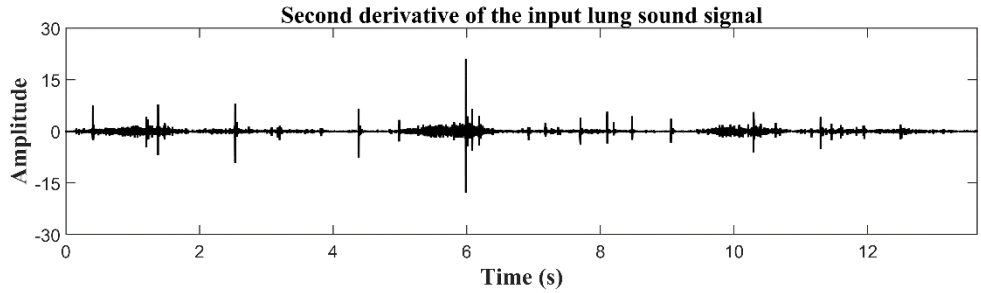
The verification process again starts with the new assigned potential inspiratory phases starting from the new first potential inspiratory phase. This verification process ends when all positive rectangles consist of an estimated peak with amplitude greater than the peak threshold. Hence, at the end of the process all positive rectangles with estimated peaks greater than the peak threshold are considered as inspiratory phases and all negative rectangles with or without estimated peaks are considered as expiratory phases.

Based on the estimated breathing phases the number of breath cycles is calculated. The combination of two phases one estimated inspiratory phase and one estimated expiratory phase (starting with an inspiratory phase because the process of breathing is considered to start with an inspiration phase) makes up one breath cycle and the total of all such combinations represents the number of breath cycles.

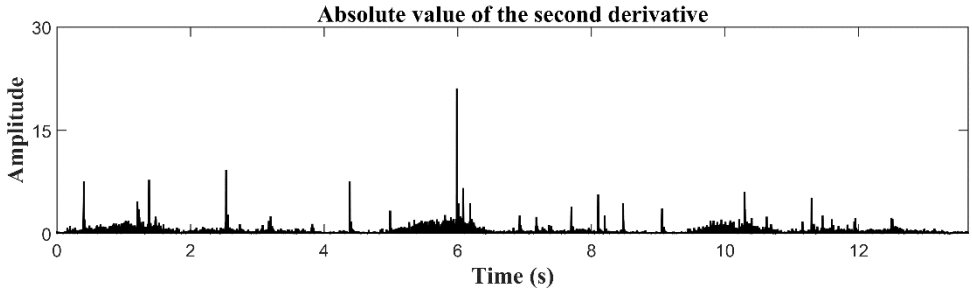
After applying the verification process the estimated inspiratory and expiratory phases are shown in Figure 48 (o) using positive and negative rectangles, respectively. The breath cycles marked from audio-visual analysis of the signals are also indicated in Figure 48 (o). Moreover, the estimated inspiratory and expiratory phases with calculated breath cycles are displayed in the second derivative of the input lung sound signal in Figure 48 (p), where positive rectangles show the estimated inspiratory phases, negative rectangles indicate the estimated expiratory phases,


Input lung sound signal
BC1
BC2
BC3
I1
E1
I2
E2
I3
E3
Amplitude
Time (s)

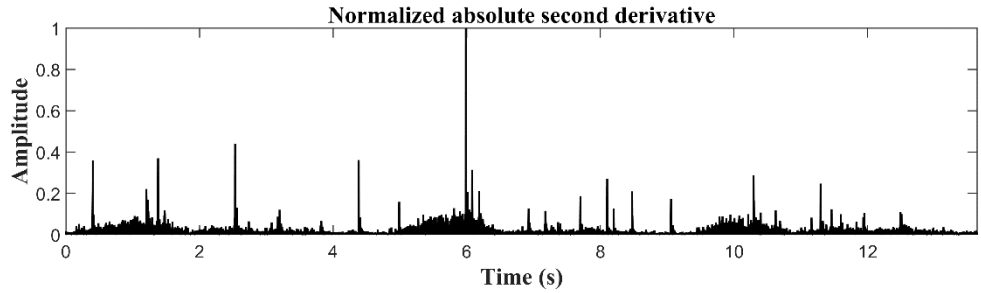
(a)


Second derivative of the input lung sound signal
Amplitude
Time (s)

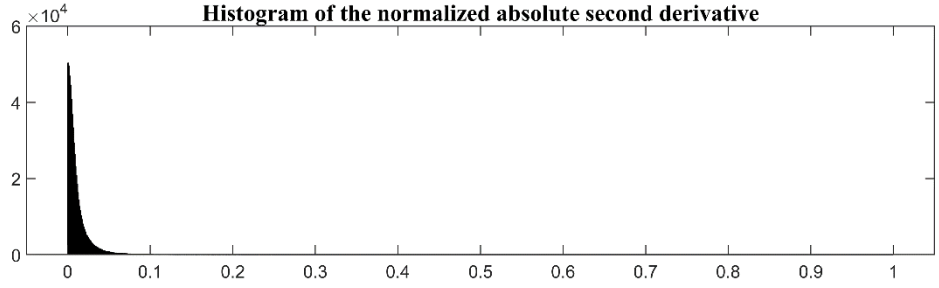
(b)


Absolute value of the second derivative
Amplitude
Time (s)

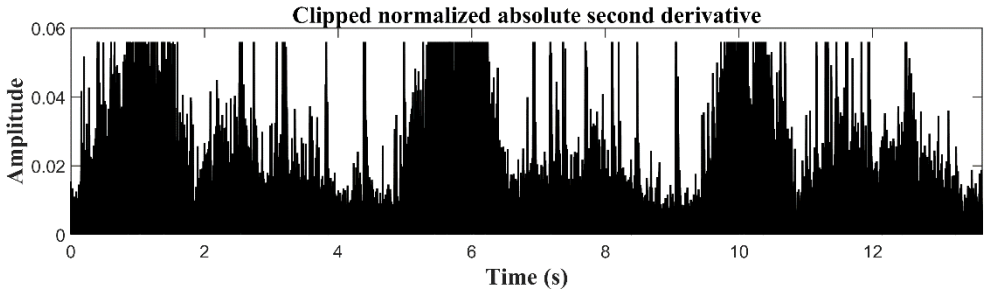
(c)


Normalized absolute second derivative
Amplitude
Time (s)

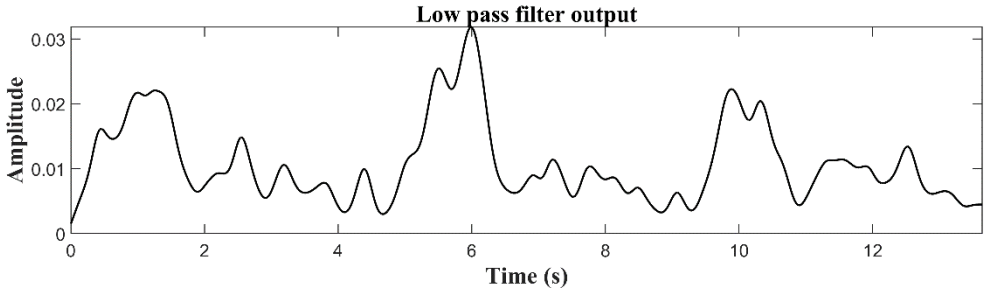
(d)


Histogram of the normalized absolute second derivative
×10^4
6
4
2
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

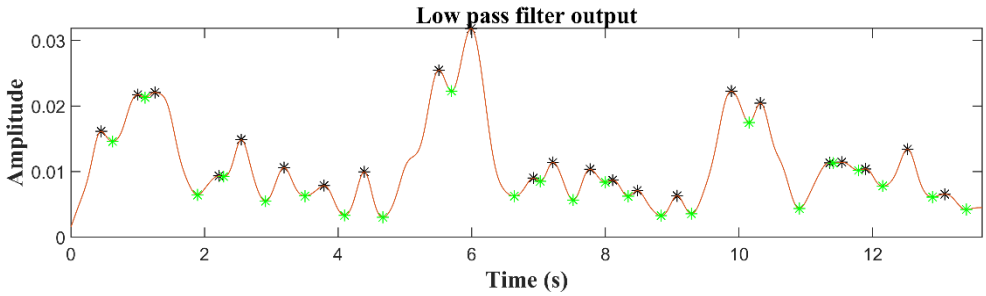
(e)


Clipped normalized absolute second derivative
Amplitude
0.06
0.04
0.02
0
0
2
4
6
8
10
12
Time (s)

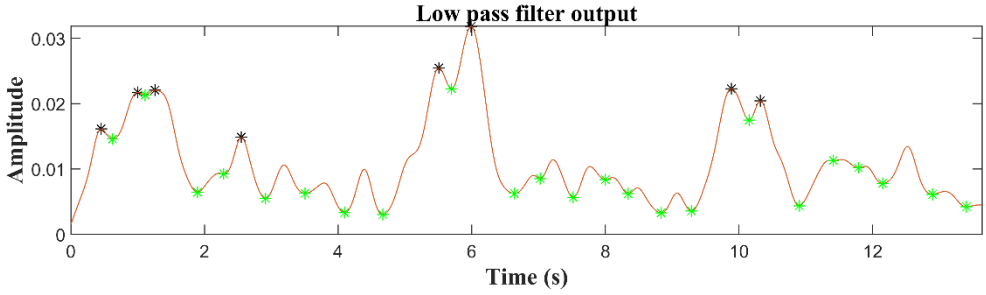
(f)


Low pass filter output
Amplitude
0.03
0.02
0.01
0
0
2
4
6
8
10
12
Time (s)

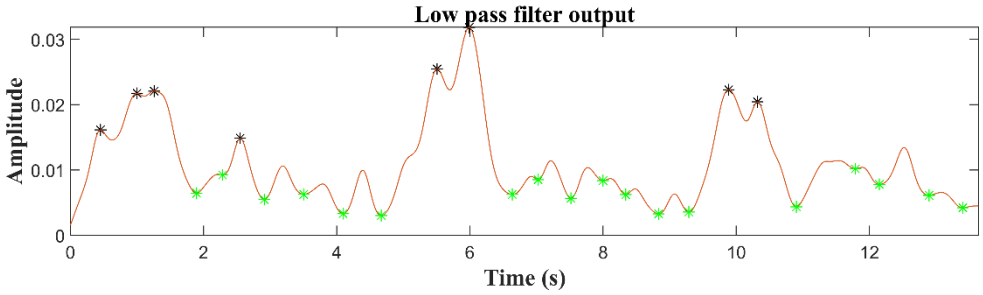
(g)


Low pass filter output
Amplitude
0.03
0.02
0.01
0
0
2
4
6
8
10
12
Time (s)

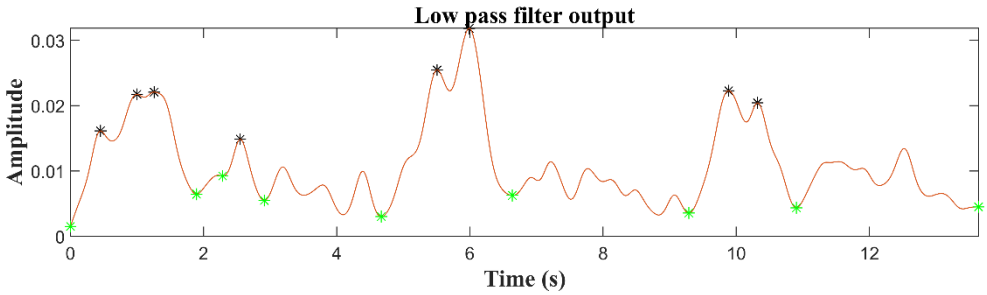
(h)


Low pass filter output
Amplitude
Time (s)
0
0.01
0.02
0.03
0
2
4
6
8
10
12

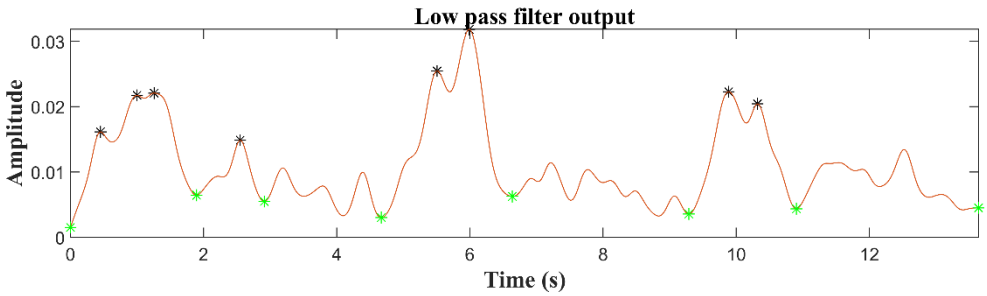
(i)


Low pass filter output
Amplitude
Time (s)
0
0.01
0.02
0.03
0
2
4
6
8
10
12

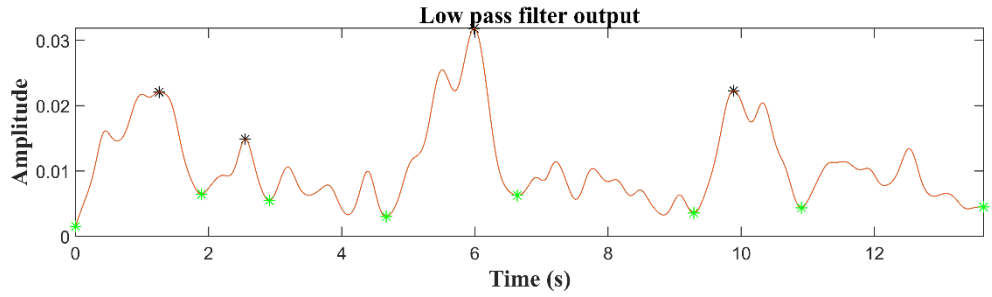
(j)


Low pass filter output
Amplitude
Time (s)
0
0.01
0.02
0.03
0
2
4
6
8
10
12

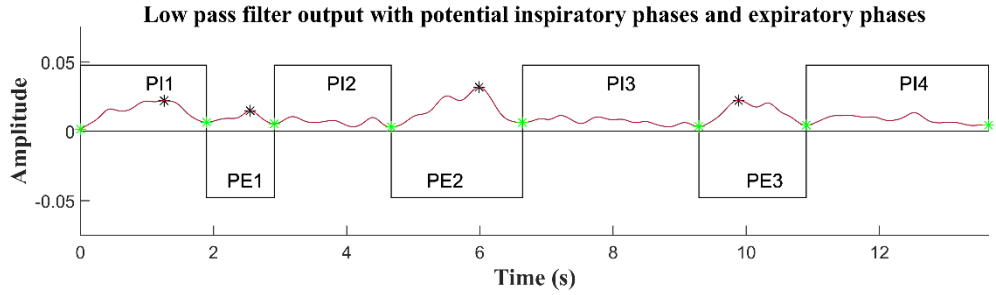
(k)


Low pass filter output
Amplitude
Time (s)
0
0.01
0.02
0.03
0
2
4
6
8
10
12

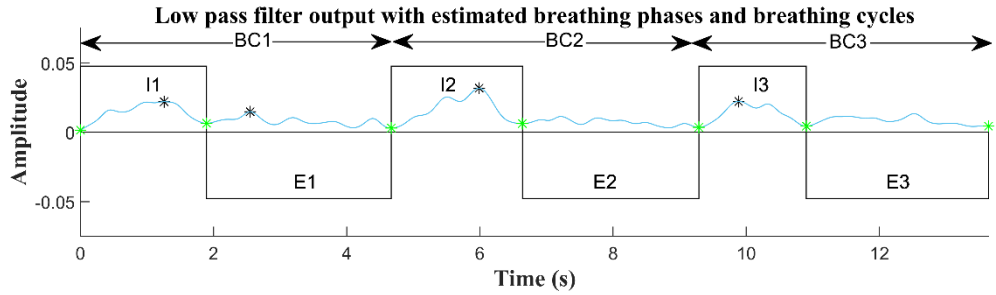
(l)



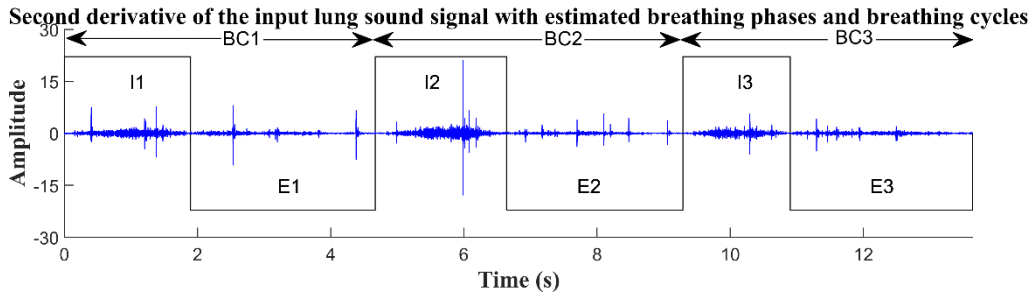
(m)



(n)



(o)



(p)

Figure 48 (a) Input lung sound signal; (b) Second derivative of the input lung sound signal; (c) Absolute value of the second derivative; (d) Normalized absolute second derivative; (e) Frequency histogram of the normalized absolute second derivative; (f) Clipped normalized absolute second derivative; (g) Low pass filter output; (h-m) Low pass filter output with different conditions of the section

9.3.5; (n) Low pass filter output with potential inspiratory phases and expiratory phases; (o) Low pass filter output with estimated breathing phases and breathing cycles; (p) Second derivative of the input lung sound signal with estimated breathing phases and breath cycles. I: Inspiratory phase; E: Expiratory phase; BC: Breath cycle.

and a combination of two phases one inspiratory and one expiratory represents a breath cycle. When we compare the automatic detection results (Figure 48 (p)) with the audio-visual marking (Figure 48 (a)) it is clear that all the true inspiratory phases, expiratory phases and breath cycles in this example have been detected.

9.4 Datasets and audio-visual marking of breathing phases and breath cycles

9.4.1 **Dataset**

The algorithm is validated using three different lung sound datasets: (a) Longitudinal dataset recorded from 19 IPF patients, (b) Cross-sectional dataset recorded from 55 patients who were referred for HRCT scan of the chest for various clinical indications, and (c) a dataset recorded from 10 healthy subjects. The complete description of the longitudinal and cross-sectional datasets is provided in Chapter 8. The healthy subjects' lung sounds dataset was recorded from 10 healthy subjects and, as for the longitudinal dataset, in this dataset each participant attended a total of 7 visits (every visit approximately in 2 months) over a 1 year period (Sgalla, 2017). In all three datasets the lung sounds recorded from 6 posterior locations using a digital stethoscope are analysed. See Chapter 8, Figure 37 for the 6 posterior locations (in green: L1-L6) in the case of longitudinal and healthy subjects datasets, and Figure 44 for the 6 posterior locations (L1-L6) in the case of cross-sectional dataset.

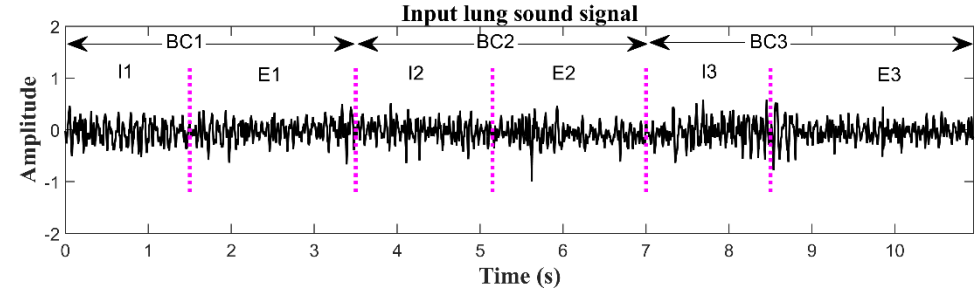
9.4.2 **Audio-visual marking of breathing cycles, inspiratory phases and expiratory phases**

In all three datasets each lung sound file was audio-visually marked for the number of breathing cycles, number of inspiratory phases, and number of expiratory phases by the author. In each lung sound file only the full breathing cycles (one inspiratory phase followed by one expiratory phase) were considered for the analysis. In the longitudinal dataset out of 19 IPF patients 13 patients completed the total 7 visits with lung sound recordings at all 6 posterior locations. 2 patients withdrew from the study due to poor health and 3 patients died during the observation period (Sgalla, 2017). One patient completed the total 7 visits but missed one lung sound recording at

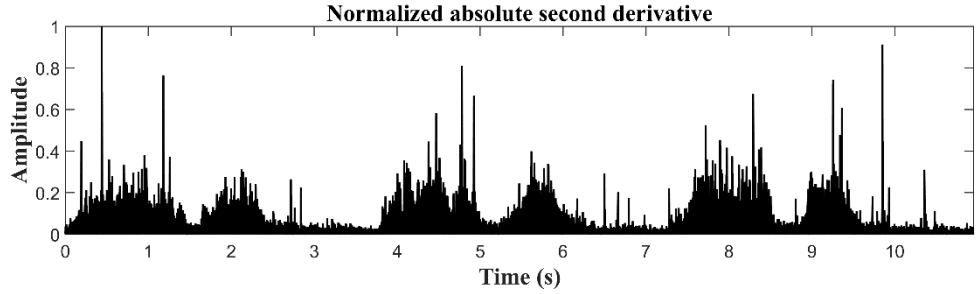
location L1 (see Chapter 8, Figure 37) in one of the 7 visits. Therefore, in the longitudinal dataset, in total 689 lung sound files (546 lung sounds recorded from 13 patient's×6 recording sites×7 visits, 41 lung sounds recorded from: 1 patient×6 recording sites×6 visits and 1 patient×5 recording sites×1 visit, 60 lung sounds recorded from 2 patient's×6 recording sites×5 visits, 24 lung sounds recorded from 2 patient's×6 recording sites×2 visits, 18 lung sounds recorded from 1 patient×6 recording sites×3 visits) were analysed. In the case of healthy subjects dataset, out of 10 subjects 7 subjects completed the study. 3 subjects withdrew from the study at some point due to personal or otherwise non-specified reasons (Sgalla, 2017). From 7 subjects, 6 completed the total 7 visits and one patient skipped a visit during the observation period. Therefore, in the healthy subjects dataset, out of 336 lung sound files (252 lung sounds recorded from 6 patient's×6 recording sites×7 visits, 36 lung sounds recorded from 1 patient×6 recording sites×6 visits, 30 lung sounds recorded from 1 patient×6 recording sites×5 visits, 12 lung sounds recorded from 1 patient×6 recording sites×2 visits, 6 lung sounds recorded from 1 patient×6 recording sites×1 visits), 282 lung sound files were analysed. 54 lung sound files were excluded from the study due to inaudibility of breathing phases or breath cycles. In the cross-sectional dataset out of 55 patients, lung sounds recorded from 48 patients were used for the analysis. 5 patients were excluded from the analysis due to inaudibility of breathing phases or breath cycles in their recorded lung sounds. In the excluded lung sound files, due to very quiet breathing the author was not able to mark either breathing phases or breath cycles. Therefore, out of 262 available lung sound files recorded from 48 patients, 258 lung sound files were analysed. With 4 lung sound files excluded from the study due to inaudibility of breathing phases or breath cycles. The different lung sound datasets used for evaluating the performance of the developed algorithm are shown in Table 30.

Table 30 Different lung sound datasets are used for evaluating the algorithm.

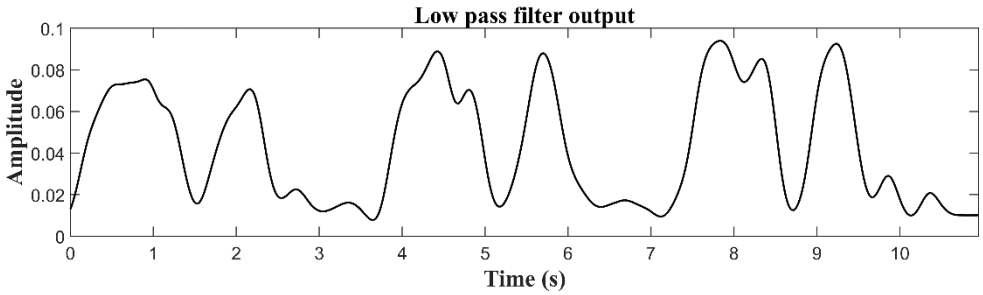
Datasets	Number of lung sound files	Number of inspiratory phases	Number of expiratory phases	Number of breath cycles
Longitudinal dataset	689	2416	2416	2416
Cross-sectional dataset	258	741	741	741
Healthy subjects dataset	282	905	905	905



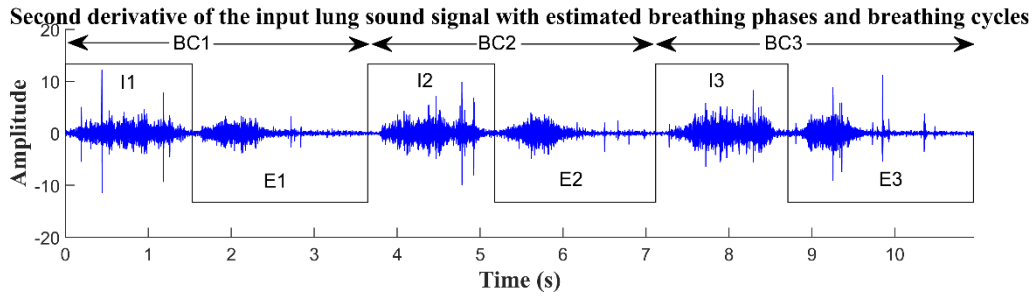
(a)



(b)

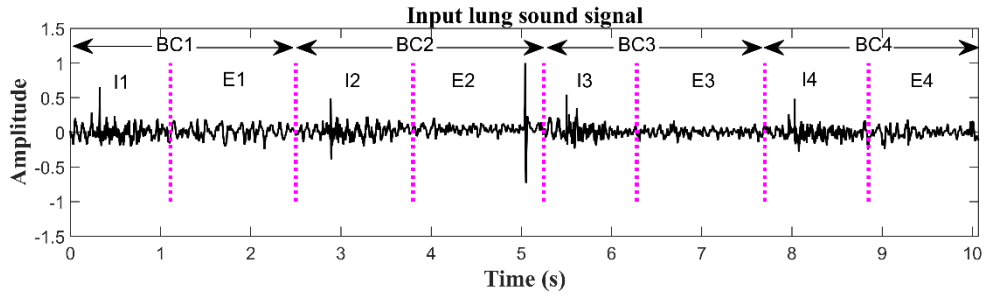


(c)

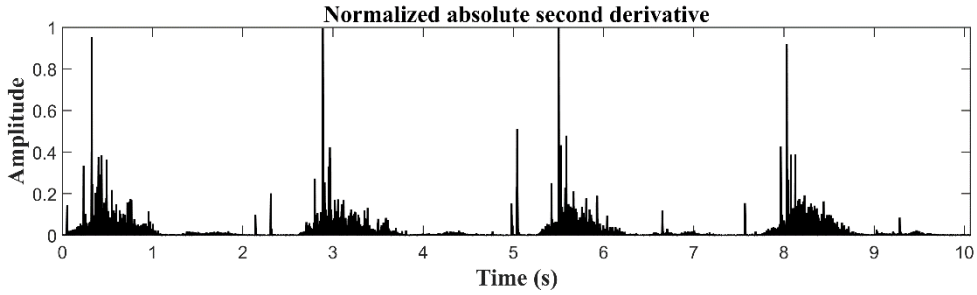


(d)

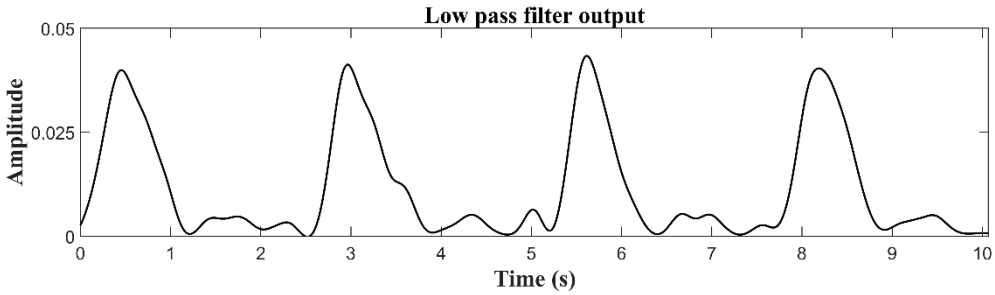
Figure 49 (a) Input lung sound signal selected from the longitudinal dataset; (b) Normalized absolute second derivative; (c) Low pass filter output; (d) Second derivative of the input lung sound signal with estimated breathing phases and breath cycles. I: Inspiratory phase; E: Expiratory phase; BC: Breath cycle.



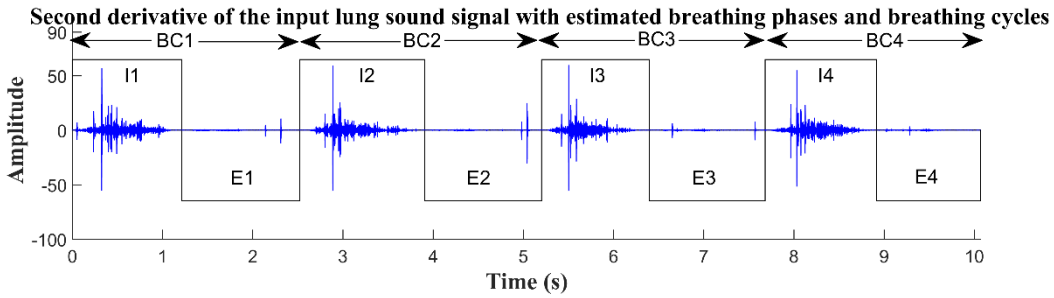
(a)



(b)

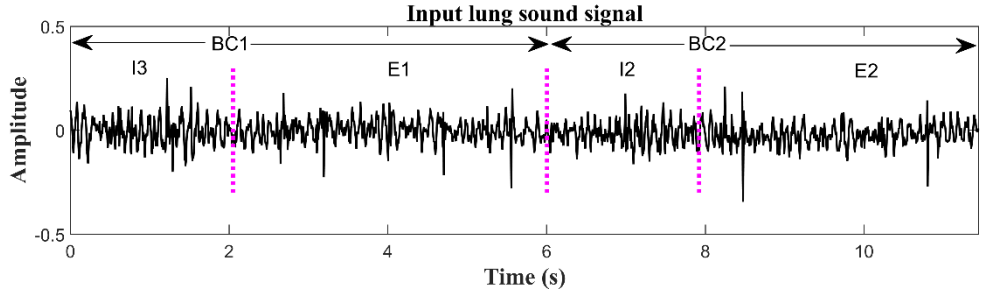


(c)

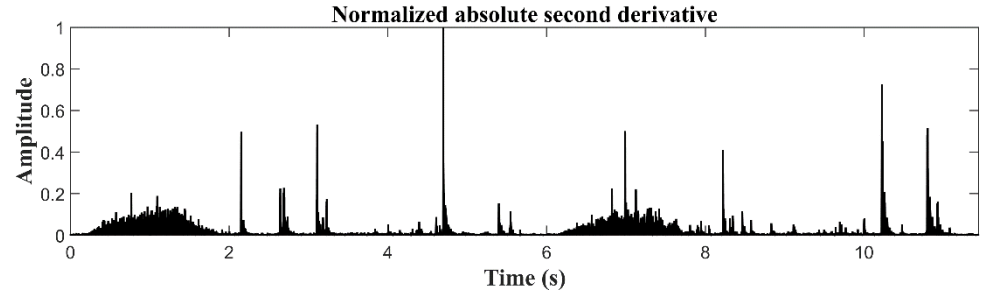


(d)

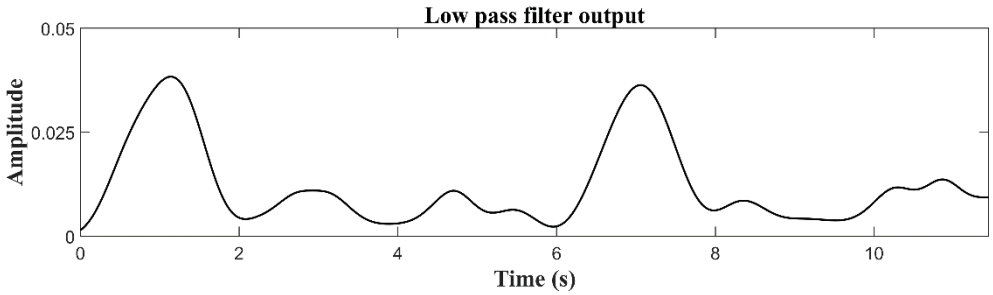
Figure 50 (a) Input lung sound signal selected from the cross-sectional dataset; (b) Normalized absolute second derivative; (c) Low pass filter output; (d) Second derivative of the input lung sound signal with estimated breathing phases and breath cycles. I: Inspiratory phase; E: Expiratory phase; BC: Breath cycle.



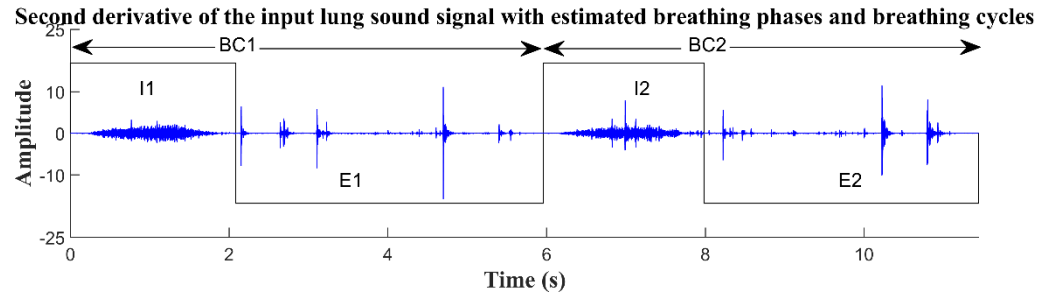
(a)



(b)



(c)



(d)

Figure 51 (a) Input lung sound signal selected from the healthy subjects dataset; (b) Normalized absolute second derivative; (c) Low pass filter output; (d) Second derivative of the input lung sound signal with estimated breathing phases and breath cycles. I: Inspiratory phase; E: Expiratory phase; BC: Breath cycle.

9.5 Performance evaluators

Three parameters are used for evaluating the breath cycle detection performance of the algorithm: sensitivity (SE), positive predictive value (PPV), and F-score (F_1).

$$SE = \frac{TP}{TP + FN} \tag{74}$$

$$PPV = \frac{TP}{TP + FP} \tag{75}$$

$$F_1 = 2 \times \frac{SE \times PPV}{SE + PPV} \tag{76}$$

Where TP, FN and FP are true positive (Number of breath cycles audio-visually counted), false negative (number of counted breath cycles which are not detected) and false positive (number of detected breath cycles which were not counted), respectively. The ability of the algorithm to identify the inspiratory and expiratory phases is also evaluated using the above mentioned parameters (Equations 74-76), where true positive is number of inspiratory or expiratory phases audio-visually counted, false negative is number of inspiratory or expiratory phases are not detected, and false positive is number of incorrectly detected inspiratory or expiratory phases. The obtained experimental results are presented in the next section.

9.6 Experimental Results

To show the ability of the algorithm to automatically estimate the breathing phases one example from each of the three datasets is shown in Figure 49, Figure 50, and Figure 51, respectively. Figure 49 Indicates one example from longitudinal dataset, Figure 50 shows one example from cross-sectional dataset, and Figure 51 illustrates one example from healthy subjects dataset. Plots labelled (a) show curves for input lung sound signals where breathing cycles and breathing phases are audio-visually marked by the author, plots labelled (b) show curves for the normalized absolute second derivative signals, plots labelled (c) show curves for the low pass filter outputs, plots labelled (d) show curves for the second derivative of the input lung sound signals with estimated breathing cycles and breathing phases where positive rectangles show the estimated inspiratory phases, negative rectangles display the estimated expiratory phases, and the combination of one inspiratory phase and one expiratory phase represents a breath cycle. Note that the complete process of the algorithm is discussed in section 9.3. By comparing the estimated breathing cycles or breathing phases using the algorithm (plots labelled (d)) with the audio-visually marked on the input lung sound signal (plots labelled (a)) it can be notice that all the breathing cycles and breathing phases are correctly detected by the algorithm.

Table 31 Sensitivity, Positive predictive value and F-score for number of breath cycles.

Dataset	SE	PPV	F-score
Longitudinal dataset	94.41	93.25	93.83
Cross-sectional dataset	97.37	94.76	96.05
Healthy subjects dataset	88.81	93.01	90.86
Average	93.53	93.67	93.58

Table 32 Sensitivity, Positive predictive value and F-score for inspiratory phases.

Dataset	SE	PPV	F-score
Longitudinal dataset	91.14	88.95	90.03
Cross-sectional dataset	94.52	89.28	91.83
Healthy subjects dataset	86.35	88.04	87.19
Average	90.67	88.76	89.68

Table 33 Sensitivity, Positive predictive value and F-score for expiratory phases.

Dataset	SE	PPV	F-score
Longitudinal dataset	90.96	89.68	90.32
Cross-sectional dataset	93.56	90.92	92.22
Healthy subjects dataset	85.62	89.52	87.53
Average	90.05	90.04	90.02

In all three datasets the values of SE, PPV, and F-score of the algorithm in estimating the number of breath cycles are presented in Table 31. On average, the algorithm achieved a SE of 93.53 %, a PPV of 93.67 %, and F-score of 93.58 %.

As mentioned earlier, the breathing cycles are calculated based on the estimation of the inspiratory and expiratory phases. Hence, in all three datasets the values of SE, PPV, and F-score of the algorithm in identifying inspiratory phases are presented in Table 32 and identifying expiratory phases are shown in Table 33. In terms of the estimation of the inspiratory phases the algorithm obtained on average, SE of 90.67 %, PPV of 88.76 %, and F-score of 89.68 %. On the

other hand, in the case of estimation of the expiratory phases, the algorithm achieved SE of 90.05 %, PPV of 90.04 %, and F-score of 90.02 %.

9.7 Discussion

The algorithm is not only tested on lung sounds recorded from healthy subjects but also evaluated using the lung sounds recorded from cardiorespiratory patients such as IPF patients in order to show its potential to calculate the number of breath cycles in different breathing patterns. The individual breath phase estimation performance of the proposed algorithm was on average, SE of 90 %, PPV of 88 %, and F-score of 89 % in case of inspiratory phases and on average, SE of 90 %, PPV of 90 %, and F-score of 90 % in case of expiratory phases. However, the algorithm has shown slightly low, on averaged sensitivity for breath phase estimation compared to previously developed algorithms (Huq & Moussavi, (2012); Jacome et al., (2019)). Nevertheless, it is important to notice that the primary aim of the algorithm was to calculate the number of breath cycles. In terms of calculating the number of breath cycles the algorithm achieved on average SE of 93 %, PPV of 93 %, and F-score of 93 %. The main advantage of the proposed algorithm is its low computational cost which makes it an ideal candidate for the clinical setting where fast processing may be helpful in decision-making. To our knowledge, this is the first study to calculate the number of breath cycles based on the estimation of breathing phases using the second derivative of the input lung sound signal.

The method showed good results in terms of calculating the breath cycles, but the limitation of the algorithm is its dependency on the non-adaptive cut-off frequency of the low pass filter. The higher cut-off frequency may generate too many false peaks and valleys which may increase the chances of identifying false breathing phases. On the other hand, the lower cut-off frequency may smooth out the true peaks and valleys of the breathing phases (especially in the case of fast breathing) which may increase the chances of losing some of the true breathing phases. Therefore, it is important to notice that in our analysis we used two cut-off frequencies: 1.2 Hz and 1 Hz. As mentioned earlier, a patient with cardiorespiratory disease generally breathe faster than a healthy subject therefore the 1.2 Hz cut off frequency was used for the longitudinal and cross-sectional datasets and 1 Hz cut-off frequency was used for the healthy subjects dataset. The future research to automatically select the low pass filter cut-off frequency according to the breathing pattern may enhance the breath cycle detection potential of the proposed algorithm.

Furthermore, the algorithm was evaluated against the audio-visual assessment made by the author. As mentioned by Pinho et al., (2015), human assessment is associated with high levels of subjectivity. Therefore, future research should consider comparing the performance of the

algorithm against a multi-annotators gold standard. Moreover, future research will focus on evaluating the performance of the algorithm with lung sounds recorded from different populations: children, young adults, old subjects with or without pulmonary diseases.

9.8 Summary

Chapter 9 described a new automatic breath cycle detection algorithm based on the estimation of breathing phases. From the experimental results it can be notice that the proposed algorithm can estimate breathing phases from the recorded lung sounds and estimated breathing phases may be used for calculating the number of breath cycles. However, the individual breathing phase (inspiratory phase or expiratory phase) identification performance of the algorithm is slightly less good when compared to the estimation of the number of breath cycles. Nevertheless, the primary aim of the algorithm was to estimate the number of breath cycles. The estimated number of breath cycles can be used in calculating the NOC/BC as we have seen in two case studies in Chapter 8. Furthermore, the combination of the IEM-FD filter, crackle verification and counting criteria (see section 8.3.4), and a new breathing cycle detection method can provide a fully automatic way of estimating the NOC/BC. The next chapter will discuss the findings of this research including the limitations of this study, conclusions and future possibilities.

Chapter 10 Discussion, conclusions, and future work

10.1 Introduction

The overall aim of this research was to develop a new crackle separation technique, the IEM-FD filter, which can be used in clinical setting for recorded lung sounds analysis. The crackle separation potential of the IEM-FD filter was evaluated using a developed dataset and compared with three previously published crackle separation techniques. The IEM-FD filter was then applied to two datasets recorded in clinical settings: a longitudinal dataset recorded in patients with IPF and a cross-sectional dataset recorded from patents referred for an HRCT of the lungs. Finally, a new automatic breath cycle detection method based on the estimation of breathing phases was developed.

This chapter will start by discussing a dataset generated for systematic testing of crackle separation techniques. This will be followed by discussing the separation performance of the new IEM-FD filter. The chapter will then discuss the two case studies (longitudinal study and cross sectional study). An algorithm developed for automatic detection of number of breath cycles from the recorded lung sounds is also discussed. The limitations of the study are then discussed. This chapter finishes by presenting the conclusions of the research its main findings and some future possibilities.

10.2 Generated dataset

To provide an open platform to researchers so that they can evaluate and compare their crackle separation and detection methods, an open access dataset is presented in this study. This dataset contains real and simulated fine and coarse crackles with different IDW/2CD and real lung sounds with fine and coarse crackles recorded from an IPF patient and a BE patient, respectively, Addition of more real lung sounds with fine or coarse crackles recorded from different cardiopulmonary diseases will provide more test cases for evaluating crackle separation performance of new algorithms in future. The detailed description of the generated dataset is provided in Chapter 3 and a paper published on the development of the dataset can be found in Appendix A (A6).

10.3 Proposed IEM-FD filter for separating crackles from normal breath sounds

As mentioned in section 2.2.3, in recent years many automatic crackle separation techniques have

been proposed for separating crackles from the normal breath sounds. However, the existing crackle separation methods either suffer from high computation complexity and/or fail to preserve crackle morphology after separation and/or have low objectivity etc. (see section 2.2.4.). Therefore, in this study, a new IEM-FD filter was proposed for automatically separating crackles from normal breath sounds. See Chapter 5 for the detailed working process of the IEM-FD filter.

As mentioned in section 1.2, the crackle time features such as number of crackles in inspiratory phase, number of zero line crossings, first half period of the crackles etc. can be used for differentiating IPF patients from patients with pneumonia and congestive heart failure (Flietstra et al., 2011). However, the presence of normal breath sounds may introduce errors when calculating crackle time features. For example, Munakata et al., (1991), identified that baseline drift over the duration of a crackle (a consequence of over-estimation) may introduce errors when calculating IDW and 2CD leading to incorrect classification of crackle type (fine or coarse) and increasing the potential for misdiagnosis. Therefore, for any automatic crackle separation technique it is not only important to separate crackles from normal breath sounds with high rate of detectability but it is also important to separate crackles with high separation quality (less over or underestimation). The separation of crackles from normal breath sounds with high quality can lead to extraction of accurate crackle time domain features, which can help to differentiate between cardiopulmonary diseases with high sensitivity and specificity and reduce the chances of medication error or mistreatment.

The performance of the IEM-FD filter with low computational load, high rate of detectability, low over or under estimation and ability to preserve crackle morphology after separation shows its potential for automatic crackle separation. Separation of crackles from normal breath sounds is an initial processing stage towards better estimation of number of crackles and their time domain features. In comparison with the established WTST-NST filter (Hadjileontiadis & Panas, 1997), WT-FD filter (Hadjileontiadis, 2005 (I); Hadjileontiadis, 2005 (II)) and EMD-FD filter (Hadjileontiadis, 2007) the IEM-FD filter has an equally high Rate of Detectability and total performance for both fine and coarse crackles (see Chapter 7). Further the IEM-FD filter has fewer data-dependent optimization parameters compared to the WTST-NST filter, WT-FD filter, and the EMD-FD filter making it generally applicable to signals recorded from cardiopulmonary patients with different diagnoses without the need for data dependent customization to optimize its performance. The separation performance of the IEM-FD filter and its systematic comparison with the selected methods is presented in Chapter 7.

The IEM-FD filter showed good results in terms of separating crackles from normal breath sounds. However, the IEM-FD filter, has several limitations: First, the selection of the SG filter parameters

in the IEM method is not adaptive which may affect its performance when, due to high frequency background noise, the envelope mean value is not properly estimated; Second, the dependency of the IEM-FD filter stopping criteria on the three non-adaptive accuracy levels i.e. β_1 for the IEM method, β_2 for the IEM-FD filter and β_3 for the FDPP algorithm may limit its overall performance.

The automatic separation of crackles from lung sounds using the IEM-FD filter not only provides an objective way of analyzing recorded lung sounds but also shows the future possibilities of computer based lung sounds for pulmonary disease diagnosis or monitoring. A paper published on the IEM-FD filter can be found in Appendix A (A7).

10.4 Two case studies

The IEM-FD filter was then applied to two case studies: longitudinal dataset and cross-sectional dataset. In both studies, NOC/BC was calculated with the separation by the IEM-FD filter the fundamental first step in processing the signals. The process of estimating NOC/BC was discussed in Chapter 8.

The longitudinal study showed that reproducible acoustic features generated from original signal (11 out of 13) significantly correlate with the NOC/BC, but only 2 reproducible acoustic features: zero-cross of the original signal (sig_zerocross) (r=0.628, p<0.001) and zero-cross of the original signal in the frequency range of 200-500 Hz (sig_200_500 Hz_zerocross) (r=0.590, p<0.001) indicated moderate correlation at the 0.01 significance level with the NOC/BC. Hence, these may be the most relevant features for monitoring IPF patients. However, it is important to notice that this study was tested on a small population of longitudinal data collected from IPF patients, larger IPF longitudinal datasets are required to further investigate the potential for monitoring disease severity in IPF patients. The more detailed discussion on the longitudinal study is provided in section 8.4.

In the cross-sectional study the developed automatic system showed the potential to differentiate IPF patients from non-IPF patients based on the average NOC/BC, it should be noted that, the automatic system was tested on a small population and a larger dataset will be needed for further investigation of its potential for differentiating IPF patients from patients with other pathologies. The more detailed discussion of the cross-sectional study is provided in section 8.6.

10.5 An algorithm for automatic detection of number of breath cycles based on the estimation of breathing phases

Although, in terms of estimating the number of breath cycles the algorithm achieved, on average SE of 93 %, PPV of 93 %, and F-score of 93 % (see Chapter 9). There are several points which should be noted: (1) the selected SG filter parameters used for estimating the second derivative of the input signal worked well in our datasets in terms of revealing breathing phases (except in some cases where breathing is very quiet). However, the selection of these parameters is not adaptive in our study hence it is possible that these selected parameters may not work as well as they performed in our datasets when applied to a new dataset with different breathing pattern. Therefore, the future research should explore the performance of these parameters on a larger dataset recorded from different age populations to see their ability to reveal breathing phases in different breathing pattern; especially when breathing is very quiet; (2) the non-adaptive selection of the low pass filter frequency may affect the overall performance of the algorithm therefore an adaptive approach is needed for selecting low pass filter cut-off frequency according to breathing pattern in future work and; (3) there are several methods available in the literature for automatically estimating breathing phases (Chuah & Moussavi, (2000); Huq & Moussavi, (2012); Reyes et al., 2016; Jacome et al., 2019), which may be used for calculating number of breath cycles in the input lung sound signal. These methods were not explored in our study but might be compared with the developed method in future work. A more detailed discussion of the developed method is provided in section 9.7.

10.6 Limitations

This study has several limitations:

Firstly, the IEM-FD filter was tested using a dataset (Chapter 3, Table 2) largely consisting of simulated signals, but then applied to the real signals, of the longitudinal and cross sectional datasets (see Chapter 8), It is important to notice that the test dataset only contains two real lung sounds samples (RBFC and RBCC) hence in this study the IEM-FD filter was only tested on a very limited number of real lung sounds samples before applying to the real signals. Therefore, future research needs to validate the performance of the IEM-FD filter on larger set of real lung sounds signals recorded at different locations (posterior, anterior, and lateral) from different cardiopulmonary diseases which can further evaluate the crackle separation performance of the IEM-FD filter and provide better support to the IEM-FD filter for using in real lung sound recordings.

Secondly, both the: IEM-FD filter and the automatic breath cycle detection method depend on empirical setting of parameters: accuracy levels (β_1, β_2, and β_3) in the IEM-FD filter and setting a low pass filter cut-off frequency (1.2 Hz or 1 Hz) in the breath cycle detection method. Therefore, future research should focus on making these parameters data driven which will increase the objectivity of these methods.

10.7 Conclusions and future work

This study proposed an automatic method for separating pulmonary crackles from normal breath sounds i.e. IEM-FD filter. The performance of the IEM-FD filter was evaluated using the developed dataset for systematic testing of crackle separation techniques and compared with the three previously published methods. Key findings of this research were: (1) The IEM-FD filter can achieve high accuracy for the number of crackles identified with 99.98 % of fine crackles and 99.80 % of coarse crackles identified in our test samples; (2) The IEM-FD filter has low computational cost/separation time compared to the established WTST-NST filter, WT-FD filter, and EMD-FD filter; (3) The IEM-FD filter can provide crackle separation with less over-estimation compared to the WTST-NST filter, WT-FD filter, and EMD-FD filter and (4) The IEM-FD filter can better preserve crackle morphology after separation compared to the WTST-NST filter, WT-FD filter, and EMD-FD filter in both fine and coarse crackle test signals. We concluded that the IEM-FD filter would be suitable for use in a clinical context for estimating number of crackles or as a first step in classifying crackles (fine or coarse) on the basis of their time domain features. Future research will focus on developing filter parameters that are fully adaptive and on evaluating the operation of the IEM-FD on a more diverse dataset recorded from cardiopulmonary patients, which can further test its ability to identify crackles in different pulmonary conditions.

To test the potential of the IEM-FD filter on real data it was applied to two case studies: longitudinal study and cross sectional study. The main finding of the longitudinal study was that reproducible acoustic features calculated from original lung sound signals correlate with the NOC/BC. This shows these reproducible can be used in clinics for assessing disease severity in IPF patients. The future research should consider larger longitudinal IPF datasets which can further investigate the potential of the reproducible acoustic features generated from original signal to monitor disease severity in IPF patients.

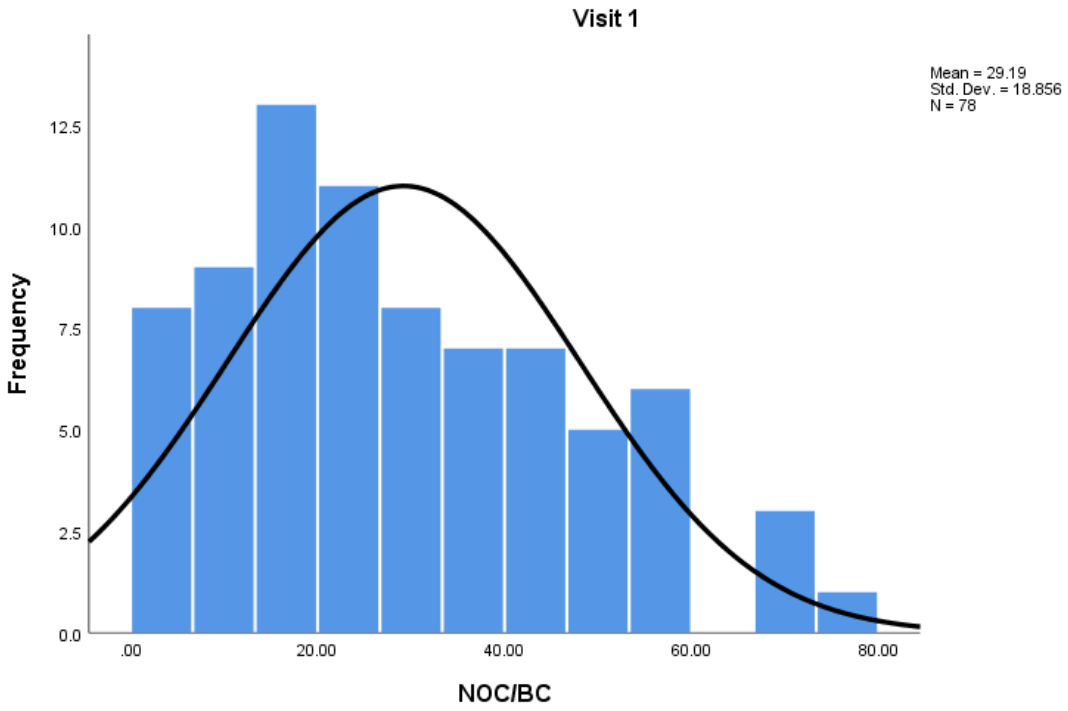
The main finding of the cross sectional study was that an automatic system based on the average NOC/BC calculated from the recorded lung sounds at lung bases can be used for differentiating IPF patients from non-IPF patients. Experimental results indicated that the automatic system can match the performance of the individual or average assessment of expert physicians in terms of

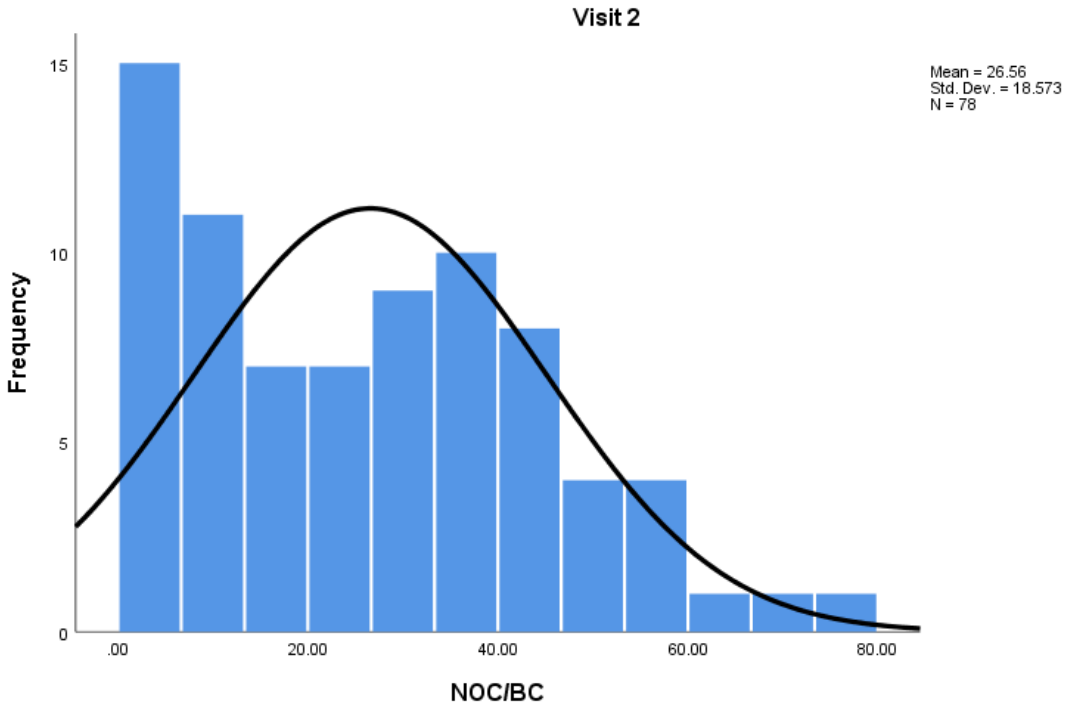
separating IPF patients from patients with other lung pathology. Although, assessment of HRCT scans should remain the gold standard for differentiating IPF patients from non IPF patients, it can be noted that the automatic system may have potential to help less expert clinicians to interpret their auscultatory findings of lung sounds. The performance evaluation of the automatic system on larger populations may further explore its usefulness in clinical environment. Note that both case studies were analyzed using the IEM-FD filter therefore these findings indicated that the IEM-FD filter has the potential to work on real data hence may play an important role in clinical findings.

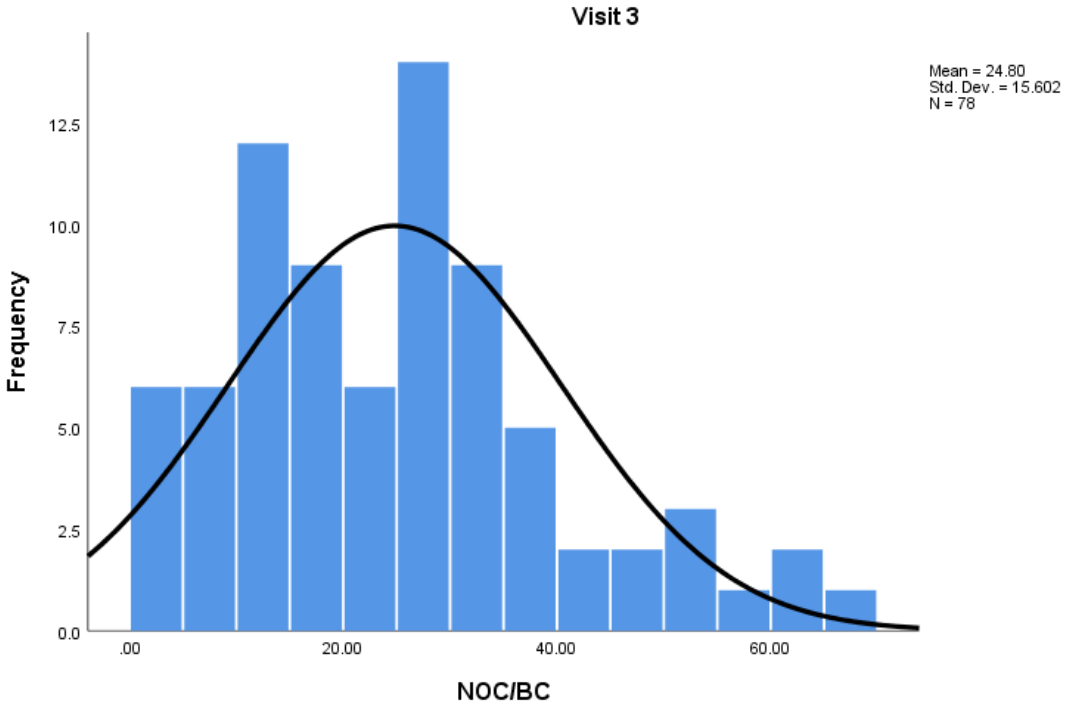
Finally, an automatic breath cycle detection algorithm was developed based on the estimation of the breathing phases. The algorithm was evaluated using the longitudinal dataset, cross sectional dataset, and a dataset recorded from healthy subjects. The key finding was that the automatic breath cycle detection algorithm can achieve a sensitivity ranging from 88.81 % to 97.37 %, a positive predictive value ranging from 93.01 % to 94.76 %, and F-score ranging from 90.86 % to 96.05 %. These results indicated the ability of the automatic breath cycle detection algorithm to automatically detect and count the number of breath cycles in recorded lung sounds. The future research should focus on evaluating the performance of the algorithm on datasets recorded from different populations (children, young, old) with or without cardiopulmonary diseases, which will further investigate the effect of different breathing patterns on the algorithm.

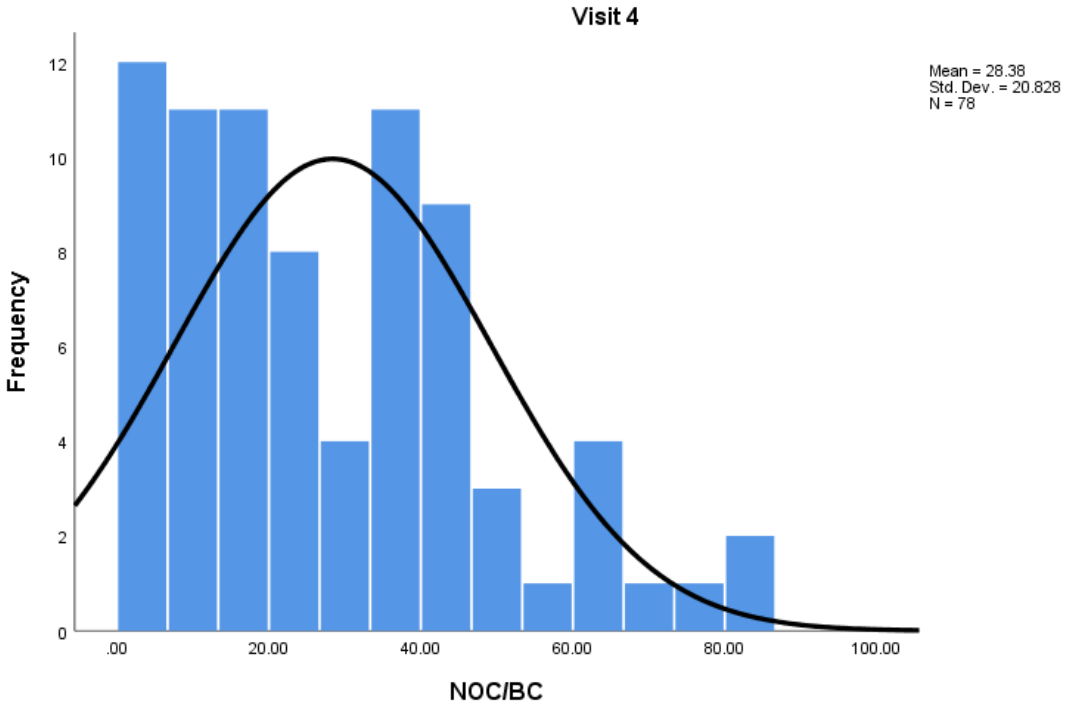
Appendix A

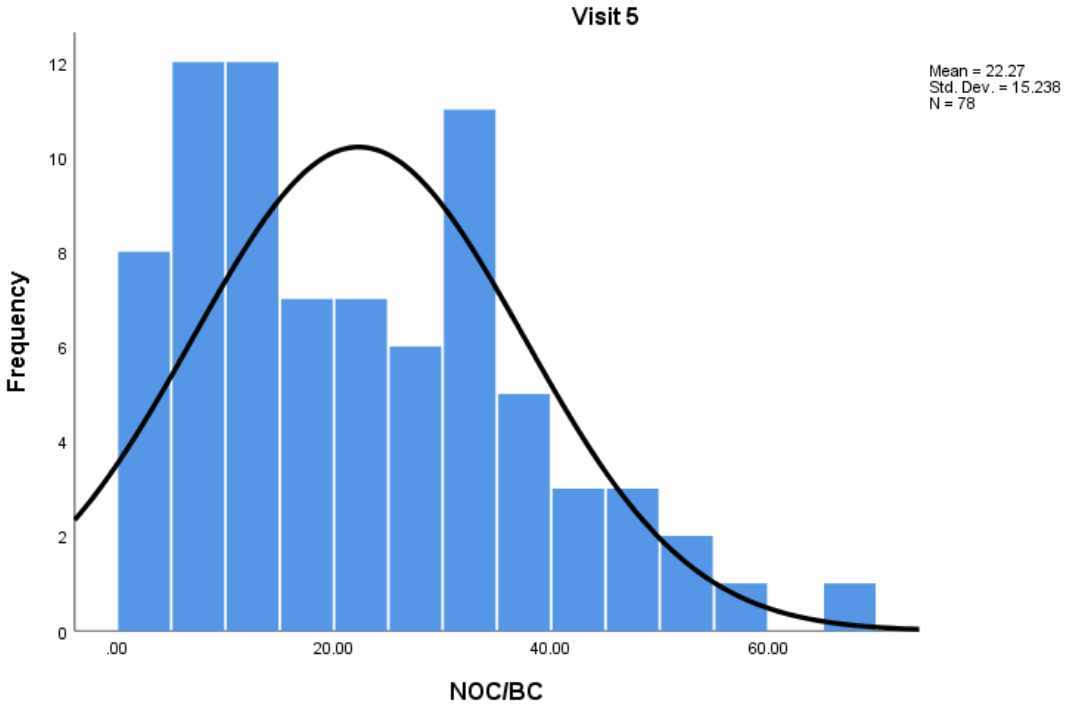
A.1 Histogram plots of the distribution of NOC/BC across 7 visits

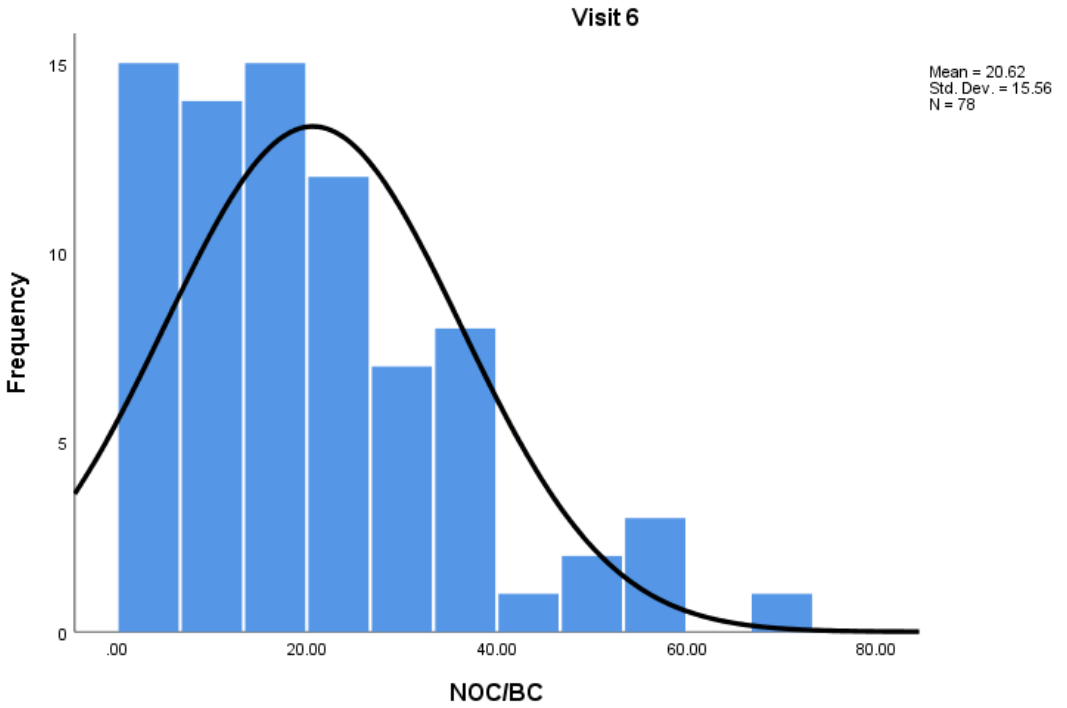





Visit 3
Mean = 24.80
Std. Dev. = 15.602
N = 78
Frequency
12.5
10.0
7.5
5.0
2.5
0.0
.00
20.00
40.00
60.00
NOC/BC


Visit 4
Mean = 28.38
Std. Dev. = 20.828
N = 78
Frequency
12
10
8
6
4
2
0
.00
20.00
40.00
60.00
80.00
100.00
NOC/BC


Visit 5
Mean = 22.27
Std. Dev. = 15.238
N = 78
Frequency
0
2
4
6
8
10
12
.00
20.00
40.00
60.00
NOC/BC


Visit 6
Mean = 20.62
Std. Dev. = 15.56
N = 78
Frequency
0
5
10
15
.00
20.00
40.00
60.00
80.00
NOC/BC

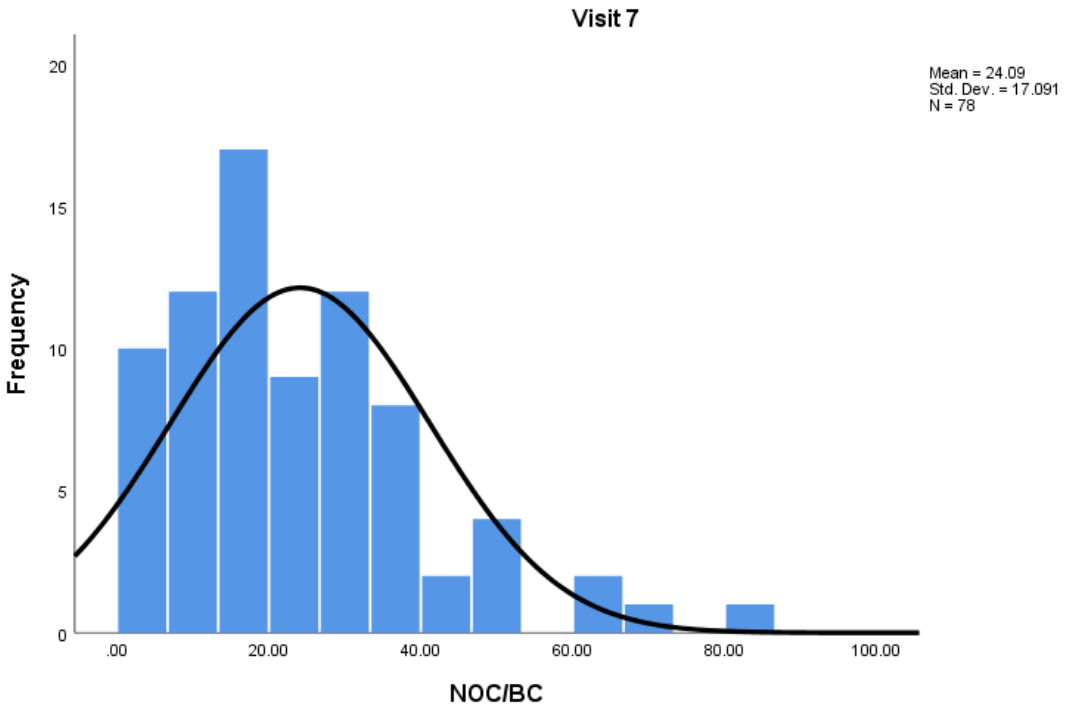


Figure A1 Histogram plots of the distribution of NOC/BC across 7 visits. The black line shows normal distribution.

A.2 Wilcoxon signed-rank test for NOC/BC across 7 visits

Table A1 Pairwise comparison over 7 visits using the Wilcoxon signed-rank test.

Pairwise comparison	Z	*p* value	*p* value after the Bonferroni correction
Visit 2 - Visit 1	-1.376[b]	0.169	1.000
Visit 3 - Visit 1	-1.910[b]	0.056	1.000
Visit 4 - Visit 1	-.194[b]	0.846	1.000
Visit 5 - Visit 1	-2.697[b]	0.007	0.147
Visit 6 - Visit 1	-3.481[b]	0.001	0.021[*]
Visit 7 - Visit 1	-2.670[b]	0.008	0.168
Visit 3 - Visit 2	-.981[b]	0.327	1.000
Visit 4 - Visit 2	-.483[c]	0.629	1.000
Visit 5 - Visit 2	-1.302[b]	0.193	1.000
Visit 6 - Visit 2	-2.668[b]	0.008	0.168
Visit 7 - Visit 2	-1.389[b]	0.165	1.000
Visit 4 - Visit 3	-1.375[c]	0.169	1.000
Visit 5 - Visit 3	-1.243[b]	0.214	1.000
Visit 6 - Visit 3	-2.147[b]	0.032	0.672
Visit 7 - Visit 3	-.868[b]	0.385	1.000
Visit 5 - Visit 4	-2.306[b]	0.021	0.441
Visit 6 - Visit 4	-2.917[b]	0.004	0.084
Visit 7 - Visit 4	-1.905[b]	0.057	1.000
Visit 6 - Visit 5	-.324[b]	0.746	1.000
Visit 7 - Visit 5	-.416[c]	0.677	1.000

Visit 7 - Visit 6	-1.382[c]	0.167	1.000
	b. Based on positive ranks. c. Based on negative ranks. *= Significant at $p < 0.05$		

A.3 19 reproducible acoustic features from sound files recorded at 6 posterior locations over complete study of 1 year in the IPF dataset

Table A2 The data in terms of mean, standard deviation, maximum, and minimum values of all 19 reproducible acoustic features from sound files recorded at 6 posterior locations over complete study of 1 year in the IPF dataset (Sgalla, 2017).

Repeatable acoustic features	**No. of patients**	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
C3 EW_200_500 Hz	19	0.000	0.594	0.226	0.107
C4 EW_75_200 Hz	19	0.000	0.789	0.573	0.105
C4 EW_200_500 Hz	19	0.000	0.487	0.164	0.095
sig_zerocross	19	0.012	0.077	0.038	0.009
sig_mfcc02	19	-12.280	-8.768	-10.158	0.576
sig_75_200 Hz_zerocross	19	0.023	0.031	0.027	0.001
sig_75_200 Hz_centroid	19	103.710	143.597	122.083	5.252
sig_200_500 Hz_rms	19	0.043	0.575	0.246	0.091
sig_200_500 Hz_lowenergy	19	0.461	0.928	0.692	0.077
sig_200_500 Hz_lowenergyASR	19	0.288	0.970	0.726	0.125
sig_200_500 Hz_zerocross	19	0.056	0.096	0.076	0.006
sig_200_500 Hz_std_meanframes	19	0.002	0.058	0.012	0.007
sig_200_500 Hz_std_medianframes	19	0.000	0.015	0.003	0.002
sig_500_1000 Hz_zerocross	19	0.140	0.160	0.148	0.002
sig_500_1000 Hz_rolloff85	19	635.193	847.290	750.090	34.009
sig_500_1000 Hz_centroid	19	572.324	699.717	630.687	17.062

C3_mfcc02	19	-12.304	-8.770	-10.164	0.578
C4_zerocross	19	0.012	0.083	0.040	0.010
C4_mfcc02	19	-12.281	-8.767	-10.157	0.576

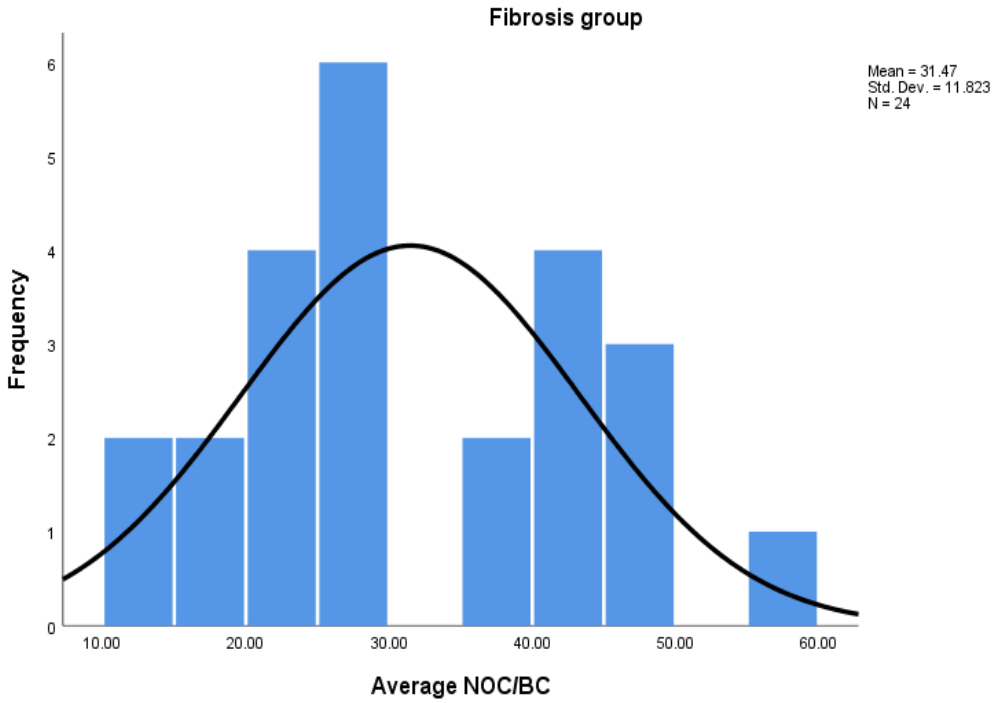
A.4 Average NOC/BC at each patient in fibrosis and non-fibrosis groups

Table A3 Average NOC/BC at each patient in fibrosis and non-fibrosis groups.

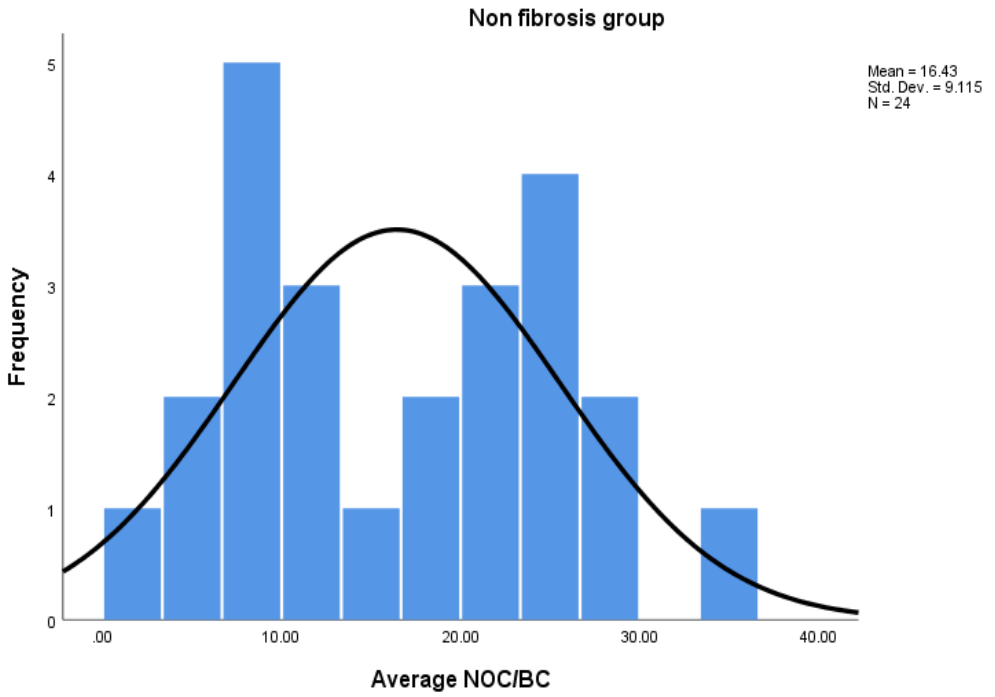
Fibrosis group				**Non fibrosis group**			
Patient Number	**Patient Gender**	**Patient age**	**Average NOC/BC**	**Patient Number**	**Patient Gender**	**Patient age**	**Average NOC/BC**
1	M	65	44.00	**1**	M	88	15.08
2	M	71	45.17	**2**	M	75	23.83
3	M	73	45.50	**3**	M	48	23.00
4	M	76	28.80	**4**	M	71	0.67
5	M	72	55.80	**5**	M	77	29.63
6	M	68	27.71	**6**	M	74	17.88
7	M	74	48.33	**7**	M	56	7.96
8	F	69	20.00	**8**	M	55	25.00
9	F	83	21.00	**9**	M	71	22.00
10	M	75	35.23	**10**	M	63	11.13
11	F	79	13.98	**11**	M	74	26.70
12	F	54	19.38	**12**	M	66	17.43
13	M	53	26.74	**13**	M	75	25.17
14	M	73	12.36	**14**	F	66	11.00
15	F	73	42.75	**15**	F	57	5.25
16	F	67	19.92	**16**	F	53	7.42

17	M	62	24.29	**17**	F	91	22.83
18	M	87	24.77	**18**	F	76	9.25
19	M	74	35.50	**19**	M	65	8.88
20	F	68	43.25	**20**	F	78	4.79
21	M	72	27.00	**21**	F	74	8.78
22	F	79	27.22	**22**	M	71	24.00
23	M	48	25.46	**23**	F	73	11.71
24	F	61	41.21	**24**	F	78	34.88

A.5 Histogram plots of the distribution of average NOC/BC at fibrosis group and non-fibrosis group



(a)



(b)

Figure A2 Histogram plots of the distribution of average NOC/BC at (a) fibrosis group, and (b) non-fibrosis group. The black line shows normal distribution.

A.6 Published conference papers

A.6.1 A dataset for systematic testing of crackle separation techniques

A dataset for systematic testing of crackle separation techniques

R. Pal and A. Barney

Abstract—**Pulmonary crackles are indicative of lung pathology and may be used for diagnosis and monitoring of disease. Many algorithms have been proposed to separate the crackle sounds from the breath noise, but a lack of standardized processes for evaluating their performance makes comparisons difficult. In this paper we propose a standard data set to be used for systematic comparative testing.**

I. Introduction

Sounds that come from the lungs can provide non-invasive diagnosis of pulmonary diseases. Lung sounds are divided into two categories: breath sounds and adventitious lung sounds. Breath sounds are heard in healthy as well as pathological lungs and result from the flow of air through the airways. Adventitious lung sounds are superimposed on breath sounds and can be an indication of pulmonary disorder [1]. Based on time domain characteristics, adventitious lung sounds can be further categorized into two classes: continuous and discontinuous. Continuous adventitious sounds with longer time duration (250 ms), are known as 'wheezes' [2]. Discontinuous adventitious sounds are called crackles and have a broadband frequency content ranging from 50 to 2000 Hz. The time duration of crackles is relatively short (less than 20 ms) [3] and they may be further characterized as either fine (high pitched) or coarse (low-pitched) sounds based on initial deflection width (IDW) two cycle deflection (2CD) and, occasionally, total deflection width (TDW) [4].

According to the American Thoracic Society, the average IDW and 2CD of fine crackles are 0.7 ms and 5 ms, and for coarse crackles are 1.5 ms and 10 ms, respectively [4]. According to Hoevers et al. [5] fine crackles have on average IDW = 0.5 ms, 2CD = 3.3 ms and total deflection width TDW = 4 ms, whereas coarse crackles have on average IDW = 1.0 ms, 2CD = 5.1 ms and TDW = 6.7 ms. According to Cohen et al. [6], fine crackles have average IDW = 0.9 ms and 2CD = 6.0 ms, and coarse crackles have average IDW = 1.25 ms and 2CD = 9.50 ms. Fine crackles are thought to be generated due to the explosive reopening of small airways that closed during the previous expiration [7]. Fine crackles are usually mid- to late-inspiratory events that follow a similar pattern for each consecutive inhalation. Fine crackles can be an indication of pneumonia, congestive heart failure and various pulmonary fibrotic diseases [8]. Coarse crackles are typically early inspiratory and expiratory events. Coarse crackles may be generated from fluid in small or medium airways and have a popping quality. These crackles can change pattern after coughing and are a symptom of, for example, chronic bronchitis, bronchiectasis and cystic fibrosis [9].

Much research effort has been expended in recent years in the automatic detection of crackles and their classification into inspiratory/expiratory, fine/coarse for the purpose of diagnosis or disease progression monitoring. Separation of crackles from normal breath sounds, as an initial processing stage can lead to better crackle characterization. Many methods have been proposed for this purpose including: the wavelet transform stationary non-stationary (WTST-NST) filter [2], the non-linear stationary non-stationary (ST-NST) filter [10], the wavelet packet transform (WPT) [11], the wavelet transform fractal dimension (WT-FD) [12,13], and the empirical mode decomposition fractal dimension (EMD-FD) filter [14].

Clearly, validation of algorithms is critical to their application in a clinical context, but a limitation in the field to date is that every study uses its own data and methodology for testing and therefore comparison of the relative strengths of different approaches under different conditions is limited. In this paper, we propose a standard data set for testing the performance of crackle separation techniques.

II. Dataset

Our dataset consists of three subsets.

- Test crackles comprising: (1) simulated fine crackles, (2) simulated coarse crackles, (3) real fine crackles, and (4) real coarse crackles.
- Test noise comprising: (1) white Gaussian noise, (2) colored noise with a spectrum matching that of a healthy breath sound recorded at the posterior right chest location.
- Test samples comprising: (1) real lung sounds with fine crackles, and (2) real lung sounds with coarse crackles.

A. Simulated fine and coarse crackles

The simulated fine and coarse crackles are generated using the mathematical function defined by Kiyokawa et al. [15], where the crackle waveform y(t) is given by:

$$y(t)=y_0(t)m(t) \quad (1)$$

with

$$y_0(t)=\sin(4\pi t^{\alpha}),\ \alpha=\frac{\log(0.25)}{\log(t_0)} \quad (2)$$

$$m(t)=0.5(1+\cos[2\pi(t^{0.5}-0.5)]) \quad (3)$$

On the basis of IDW and 2CD, three cases of fine and coarse crackles are generated: (1) American Thoracic Society definition [4] with: fine crackle IDW = 0.7 ms and

*Research supported by the NIHR Southampton Biomedical Research Centre, AAIR Charity and the Engineering and Physical Sciences Research Council.

R. Pal is with ISVR, University of Southampton, Southampton, UK (phone: +44 023 805937374 e-mail: r.pal@ soton.ac.uk).

A. Barney is with ISVR, University of Southampton, Southampton, UK (e-mail: ab3@ soton.ac.uk).

978-1-5386-1311-5/19/$31.00 ©2019 IEEE

2CD = 5 ms, coarse crackle IDW = 1.5 ms and 2CD = 10 ms; (2) based on Hoevers et al. [5] with fine crackle IDW = 0.5 ms and 2CD = 3.3 ms, coarse crackle IDW = 1 ms and 2CD = 5.1 ms and (3) based on Cohen [6] with fine crackle IDW = 0.9 ms and 2CD = 6 ms and coarse crackle IDW = 1.25 ms and 2CD = 9.5 ms. For each simulated crackle case, a set of ten identical crackles was generated.

B. Real fine and coarse crackles

We selected 10 real fine crackles from our lung sound database recorded from a patient with Idiopathic Pulmonary Fibrosis (IPF) and 10 real coarse crackles recorded from a patient with Bronchiectasis (B_r). All selected crackles followed at least three of the criteria defined by Murphy et al. [16] as characteristic of a pulmonary crackle.

C. Simulated test signals

Each set of crackles was buried in two types of noise: 1) white Gaussian noise and 2) noise with the same spectrum as breath noise from a healthy subject measured over the lung bases on the right-hand side of the back. Signal to noise ratio ranged from -10 to 10 dB in steps of 1 dB and for each SNR, 201 test signals were generated each containing 10 crackles.

D. Real lung sounds with fine and coarse crackles

In addition to the simulated data, two examples of real lung sounds, one from a patient with Idiopathic Pulmonary Fibrosis (IPF) with predominantly fine late-inspiratory crackles and one from a patient with Bronchiectasis (B_r) with coarse, mainly expiratory crackles, are included in the dataset. The precise location of each crackle in these sound samples has not been clinically verified; nevertheless, they can provide useful comparisons between separation techniques to evaluate their performance in separating the stationary from the non-stationary parts of the lung sound.

Table 1 summarizes the different test cases.

TABLE I. SUMMARY OF THE TEST DATASET

Cases	Nc	Diagnosis	IDW& 2CD (ms)	BN	SNR
Simulated fine crackles	10	NA	0.7&5 [4]	BR_N	-10 to 10 dB
				W_N	
			0.5& 3.3 [5]	BR_N	
				W_N	
			0.9&6 [6]	BR_N	
				W_N	
Simulated coarse crackles	10	NA	1.5&10 [4]	BR_N	
				W_N	
			1&5.1 [5]	BR_N	
				W_N	
			1.25& 9.5 [6]	BR_N	
				W_N	
Real fine crackles	10	IPF	ND	BR_N	
				W_N	
Real coarse crackles	10	B_r	ND	BR_N	
				W_N	
Real breath sound with fine crackles	ND	IPF	ND	NBS	ND
Real breath sound with coarse crackles	ND	B_r	ND	NBS	ND

Nc: Number of crackles; ND: Not defined; NA=Not applicable; IPF: Idiopathic pulmonary fibrosis; B_r: Bronchiectasis; BN: Background noise; BR_N: Breath noise; W_N: White Gaussian noise; NBS: Normal breath sound.

As a case study we have analyzed the crackle separation performance of the wavelet transform stationary non-stationary filter (WTST-NST) [2], and the wavelet transform fractal dimension filter (WT-FD) [12,13] and we present here the results of that part of the analysis using the simulated fine and coarse crackles, and the real fine and coarse crackles buried in noise with the spectrum of breath noise. For the detailed working process of the WTST-NST filter and the WT-FD filter we refer readers to [2], [12,13].

III. METHODS

The WTST-NST and the WT-FD filters separate the signal into non-stationary and stationary parts. In a successful separation, crackles would appear only in the non-stationary part of the signal and the normal breath sounds, which are considered to be wide sense stationary, within the stationary part of the signal. The WT is computed using the Daubechies-4 wavelet [13]. The sound files are sampled at 44100 Hz and the separation algorithms are implemented using the MATLAB programming language. Table II. Shows the parameters used for implementing the WTST-NST and the WT-FD filters.

The WTST-NST [2] is an iterative multi-resolution decomposition and multi-resolution reconstruction (MRD-MRR) scheme which provides separation of WT coefficients related to a signal of interest from the background noise at each level using some threshold value (Threshold= F_{adj} . SD, where F_{adj} is a multiplication factor and SD is the standard deviation of the waveform).

The WT-FD filter [12,13] overcomes the need for empirical setting of a multiplication factor (F_{adj}) in the WTST-NST filter and provides automatic separation of WT coefficients related to signal of interest and background noise using the fractal dimension technique. The fractal dimension of a signal is estimated using a sliding window of length W_{FD}=int(kf_s), where 'int' indicates rounding to integer value, f_s is the sampling frequency and k is a multiplication factor.

TABLE II. PARAMETERS USED FOR THE WTST-NST AND THE WT-FD METHODS

Parameters	WTST-NST	WT-FD
Number of samples (N)	32,768	32,768
Number of decomposition levels (M)	$M=\log_2(N)=15$	M =1[13]
Sampling frequency (f_S)	44,100 Hz	
Accuracy level (β)	0.1	0.1
Number of iterations (L)	1	1
Multiplication factor	F_{adj}=3 [2]	k=0.006 [13]

We propose a measure of the quality of separation using the average power of the non-stationary signal estimated from the separation algorithm compared to the average power of the crackle signal before insertion into the noise.

The average power for each test signal is calculated using Eq. (4).

$$\text{Average power}=\left\lfloor\frac{\left(\sum_{n=1}^{N}|Y(n)|^2\right)}{N}\right\rfloor \quad (4)$$

TABLE III. PERFORMANCE OF THE WTST-NST FILTER

Cases	BN	WTST-NST							
		$\bar{R}_{QF}$ (SD)		$\bar{E}_{QF}$ (SD)		Ovr_{Est} (if $\bar{R}_{QF} > \bar{E}_{QF}$)		Unr_{Est} (if $\bar{R}_{QF} < \bar{E}_{QF}$)	
						$\%Ovr_{Est}$		$\%Unr_{Est}$	
		S^-	S^+	S^-	S^+	S^-	S^+	S^-	S^+
Simulated fine crackles	BR_N	0.7367 (0.1384)	0.2393 (0.1201)	0.5744 (0.1036)	0.1995 (0.1001)	22.0	16.6	NA	
	BR_N	0.7368 (0.1383)	0.2396 (0.1201)	0.6057 (0.1324)	0.1902 (0.0958)	17.7	20.6		
	BR_N	0.7367 (0.1386)	0.2399 (0.1199)	0.5774 (0.0993)	0.2002 (0.0995)	21.6	16.5		
Simulated coarse crackles	BR_N	0.7368 (0.1384)	0.2393 (0.1200)	0.4929 (0.1357)	0.1230 (0.1057)	33.1	48.6		
	BR_N	0.7369 (0.1383)	0.2395 (0.1202)	0.5667 (0.1128)	0.1928 (0.0971)	23.0	19.4		
	BR_N	0.7370 (0.1381)	0.2406 (0.1204)	0.5256 (0.1239)	0.1258 (0.0980)	28.6	47.7		
Real fine crackles	BR_N	0.7369 (0.1383)	0.2396 (0.1202)	0.5949 (0.1295)	0.1907 (0.0918)	19.2	20.4		
Real coarse crackles	BR_N	0.7371 (0.1383)	0.2398 (0.1204)	0.4722 (0.2337)	0.0415 (0.0074)	35.9	82.6		

BN: Background noise, BR_N: Breath noise, $\bar{R}_{QF}$: mean of the reference quality factor, $\bar{E}_{QF}$: mean of the estimated quality factor, SD: Standard deviation, S^-: Signal to noise ratio from -10 to 0 dB, S^+: Signal to noise ratio from 1 to 10 dB, Ovr_{Est}: Overestimation, Unr_{Est}: Underestimation, NA: Not applicable.

TABLE IV. PERFORMANCE OF THE WT-FD FILTER

Cases	BN	WT-FD							
		$\bar{R}_{QF}$ (SD)		$\bar{E}_{QF}$ (SD)		Ovr_{Est} (if $\bar{R}_{QF} > \bar{E}_{QF}$)		Unr_{Est} (if $\bar{R}_{QF} < \bar{E}_{QF}$)	
						$\%Ovr_{Est}$		$\%Unr_{Est}$	
		S^-	S^+	S^-	S^+	S^-	S^+	S^-	S^+
Simulated fine crackles	BR_N	0.7368 (0.1383)	0.2396 (0.1203)	0.2110 (0.0586)	0.1849 (0.0850)	71.3	22.8	NA	
	BR_N	0.7368 (0.1384)	0.2395 (0.1200)	0.3156 (0.0930)	0.1948 (0.0977)	57.1	18.6		
	BR_N	0.7368 (0.1383)	0.2394 (0.1200)	0.1700 (0.0479)	0.1748 (0.0730)	76.9	26.9		
Simulated coarse crackles	BR_N	0.7372 (0.1381)	0.2393 (0.1189)	0.0828 (0.0360)	0.1235 (0.0285)	88.7	48.3		
	BR_N	0.7368 0.1383)	0.2394 (0.1202)	0.2176 (0.0683)	0.1868 (0.0881)	70.4	21.9		
	BR_N	0.7369 (0.1382)	0.2394 (0.1196)	0.0899 (0.0368)	0.1267 (0.0305)	87.8	47.0		
Real fine crackles	BR_N	0.7370 (0.1382)	0.2394 (0.1200)	0.3884 (0.1008)	0.1944 (0.0975)	47.2	18.7		
Real coarse crackles	BR_N	0.7369 (0.1382)	0.2404 (0.1201)	0.1343 (0.0645)	0.1713 (0.0685)	81.7	28.7		

BN: Background noise, BR_N: Breath noise, $\bar{R}_{QF}$: mean of the reference quality factor, $\bar{E}_{QF}$: mean of the estimated quality factor, SD: Standard deviation, S^-: Signal to noise ratio from -10 to 0 dB, S^+: Signal to noise ratio from 1 to 10 dB, Ovr_{Est}: Overestimation, Unr_{Est}: Underestimation, NA: Not applicable.

Where $|Y(n)|$ is the absolute value of the test signal and N =32,768 is the number of samples.

For each SNR the 201 test signals are passed through each algorithm, the average power is calculated and the mean over all test signals is taken.

Two quality factors are then calculated at each SNR: (1) reference quality factor (R_{QF}) using Eq. (5), and (2) estimated quality factor (E_{QF}) using Eq. (6):

$$R_{QF}=\frac{I_{AVG}(W)-R_{AVG}(W)}{I_{AVG}(W)} \quad (5)$$

$$E_{QF}=\frac{I_{AVG}(W)-NS_{AVG}(W)}{I_{AVG}(W)} \quad (6)$$

Where I_{AVG} is the average power of the test signals, R_{AVG} represents the reference power calculated from the crackle signal before inserting into the noise and NS_{AVG} is the average power of the non-stationary signal estimate.

If $E_{QF}<R_{QF}$ for a given SNR then the non-stationary signal is considered overestimated and likewise for $E_{QF}>R_{QF}$ the non-stationary signal is considered under-estimated. The percentage of overestimation ($\%Ovr_{Est}$) and underestimation ($\%Unr_{Est}$) are calculated using the Eq. (7) and Eq. (8), respectively.

$$\%\ Ovr_{Est}=\left|\frac{mean(R_{QF})-mean(E_{QF})}{mean(R_{QF})}\right|\times 100 \quad (7)$$

$$\%\ Unr_{Est}=\left|\frac{mean(R_{QF})-mean(E_{QF})}{1-mean(R_{QF})}\right|\times 100 \quad (8)$$

Where the mean is calculated over each SNR value in the ranges -10 to 0 dB and 1 to 10 dB.

IV. RESULTS

Input signal average power, estimates of non-stationary signal average power & stationary signal average power and reference average power for one typical set of test signals

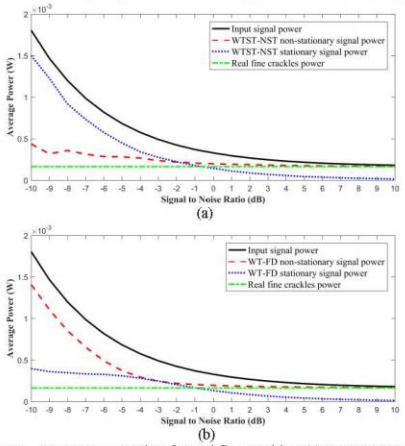


Figure 1. Average power plots for real fine crackles, (a) WTST-NST filter (Table III, Case Real fine crackles), and (b) WT-FD filter (Table IV, Case Real fine crackles).

(Table III and Table IV, Case Real fine crackles) are shown in Fig. 1. The standard deviation at each SNR for the input signal average power was between 0.49 and 0.92% of the signal power and for the non-stationary signal estimate was between 0.63 and 98.82% of the signal power.

Reference quality factor and estimated quality factor for one typical set of test signals (Table III and Table IV, Case Real fine crackles) are shown in Fig. 2. The standard deviation at each SNR for the reference quality factor was between 0.05 and 5.18% of the quality factor and for the estimated quality factor was between 3.97 and 39.72% of the quality factor.

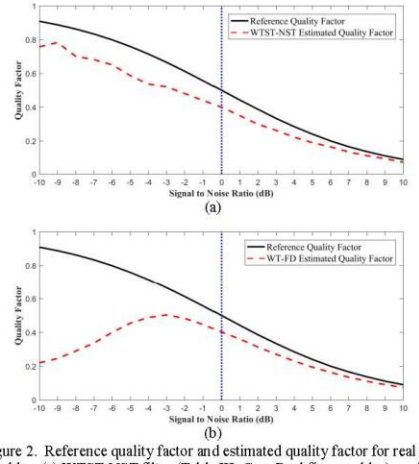


Figure 2. Reference quality factor and estimated quality factor for real fine crackles, (a) WTST-NST filter (Table III, Case Real fine crackles), and (b) WT-FD filter (Table IV, Case Real fine crackles).

From Table III and Table IV, and Figs 1 & 2, it can be seen that neither the WTST-NST nor the WT-FD underestimates the non-stationary signal at any SNR. Generally, for noise with the spectrum of a real breath sound, WTST-NST overestimates by no more than 49% over the whole range -10 dB to 10 dB (except in the case of real coarse crackles in the range 1 to 10 dB where overestimation is much greater). WT-FD produces far higher over estimates over the range -10 to 0 dB but performs broadly equivalently to WTST-NST between 1 and 10 dB. A more detailed comparison could be made if the performance on the whole test data set were considered.

For the test dataset we have selected two types of background noise: white Gaussian noise to provide a standard test of SNR performance and colored noise matched to breath sound characteristics. We note that the spectrum of the colored noise is derived from measurements on a healthy person and may not give an ideal measure of absolute performance under pathological conditions if the specific disease under consideration modifies the breath sounds as well as adding crackles. Noise modeled from breath sounds measured at different chest locations and from subjects with different pathological conditions could provide additional information on algorithm performance.

V. Conclusion

We conclude that our data set and quality criteria can provide a useful comparative performance analysis of different crackle separation algorithms. We offer this as a standard data set for testing separation algorithms which is available at https://doi.org/10.5258/SOTON/D0801. A future extension to the quality criteria will allow the comparative performance of crackle identification (number of crackles detected) and characterization (performance as an estimator of IDW, 2CD) algorithms to be similarly evaluated.

References

[1] A. R. A. Sovijärvi, F. Dalmasso, J. Vanderschoot, L. P. Malmberg, G. Righini, and S. A. T. Stoneman, "Definition of terms for applications of respiratory sounds," *Eur Respir. Rev*, vol. 10, no. 77, pp. 597–610, 2000.

[2] L. J. Hadjileontiadis and S. M. Panas, "Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 12, pp. 1269–1281, 1997.

[3] R. Loudon and R. L. Murphy, "Lung sounds," *Am. Rev. Respir. Dis.*, vol. 130, no. 4, pp. 663–73, 1984.

[4] G. Charbonneau, E. Ademovic, B. M. G. Cheetham, L. P. Malmberg, J. Vanderschoot, and A. R. A. Sovijarvi, "Basic techniques for respiratory sound analysis," *Eur Respir. Rev*, vol. 10, no. 77, pp. 625–635, 2000.

[5] J. Hoevers and R. G. Loudon,"Measuring crackles," *Chest*, vol. 98, no. 5, pp. 1240–1243, 1990.

[6] A. Cohen, "Signal Processing Methods for Upper Airway and Pulmonary Dysfunction Diagnosis," *IEEE Eng. Med. Biol. Mag.*, vol. 9, no. 1, pp. 72–75, 1990.

[7] M. Munakata, H. Ukita, I. Doi, Y. Ohtsuka, and Y. Homma, "Spectral and waveform characteristics of fine and coarse crackles," *Thorax*, vol. 46, no. 9, pp. 651–657, 1991.

[8] R. X. A. Pramono, S. Bowyer, and E. Rodriguez-Villegas, "*Automatic adventitious respiratory sound analysis: A systematic review*," *PLoS ONE*, vol. 12, no. 5, 2017.

[9] S. S. Kraman, "Lung sounds: An Introduction to the Interpretation of the Auscultatory Finding," *Northbrook, IL: Amer. College of Chest Physician*, pp. 19-21, 1993.

[10] M. Ono, K. Arakawa, M. Mori, T. Sugimoto, and H. Harashima, "Separation of fine crackles from vesicular sounds by a nonlinear digital filter," *IEEE Trans. Biomed. Eng.*, vol. 36, no. 2, pp. 286–291, 1989.

[11] M. Bahoura, and X. Lu, "Separation of Crackles From Vesicular Sounds Using Wavelet Packet Transform," *2006 IEEE Int. Conf. Acoust. Speed Signal Process. Proc.*, vol. 2, pp. 1076-1079, 2006.

[12] L. J. Hadjileontiadis, "Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding - Part I: Methodology," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 6, pp. 1143–1148, 2005.

[13] L. J. Hadjileontiadis, "Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding - Part II: Application results," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 6, pp. 1050–1064, 2005.

[14] L. J. Hadjileontiadis, "Empirical Mode Decomposition and Fractal Dimension Filter: A Novel Techinque for Denoising Explosive Lung Sounds," *IEEE Eng. Med. Biol. Mag.*, vol. 26, no.1, pp. 30–39, 2007.

[15] H. Kiyokawa, M. Greenberg, K. Shirota, and H. Pasterkamp, "Auditory detection of simulated crackles in breath sounds," *Chest*, vol. 119, no. 6, pp. 1886–1892, 2001.

[16] R. L. H. Murphy, JR., EA. D. Bono, and F. Davidson, "Validation of an automatic crackle (rale) counter," *Am. Rev. Respir.*, vol. 140, no. 4, pp. 1017-1020, 1988.

A.6.2 Pulmonary Crackle Detection Using the Hilbert Energy Envelope

Pulmonary Crackle Detection Using the Hilbert Energy Envelope

Ravi Pal(✉) and Anna Barney

Institute of Sound and Vibration Research, University of Southampton, Southampton SO17 1BJ, UK
r.pal@soton.ac.uk

Abstract. This paper presents a method for automatic pulmonary crackle detection based on the Hilbert energy envelope (HEE). Automatic detection of crackles in lung sounds offers a non-invasive way of monitoring or diagnosing cardiopulmonary diseases. The algorithm is divided into four main steps: (a) preprocessing, (b) estimation of HEE, (c) thresholding, and (d) applying time width conditions based on crackle two-cycle deflection and initial deflection width. Its performance is tested using a publicly available lung sound dataset of fine and coarse crackles and evaluated by the sensitivity (95.7%), positive predictive value (89.5%), and F-score (91.7%) for crackle detection. The good detection performance indicates the potential of the HEE-based algorithm as an automatic method for crackle detection in lung sound recordings.

Keywords: Pulmonary crackle · Automatic pulmonary crackle detection · Hilbert Energy Envelope (HEE) Algorithm

1 Introduction

In this paper we present, a method for automatic pulmonary crackle detection based on the Hilbert energy envelope (HEE).

Pulmonary crackles are short-lived, explosive lung sounds which are superimposed on normal breath sounds in some pathological lung conditions [1]. Crackles can provide valuable diagnostic information regarding different cardiopulmonary diseases including cystic fibrosis, pneumonia, fibrosing alveolitis, bronchiectasis, sarcoidosis, congestive heart failure, and asbestosis [2].

A traditional stethoscope offers a non-invasive way of examining the lung condition by listening to lung sounds through the chest wall, however the interpretation of the sound is highly subjective and depends on the expertise and hearing ability of the physician [3]. Visual detection of crackle sounds through analysis of a recording of the acoustic lung sound signal made with an electronic stethoscope [4], can also be very dependent on the expertise of the analyst. Computerized detection of crackles can, however, overcome these limitations and providing an objective way to detect crackle sounds [5].

In recent years, several automatic methods have been proposed for detecting crackles in lung sounds: time-varying autoregressive algorithm [3], fractal dimension and box filtering algorithm [5], first derivative absolute value (FDAV) based time

T. Jarm et al. (Eds.): EMBEC 2020, IFMBE Proceedings 80, pp. 994–1003, 2021.
https://doi.org/10.1007/978-3-030-64610-3_111

domain analysis [6], and fractal dimension detector (FDD) [7]. These methods show good crackle detection ability for high quality sound recordings, but as mentioned in [5] have not been widely tested with recordings made in clinical settings where movement artefacts and environmental noise are commonplace. A crackle detector for a clinical setting needs high crackle detection accuracy and robustness to noise, further, low computational complexity can be advantageous for rapid processing to support clinical decision-making.

Figure 1 displays the waveform of a typical crackle. Crackles can be characterized in the time domain by their initial deflection width (IDW) and their two-cycle deflection width (2CD) and may be divided into fine crackles (mean IDW = 0.7 ms; mean 2CD = 5 ms) and coarse crackles (mean IDW = 1.5 ms; mean 2CD = 10 ms) [8].

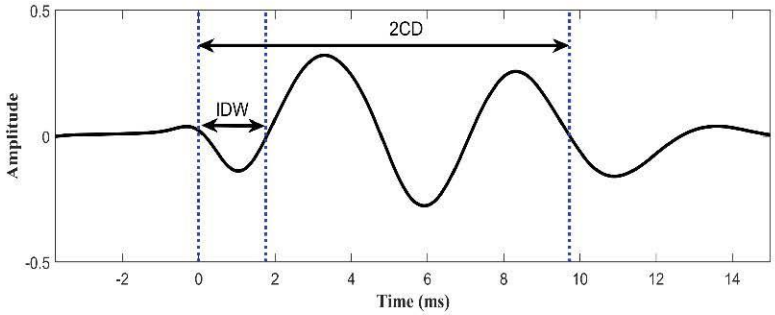


Fig. 1. The waveform of a typical crackle showing the characteristic time domain features: initial deflection width (IDW) and two-cycle deflection width (2CD).

The next section of this paper presents the HEE algorithm for crackle detection. Test data and the quantitative evaluators used for performance analysis are discussed in Sect. 3. Section 4 contains the results and discussion with conclusions presented in Sect. 5.

2 Hilbert Energy Envelope Algorithm

The Hilbert energy envelope (HEE) constructs an energy envelope for a signal using the Hilbert transform to determine its instantaneous characteristics. Sharma et al. [9] presented an algorithm based on the HEE for heart rate extraction from acoustic recordings at the neck. Here the concept is adapted for automatic crackle detection. The process is shown schematically in Fig. 2. The HEE algorithm was developed using Matlab (R2019a) programming language.

Figure 3 shows a worked example for a 0.137 s section of a lung sound signal recorded from a patient with idiopathic pulmonary fibrosis (Fig. 3(a)). The location of each crackle has been audio-visually identified by an experienced pulmonary acoustics researcher and marked with an arrowhead.

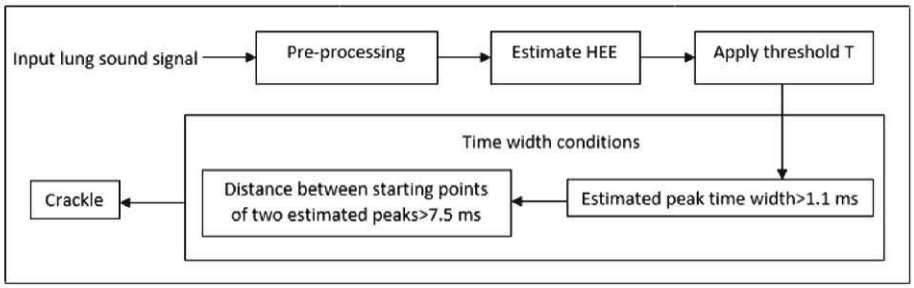


Fig. 2. Block diagram of the HEE algorithm for crackle detection.

The steps of the algorithm are as follows:

2.1 Pre-processing

The lung sound signal ($cy[n]$) is pre-processed using a 6th order Butterworth high pass filter with cut off frequency 75 Hz, (Fig. 3(b)).

2.2 Estimation of Hilbert Energy Envelope

The instantaneous characteristics of the preprocessed signal ($y_q[n]$) are calculated using the analytical function $z[n]$ [9].

$$z[n] = y_q[n] + ix[n] \tag{1}$$

Where, $y_q[n]$ is the preprocessed input signal, $x[n]$ is the Hilbert transform of the preprocessed input signal and $i = \sqrt{-1}$.

Using the analytical function in (1), the instantaneous amplitude ($a[n]$) and the instantaneous frequency ($\omega[n]$) are estimated.

$$a[n] = \sqrt{(y_q[n])^2 + (x[n])^2} \tag{2}$$

$$\omega[\mathrm{n}] = \tan^{-1}\left(\frac{x[n]}{y_q[n]}\right) \tag{3}$$

The energy envelope $H(y_q[n])$ is then calculated using (4) and is illustrated in Fig. 3(c):

$$H(y_q[n]) = |a[n]|^2 = (y_q[n])^2 + (x[n])^2, \quad 1 \leq \mathrm{n} \leq \mathrm{N} \tag{4}$$

where N is the total number of samples in the input signal.

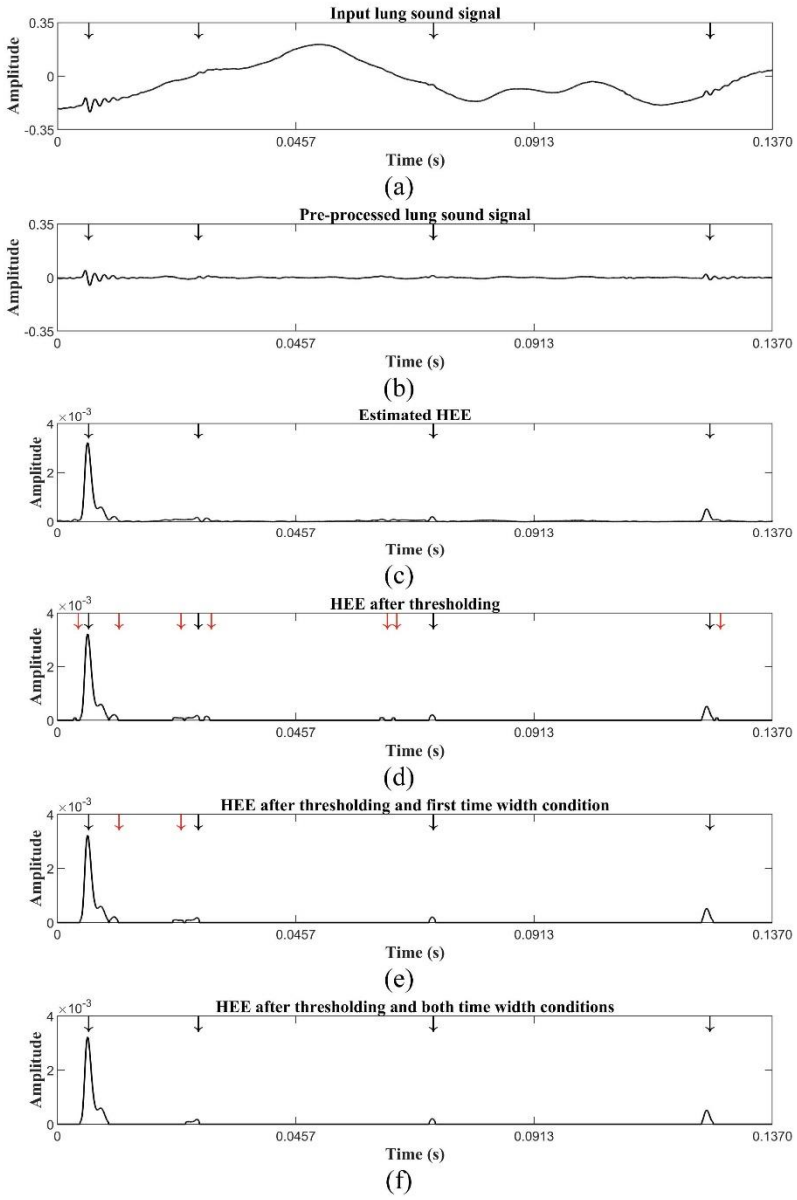


Fig. 3. Worked example of the HEE algorithm based on crackles time domain features. (a) A time section of 0.137 s lung sound data recorded from a patient with idiopathic pulmonary fibrosis (location of the crackles is marked with black arrowheads); (b) Pre-processed signal; (c) Estimated Hilbert energy envelope of the pre-processed signal; (d) Output after thresholding; (e) Removal of unwanted envelope peaks using first time width condition; (f) Eliminating remaining unwanted peaks using the second time width condition.

2.3 Thresholding

An energy threshold T is calculated using the estimated energy envelope $H(y_q[n])$:

$$T = m_f \times max\{H(y_q[n])\}, \quad 1 \leq n \leq N \tag{5}$$

where m_f is a multiplication factor We have selected the value of the multiplication factor to be m_f = 0.025. This choice is justified in Sect. (3.4). Threshold T is applied to the energy envelope $(H(y_q[n])$ and only values greater than T are kept. The estimated envelope signal after thresholding is displayed in Fig. 3(d) where the peaks related to crackles are shown using the black arrowheads and the peaks not related to crackles are marked using red arrowheads.

2.4 Time Width Conditions

To reduce the number of false crackle detections, two conditions are empirically selected based on typical values for the crackle time domain features: IDW and 2CD [8]:

(1) HEE peaks with time width less than 1.1 ms (i.e. the average of the mean IDW for fine and coarse crackles respectively) are discarded, (Fig. 3(e)).
(2) To prevent multiple detections of the same crackle and to eliminate any remaining false peaks after applying the first time width condition, the distance between the starting points of two envelope peaks is considered. If the distance between those starting points is greater than 7.5 ms (i.e. the average of the mean 2CD for fine and coarse crackles respectively), both peaks are considered to be crackles; otherwise only the peak with longer duration is considered to be a crackle and the other is discarded. The output, after applying this condition, is shown in Fig. 3(f), where it can be seen that all false peaks have been eliminated.

3 Dataset and Performance Evaluators

3.1 Dataset

The HEE algorithm was tested using a publicly available data set [10] of simulated and real fine and coarse crackles which can be embedded in two types of background noise: breath noise, and Gaussian white noise. The dataset also contains a sample of a real lung sound with fine crackles recorded from a patient with idiopathic pulmonary fibrosis and a sample of a real lung sound with coarse crackles recorded from a patient with bronchiectasis. The real lung sound files were recorded using an electronic stethoscope and all files in the dataset are sampled at 44100 Hz. Table 1 shows the cases selected from the dataset for performance analysis, SNRs tested ranged from 0 to 10 dB because for SNR levels less than 0 dB either a large number of false crackles were detected or estimated peaks, which may correspond to crackles, started to combine with each other due to background noise. Note that although the data set also

provides the option of embedding the crackles in Gaussian white noise, performance in that condition was not tested in this study.

Table 1. Cases generated from the test dataset [10].

Cases	IDW & 2CD (ms)	N_C	D_g	BN	SNR
Simulated fine crackles	0.7 & 5 [8]	10	NA	BR_N	0 to 10 dB
	0.5 & 3.3 [11]	10	NA	BR_N	
	0.9 & 6 [12]	10	NA	BR_N	
Simulated coarse crackles	1.5 & 10 [8]	10	NA	BR_N	
	1 & 5.1 [11]	10	NA	BR_N	
	1.25 & 9.5 [12]	10	NA	BR_N	
Real fine crackles	ND	10	IPF	BR_N	
Real coarse crackles	ND	10	B_r	BR_N	
Real breath sound with fine crackles	ND	ND	IPF	NBS	ND
Real breath sound with coarse crackles	ND	ND	B_r	NBS	ND

N_C: Number of crackles; Dg: Diagnosis; ND: Not defined; NA = Not applicable; IPF: Idiopathic pulmonary fibrosis; Br: Bronchiectasis; BN: Background noise; BR_N: Breath noise; NBS: Normal breath sound; SNR: Signal to noise ratio.

3.2 Performance Evaluators

Three parameters were used to evaluate the crackle detection performance of the algorithm: sensitivity (SE), positive predictive value (PPV), and F-score (F_1) [13] where:

$$F_1 = 2 \times \frac{SE \times PPV}{SE + PPV} \tag{6}$$

For each SNR tested, 501 test samples were generated (see Sect. 3.3). The SE, PPV and F_1 were calculated for each sample and the average over all test samples was generated. Note that the case of temporally overlapped crackles was not considered and if more than one crackle lay under a single HEE peak, or if, due to background noise, two or more peaks were connected with each other, then they were considered as one crackle.

3.3 Selection of the Number of Test Samples

Figure 4, shows the average F_1 of the HEE algorithm in the cases of real fine crackles (RFC) and real coarse crackles (RCC) at 5 dB SNR as the number of test samples is increased from 1 to 2001 in steps of 1. We note that average F-score is approximately independent of the number of samples when the number of samples exceeds 500. The blue vertical dotted line indicates the selected number of samples for all tests reported here.

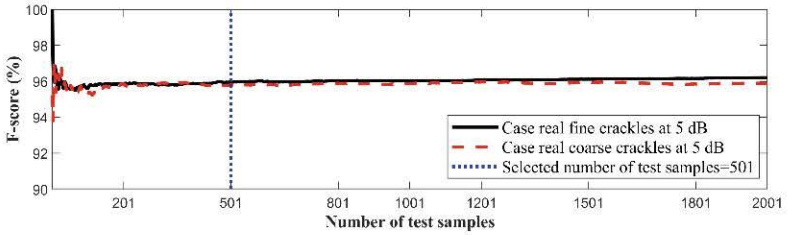


Fig. 4. Selection of number of test samples to eliminate random variation in the HEE algorithm crackle detection performance using real fine crackles case and real coarse crackles case at an SNR of 5 dB.

3.4 Selection of Multiplication Factor (m_f)

As mentioned in Sect. 2.3, those HEE peaks, which are likely to correspond to crackles, are selected using a threshold, T. The threshold value is calculated (Eq. (5)) using a multiplication factor m_f.

Figure 5 shows the performance of the HEE algorithm in terms of SE and PPV in the case of real fine crackles and real coarse crackles at an SNR of 5 dB as the multiplication factor ranges from 0.01 to 0.03, in steps of 0.001. We observe that, for $m_f > 0.025$ the SE drastically decreases and for $m_f < 0.025$ the PPV gradually starts to decrease. The selected multiplication factor $m_f = 0.025$ is marked on Fig. 5 and is chosen to avoid both these regions.

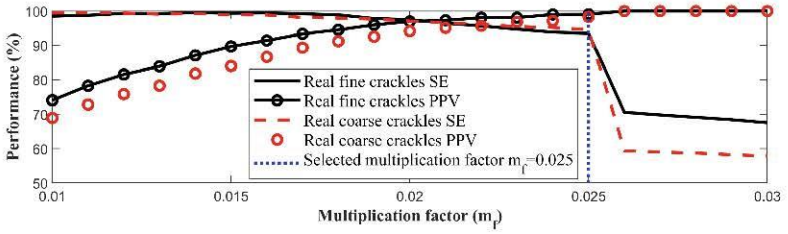


Fig. 5. Selection of multiplication factor m_f using real fine and coarse crackles cases (Table 1) at 5 dB SNR.

4 Results and Discussion

Figures 6(a) and (b) show the F-score plots for the HEE algorithm for fine (real and simulated) and coarse (real and simulated) crackles, respectively in the SNR range of 0 to 10 dB. It can be observed that for SNR > 5 dB the F-score plots start to converge to 100% in both cases, with generally faster convergence fine crackles. The performance of the HEE algorithm in terms of SE, PPV and F_1 on the whole test data with SNR = 5 dB, is shown in Table 2.

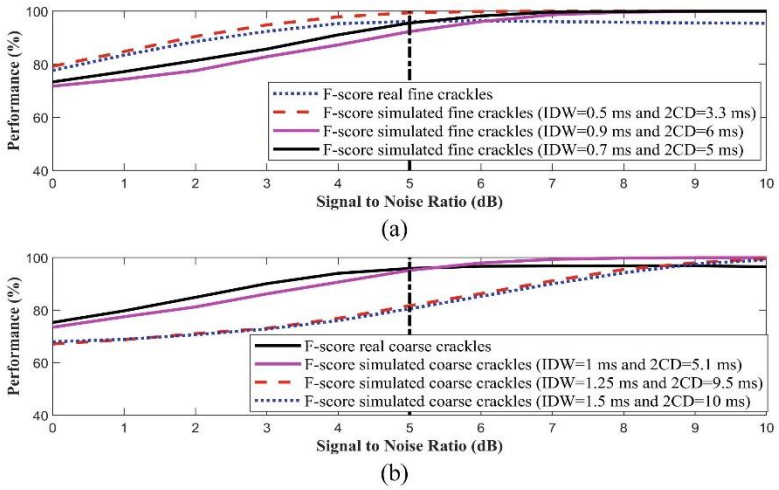


Fig. 6. The HEE algorithm F-score (F_1) plots for breath noise cases with a signal to noise ratio in the range of 0 to 10 dB (Table 1). (a) Real fine crackles and simulated fine crackles cases; (b) Real coarse crackles and simulated coarse crackles cases.

Table 2. Crackle detection performance.

Cases		N_c	SNR	NS	HEE			
					$\overline{SE}(SD)$	$\overline{PPV}(SD)$	$\bar{F}_1(SD)$	$\overline{D}_T(SD)$ s
SFC	A_F	10	5 dB	501	99.4 (2.2)	91.6 (7.3)	95.2 (4.3)	0.72 (0.03)
	H_F	10	5 dB	501	99.9 (0.9)	99.1 (2.8)	99.5 (1.7)	0.75 (0.04)
	C_F	10	5 dB	501	99.7 (1.7)	87.3 (7.9)	92.9 (4.7)	0.79 (0.04)
SCC	A_C	10	5 dB	501	99.7 (1.6)	67.6 (9.2)	80.2 (6.5)	0.78 (0.04)
	H_C	10	5 dB	501	99.8 (1.4)	91.1 (7.2)	95.1 (4.1)	0.72 (0.04)
	C_C	10	5 dB	501	99.9 (1.1)	69.5 (8.6)	81.7 (5.9)	0.77 (0.04)
RFC		10	5 dB	501	93.9 (4.2)	99.0 (3.1)	96.3 (2.7)	0.75 (0.05)
RCC		10	5 dB	501	93.9 (5.9)	98.0 (4.1)	95.8 (3.6)	0.72 (0.02)
RBFC		32	ND	1	71.1	91.4	80.0	0.24
RBCC		6	ND	1	100	100	100	0.25
Overall performance					95.7 (9.0)	89.5 (11.8)	91.7 (7.9)	0.65 (0.21)

SFC: Simulated fine crackles; A_F: IDW = 0.7 ms & 2CD = 5 ms [8]; H_F: IDW = 0.5 ms & 2CD = 3.3 ms [11]; C_F: IDW = 0.9 ms & 2CD = 6 ms [12]; SCC: Simulated coarse crackles; A_C: IDW = 1.5 ms & 2CD = 10 ms [8]; H_C: IDW = 1 ms & 2CD = 5.1 ms [11]; C_C: IDW = 1.25 ms & 2CD = 9.5 ms [12]; RFC: Real fine crackles; RCC: Real coarse crackles; RBFC: Real breath sound with fine crackles; RBCC: Real breath sound with coarse crackles; N_C: Number of crackles; SNR: Signal to noise ratio; NS: Number of test samples; $\overline{SE}$: Mean of sensitivity; $\overline{PPV}$: Mean of positive predictive value; $\bar{F}_1$: Mean of F-score; $\overline{D}_T$: Average detection time; s: Second; ND: Not defined; SD: Standard deviation; In all cases number of samples (N) = 32,768.

The overall performance of the HEE algorithm in terms of SE (95.7%), PPV (89.5%), and F_1 (91.7%) is equivalent to or better than that reported for other detection methods reported in the literature (e.g. [5, 6, 14]). Moreover, the average detection time (D_T) for the algorithm to finish is less than 1 s. The main advantage of the proposed HEE algorithm is its structural simplicity, lower computational cost (detection time) and the ability to perform in low SNRs for both fine and coarse crackles.

Although the favorable results show the potential of the HEE algorithm for automatic crackle detection, this study has some limitations. First, the overall performance of the algorithm is dependent on a suitable choice of the non-adaptive multiplication factor for calculating the threshold, T. Second, though the method works well to determine the location of the start of the crackle, its ability to extract the full duration of each crackle signal has not been examined, so it remains to be determined whether the method can be recommended where characterization of the temporal morphology of crackles is important.

5 Conclusion

In this paper a new HEE algorithm for automatic crackle detection based on crackle time domain features (IDW and 2CD) has been presented. The results suggest that the proposed method is fast ($D_T < 1$ s) and offers high detection performance (SE 95.7%, PPV 89.5%, and F_1 91.7%) even when signal to noise ratio is low.

Future research will focus on evaluating performance on a larger dataset recorded from a range of cardio pulmonary patients with different diagnoses.

Acknowledgment. This research was supported by the NIHR Southampton Biomedical Research Centre, AAIR Charity and the Engineering and Physical Sciences Research Council.

Conflict of Interest. The authors declare no conflict of interest.

References

1. Pramono, R.X.A., Bowyer, S., Rodriguez-Villegas, E.: Automatic adventitious respiratory sound analysis: a systematic review. PLoS ONE **12**(5), e0177926 (2017)
2. Kiyokawa, H., Greenberg, M., Shirota, K., Pasterkamp, H.: Auditory detection of simulated crackles in breath sounds. Chest **119**(6), 1886–1892 (2001)
3. Reyes, B.A., Olvera-Montes, N., Charleston-Villalobos, S., Gonzalez-Camarena, R., Mejia-Avila, M., Aljama-Corrales, T.: A smartphone-based system for automated bedside detection of crackle sounds in diffuse interstitial pneumonia patients. Sensors **18**(11), 3813 (2018)
4. Murphy, R.L.H., Holford, S.K., Knowler, W.C.: Visual lung-sound characterization by time-expanded wave-form analysis. N. Engl. J. Med. **296**(17), 968–971 (1977)
5. Pinho, C., Oliveira, A., Jácome, C., Rodrigues, J., Marques, A.: Automatic crackle detection algorithm based on fractal dimension and box filtering. Procedia Comput. Sci. **64**, 705–712 (2015)

6. Vannuccini, L., Rossi, M., Pasquali, G.: A new method to detect crackles in respiratory sounds. Technol. Health Care **6**(1), 75–79 (1998)
7. Hadjileontiadis, L.J., Rekanos, I.T.: Detection of explosive lung and bowel sounds by means of fractal dimension. IEEE Signal Process. Lett. **10**(10), 311–314 (2003)
8. Charbonneau, G., Ademovic, E., Cheetham, B.M.G., Malmberg, L.P., Vanderschoot, J., Sovijarvi, A.R.A.: Basic techniques for respiratory sound analysis. Eur. Respir. Rev. **10**(77), 625–635 (2000)
9. Sharma, P., Imtiaz, S.A., Rodriguez-Villegas, E.: An algorithm for heart rate extraction from acoustic recordings at the neck. IEEE Trans. Biomed. Eng. **66**(1), 246–256 (2019)
10. Pal, R., Barney, A.: A dataset for systematic testing of crackle separation techniques. In: Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS 2019, Berlin, pp. 4690–4693 (2019)
11. Hoevers, J., Loudon, R.G.: Measuring crackles. Chest **98**(5), 1240–1243 (1990)
12. Cohen, A.: Signal processing methods for upper airway and pulmonary dysfunction diagnosis. IEEE Eng. Med. Biol. Mag. **9**(1), 72–75 (1990)
13. Chinchor, N.: MUC-4 evaluation metrics. In: Proceedings of the 4th Conference on Message Understanding, pp. 22–29 (1992)
14. Quintas, J., Campos G., Marques A.: Multi-algorithm Respiratory Crackle Detection. In: Proceedings of the 6th International Conference on Health Informatics (HEALTHINF), Spain, pp. 239–244 (2013)

A.7 Published journal paper

A.7.1 Iterative envelope mean fractal dimension filter for the separation of crackles from normal breath sounds

Biomedical Signal Processing and Control 66 (2021) 102454

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Iterative envelope mean fractal dimension filter for the separation of crackles from normal breath sounds

Ravi Pal *, Anna Barney

Institute of Sound and Vibration Research (ISVR), University of Southampton, Southampton, United Kingdom

ARTICLE INFO

Keywords:
Crackles
Normal breath sounds
Iterative envelope mean fractal dimension (IEM-FD) filter

ABSTRACT

This paper presents a new method of separating pulmonary crackles from normal breath sounds: the iterative envelope mean fractal dimension (IEM-FD) filter. Crackles are an important physiological parameter for evaluating lung condition of an individual and their automatic separation from normal breath sounds can provide an objective way of diagnosing or monitoring different cardiopulmonary diseases. The filter combines the new iterative envelope mean (IEM) method with the established fractal dimension (FD) technique. The IEM method estimates the non-stationary and stationary parts of the lung sound signal and then the FD technique is applied to the estimated non-stationary output of the IEM method for further refining the separation process. The IEM-FD filter is tested using a publicly available dataset and, compared with an established crackle separation technique. The IEM-FD achieves high accuracy for crackle detection in the presence of noise with SNR >= -1 dB for fine crackles and SNR > +1 dB for coarse crackles, and has low computational cost, with minimal under- or over-estimation and good preservation of crackle morphology. The method is shown to have an overall performance suitable for automated analysis to determine accurately the number and characteristics of pulmonary crackles in a recorded lung sound.

1. Introduction

In this paper, we present a new method for separating pulmonary crackles from normal breath sounds.

During breathing, turbulence in the large airways induces vibrations in the airway walls which are transmitted through the lung tissue and chest wall. These sounds are referred to as normal breath sounds and include bronchial sounds, bronchovesicular sounds and vesicular sounds. Crackles are discontinuous, non-harmonic lung sounds, superimposed on the normal breath sounds, which can be an indication of lung abnormality or pulmonary disease [1]. Crackles are thought to be generated by the sudden opening or closing of airways [2,3]. Normal breath sounds and crackles are both audible through a stethoscope placed against the wall of the thorax.

Crackles are generally described as explosive sounds, being of short time duration, typically 20 ms or less, and having broad frequency content ranging from 100 to 2000 Hz or even higher [4]. They may be characterized into fine (high pitched) and coarse (low pitched) based on their time domain features including their initial deflection width (IDW), two cycle deflection (2CD) and, occasionally, largest deflection width (LDW) [5]. According to the American Thoracic Society, fine crackles have on average IDW ≈ 0.7 ms and 2CD ≈ 5 ms, whereas coarse crackles have on average IDW = 1.5 ms and 2CD = 10 ms [5]. Fine crackles are usually heard in the mid to late inspiratory breath phase and are thought to be generated by the explosive opening of small airways [6]. Coarse crackles are thought to be generated due to air bubbling through sputum in larger airways and can be heard in the early inspiratory and also the expiratory breath phase [7].

Fine crackles may be related to congestive heart failure, lung fibrosis and pneumonia whereas coarse crackles are associated with chronic obstructive pulmonary disease, chronic bronchitis and bronchiectasis. The timing, number and place of generation of pulmonary crackles has been shown to vary with the underlying disease and with its severity. Crackle characteristics can therefore be used in diagnosis and in monitoring of disease progression [8].

Auscultation with a conventional stethoscope placed against the exterior chest wall can be used to identify the presence of crackles, but this approach is subjective and assessment of the number of crackles present and identification of their type (fine or coarse) is highly dependent on clinician hearing ability and expertise. More recently,

* Corresponding author.
E-mail address: r.pal@soton.ac.uk (R. Pal).

https://doi.org/10.1016/j.bspc.2021.102454
Received 10 August 2020; Received in revised form 27 November 2020; Accepted 23 January 2021
Available online 14 February 2021
1746-8094/

electronic stethoscopes have offered the facility to record lung sounds and many methods for automatic processing have been developed [9].

As an initial processing stage, automatic separation of crackles from normal breath sounds can lead to better crackle characterization. On the timescale of a crackle (typically 10–20 ms) normal breath sounds may be considered quasi-stationary and therefore separation into estimates of the non- stationary and stationary elements of a lung sound signal generally sends crackle components to the non stationary signal estimate and the majority of components associated with normal breath sounds to the stationary signal estimate. Separation can reveal not only large amplitude crackles but also small amplitude crackles, which are often significantly masked by the normal breath sounds. Many recent studies have reported development and testing of such separation methods.

Visual Time Expanded Waveform analysis (VTEWA) [10] can be used to identify crackles from the time domain lung sound signal, but this approach is subjective, time consuming and has high inter- observer variability [11], and therefore automated techniques are preferred.

Level slicers and high pass filters (LS & HPFs) can separate crackles from normal breath sounds to some extent, but these methods are insufficient for separating out small amplitude crackles and further, can distort the crackle signal in the process. Katila et al. [12] presented a case study of the effect of high pass filtering on the morphology of crackles. It was observed that both the high pass filter cut off frequency and the filter type influenced the crackle waveform.

Hadjileontiadis et al. [13] proposed the wavelet transform stationary non- stationary (WTST-NST) filter. This method is based on an iterative multiresolution decomposition and multiresolution reconstruction (MRD-MRR) scheme, which separates the stationary and non stationary parts of a signal. The WTST-NST filter can be used to separate crackles from breath sounds on the basis that explosive events in the time domain, such as crackles, have large components at many wavelet scales whereas the components due to relatively stationary signals, such as breath sounds, reduce with increasing wavelet scale. This allows separation, through their amplitude, of wavelet transform coefficients related to crackles from those related to normal breath sounds at each wavelet scale using some threshold value. At each wavelet scale, the threshold value is based on the standard deviation of the wavelet transform coefficients and an empirical multiplication factor. This method can achieve separation with all crackle signals directed to the non-stationary output in most cases, but the morphology of the crackles is not generally well preserved. Further, its computational complexity makes the WTST-NST unsuitable for clinical applications where high processing speed is advantageous [11,14].

The wavelet transform fractal dimension (WT-FD) filter [15,16] has comparable outcomes to the WTST NST filter for locating the crackle peaks, but is somewhat better at preserving the morphology of the extracted crackles [16]. In applying the WT-FD filter, firstly, the input signal is decomposed into approximation and detail coefficient vectors using the WT and then the FD technique [17] is applied to the approximation and detail coefficient vectors to separate the WT coefficients related to crackles from those related to normal breath sounds. Although this method separates both fine and coarse crackles into the non stationary output with high sensitivity [16], the choice of base wavelet and number of WT decomposition scales is critical to its success in separating the crackles [18]. A need to optimize these parameters for any given data set prior to use in a clinical setting would be less than ideal.

More recently, Hadjileontiadis [18] proposed the empirical mode decomposition fractal dimension (EMD-FD) filter. In this method, firstly the input signal is adaptively decomposed into multiple intrinsic mode functions (IMFs) and a residual component using the EMD technique [19]. Next, an energy-based threshold is used for selecting the number of IMFs containing significant elements of crackle sound and then, on those selected IMFs, the FD technique [17] is applied to further refine the separation to leave just the crackles. EMD-FD was shown to perform comparably to WT-FD, but with the advantage that, since, unlike WT-FD, EMD has no underlying assumption of orthogonality, it can be applied to both linear and non-linear data. Further, as a data driven process the number of *a priori* parameter choices is minimized whereas for the WT-FD filter the choice of wavelet and number of scales must be made in advance of processing. In addition, EMD-FD was shown to be robust against the presence of extraneous environmental noise. However, the selection of the number of independent mode functions containing crackle information requires the setting of an empirical threshold. Use of too many IMFs may add components of normal breath sound to the separated signal (over estimation) whereas too few may result in crackles missing from the separated signal (under estimation) or in distorted crackle morphology. We note also that EMD is, in general, a rather slower analysis method that WT due to its relatively high computational complexity.

Other automated methods such as the stationary non-stationary (ST-NST) filter [20], mST-NST filter [21], fuzzy stationary non-stationary (FST-NST) filter [22], generalized fuzzy stationary non-stationary (GFST-NST)) filter [11], orthogonal least square fuzzy (OLSF) filter [23], the wavelet packet transform (WPT) filter [24], neurofuzzy filter [14], and Independent component analysis (ICA) [25] have also all been proposed for automatic crackle separation. Although, these methods can provide crackle separation with low computational cost, they do so with reduced sensitivity or lower quality in the reconstruction of the crackle morphology. A summary of the strengths and weaknesses of each method is given in Table 1.

Latterly, separation has received less emphasis as machine learning systems have been tested for classification of lung sounds without this pre-processing step (e.g. [26]). However, such systems might have their ability to classify accurately enhanced by using separation as a pre-processing step.

Auscultation of pulmonary crackles has been used in clinical assessment of patients for diagnosis and for monitoring of disease progression for many years. For diagnosis, an initial consideration may be the presence or absence of an unusual number of crackles [27,28]. This assessment may be confounded by the increased prevalence of crackles with age in the healthy population [29]. Further, low amplitude crackles may frequently be masked by the breath noise leading to an underestimate of the true number. Thus, an accurate assessment of the number of crackles present can assist with the diagnostic decision. In a clinical environment, background noise may also be a problem in aural assessment of crackles [30–32] so a system robust to added noise is also

Table 1
Technical characteristics for good crackle separation compared for different published separation methods.

Methods	ACC	UOE	POC	CCX	NRB	OBJ
VTEWA [10]	+	x	x	–	–	–
LS & HPFs [12]				+	x	+
ST-NST [20]	-	-	-	x	x	–
mST-NST [21]	+	x	x	x	+	–
WTST-NST [13]	+	+	-	-	+	-
WT-FD [15,16]	++	+	+	+	+	-
EMD-FD [18]	+	+	+	–	+	–
FST-NST [22]	+	+	-	+	x	-
GFST-NST [11]	+	+	-	+	x	-
OLSF [23]	+	+	-	+	x	-
Neurofuzzy filter [14]	+	x	x	+	x	-
WPT [24]	+	x	x	+	+	-
ICA [25]	+	x	x	x	x	-

ACC: Accuracy (number of crackles correctly separated); UOE: Under-, over estimation; POC: Preservation of crackle morphology; CCX: Analysis speed (computational complexity); NRB: Robustness to additive/environmental noise; OBJ: Objectivity (need to set hard thresholds and/or make decisions about process based on the data and/or requirement of training phase for estimation of the optimum model parameters); ++ = strong attribute; + = acceptable attribute; – weak attribute; – – very weak attribute; x – attribute not reported.

preferred. New methods may be more readily adopted clinically when they adapt traditional auscultation methods [33], especially if they can be carried out at the same time as traditional auscultation or with limited additional time commitment [34]. Differential diagnosis between lung conditions may be additionally facilitated by classification of the crackles into fine and coarse or by inspection of the crackle morphology in the recorded sound signal [35,36]. This requires an analysis process which gives a representation of the crackle with minimal distortion to allow its temporal characteristics to be accurately measured. For monitoring, number, classification as coarse or fine and crackle morphology are again critical parameters [37,38]. An optimal process for crackle separation should therefore have high accuracy for the number of crackles detected, high robustness to noise, low computational complexity (high processing speed), separation without significant under- or over-estimation of the crackle waveform and the ability to preserve crackle morphology after separation.

Respiratory monitoring remains a current clinical interest [39,40]. Building on a concept initially developed by Murphy et al. [41], many recent studies (e.g [42].) have considered multi- channel recordings and analysis to identify the presence of abnormal lung sounds. Garcia et al. [25] have recently used Independent Component Analysis (ICA) on a 25 channel recording system to separate crackles from normal breath sounds. However, the specialist equipment required for a multichannel approach differs substantially from the traditional stethoscope, ubiquitous in a clinical setting, leading to poor uptake of such devices in practice. The simplicity of processes that align with well-established auscultation techniques are therefore to be preferred.

In this paper, we present a new automatic single channel crackle separation technique known as the iterative envelope mean fractal dimension (IEM-FD) filter. The iterative envelope mean (IEM) method divides a signal into initial estimates of its stationary and non-stationary parts. The FD technique [17], applied to the non- stationary signal estimate, removes further elements related to normal breath sounds to refine the separation. To evaluate our new method, we compare it with, the WT-FD filter [15,16]. The WT-FD filter was used for comparison due to its high accuracy for the number of crackles detected, shown in [16] to be 100 % in both fine and coarse crackles, and good ability to meet the other advantageous criteria.

The rest of the paper is arranged as follows. Section 2 describes the IEM-FD filter. Section 3 presents the test dataset and the quantitative evaluators used to assess the outcome and Section 4 details the parameters selected for the separation filters. The experimental results are presented in Section 5 where a comparison is made with WT-FD performance on the same dataset. Section 6 presents the discussion of results and Section 7 the conclusions.

2. Iterative envelope mean fractal dimension filter

The IEM-FD filter combines two techniques: the new IEM method and the more established FD technique [17]. The IEM method estimates the stationary and non-stationary parts of the lung sound signal and, the FD technique is then applied to the non-stationary output of the IEM method to refine the separation process further.

2.1. Iterative envelope mean method

The lung sound signal is a combination of normal breath sounds and any added sounds, such as crackles, and is typically recorded over between three and twenty breath cycles depending on respiratory rate and recording duration. The IEM method subtracts the envelope mean value of the smoothed lung sound signal from the original lung sound signal. The envelope mean signal is the mean of the upper and lower envelopes of the smoothed lung sound signal. The resulting signal can then be used as the input for a subsequent iteration. After a number of iterations, Q, the IEM method will provide an estimate of the non-stationary part of the lung sound signal and also, through the summation of the envelope means from each iteration, an estimate of its stationary part.

In detail, the IEM method proceeds as follows: In the first step, a smoothed version of the lung sound signal is generated and its first and second derivatives are calculated using a filter from the Savizky Golay (SG) family. The SG filter parameters are selected according to the guidelines proposed by Vannuccini et al. [43] with degree of fitting polynomial $p_f = 4$ and number of coefficients n_c equal to approximately one to two times the half width of the shortest duration feature of interest in the signal. In the case of crackles, the first cycle of the crackle is the shortest cycle and generally has a duration of less than 2 ms. In our data, where the sampling frequency of the lung sound signal is 44,100 Hz, the half width is less than 88 samples. The SG filter parameters used here are therefore $p_f = 4$, $n_c = 89$ and order of derivation $(d_o) = 0$, 1 and 2 for smoothing the lung sound signal, and for estimating first and second derivative of that smoothed signal, respectively.

Next all the local extrema of the first derivative ($y'_s(n)$) are identified and classified as maxima or minima using sign changes over the second derivative ($y''_s(n)$) of the smoothed lung sound signal ($y_s(n)$).

The coordinates of the smoothed lung sound signal at the location of each of the first derivative local maxima and minima are then calculated. A cubic spline interpolation is used to connect the maxima in the smoothed lung sound signal to define the upper envelope ($UP_{env}(n)$), and correspondingly, the local minima are connected with each other to extract the lower envelope ($LW_{env}(n)$). The envelope mean value is then calculated using the estimated upper and lower envelopes of the smoothed lung sound signal:

$$m_q(n) = \frac{UP_{env_q}(n) + LW_{env_q}(n)}{2} \tag{1}$$

where n is the sample index in the input signal i.e. $n = 1, 2, \ldots N$ and q is the iteration number where $q = 1, 2, \ldots, Q$. The envelope mean value is then subtracted from the lung sound signal to get an estimate of the non stationary signal $R_q(n)$:

$$R_q(n) = y_q(n) - m_q(n) \tag{2}$$

where $y_q(n)$ is the lung sound signal at iteration q. Note that for each iteration, q, the envelope mean value is calculated using the smoothed lung sound signal ($y_{s_q}(n)$), and its derivatives, which is then subtracted from the un- smoothed lung sound signal $y_q(n)$.

To end the iterative process, a stopping criterion $STC1_q$ is calculated:

$$STC1_q = \left| E\left\{R^2_{q-1}(n)\right\} - E\left\{R^2_q(n)\right\} \right| \tag{3}$$

where $E\{.\}$ is the expected value and has an initial value of $R_{q-1} = 0$.

The stopping criterion ($STC1_q$) is compared with accuracy level $\beta 1$: $(1 > \beta 1 > 0)$. In this study the value of $\beta 1$ is empirically set equal to 0.01.

If $STC1_q \geq \beta 1$, a new input lung sound signal $y_{q+1}(n) = R_q(n)$ is defined and the process is repeated (usually one or two iterations are sufficient (see Table 3). Note that for the IEM method the stopping criterion is the same as that defined in [13,15,17,18].

When the stopping criterion is met, the estimates of the non-stationary and stationary parts of the lung sound signal are calculated using Eqs. (4) and (5) , respectively where Q is the total number of iterations made:

$$NSTS(n) = R_Q(n) \tag{4}$$

$$STS(n) = \sum_{q=1}^{Q} m_q(n) \tag{5}$$

Note that the first derivative local maxima and minima locations on the smoothed lung sound signal are used for estimating the upper and lower envelopes rather than using local maxima and minima points directly from the smoothed lung sound signal. If the upper and lower

envelopes, using local maxima and minima of the smoothed lung sound signal itself are used, an inefficiency arises. The upper and lower envelopes can have large separation for regions with infrequently occurring extrema points (low frequency variation), which may require a large number of iterations for separating the lung sound signal. Using instead the upper and lower envelopes derived from the local maxima and minima of the first derivative of the smoothed lung sound signal reduces this inefficiency and consequently the number of iterations needed.

As an example, Fig. 1 shows a section of duration 0.075 s of a lung sound signal recorded from a patient with idiopathic pulmonary fibrosis (Table 2, Case RBFC) where the location of the crackles has been audio-visually identified by an experienced pulmonary acoustics researcher and marked with arrowheads. Fig. 1(a) displays the non-stationary output of the IEM process after the 1st iteration, the upper and lower envelopes and the envelope mean value where upper and lower envelopes are estimated using directly the smoothed lung sound signal extrema points. It can be observed that between 0.055 s and 0.075 s, the separation between the upper and lower envelopes is large. As a result, when the envelope mean value is subtracted from the lung sound signal that region changes its shape but the non-stationary output not only contains the crackles but also consists of a large portion of normal breath sound after the first iteration. On the other hand, in Fig. 1(b) where upper and lower envelopes are estimated using the first derivative local maxima and minima locations on the smoothed lung sound signal, we observe that the envelope mean is a closer fit to the lung sound signal and when it is subtracted, the contribution of normal breath sound to the non-stationary estimate is very much less. Note that the lung sound signal is smoothed using the SG filter prior to calculating the upper and lower envelopes to remove the impulsive spikes without affecting the crackle waveform.

Huang et al. [19] proposed the EMD technique for adaptively

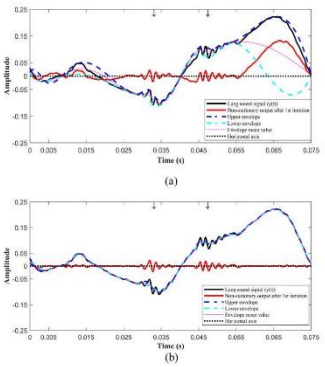


Fig. 1. Illustration of the iterative envelope mean method applied to a section of 0.075 s of lung sound data recorded from a patient with idiopathic pulmonary fibrosis; (a) estimation of the upper, lower and mean envelopes and the non-stationary signal estimate after one iteration using extrema points of the smoothed lung sound signal; (b) estimation of the upper, lower and mean envelopes and the non-stationary signal estimate after one iteration using extrema locations of the first derivative of the smoothed lung sound signal.

Table 2
Summary of the test dataset.

Cases	N_C	Diagnosis	IDW& 2CD (ms)	BN	SNR
SFC	10	NA	0.7&5 [5]	BR_N	−10 to 10 dB in steps of 1 dB
			0.5&3.3 [45]	BR_N	
			0.9&6 [46]	BR_N	
SCC	10	NA	1.5&10 [5]	BR_N	
			1&5.1 [45]	BR_N	
			1.2&8.8 [46]	BR_N	
RFC	10	IPF	ND	BR_N	
RCC	10	B_L	ND	BR_N	
RBFC	ND	IPF	ND	NBS	ND
RBCC	ND	B_L	ND	NBS	ND

SFC: Simulated fine crackles; RFC: Real fine crackles; SCC: Simulated coarse crackles; RCC: Real coarse crackles; RBFC: Real breath sound with fine crackles; RBCC: Real breath sound with coarse crackles; N_C: Number of crackles; ND: Not defined; NA = Not applicable; IPF: Idiopathic pulmonary fibrosis; B_L: Bronchiectasis; BN: Background noise; BR_N: Breath noise; NBS: Normal breath sound.

decomposing a signal into its IMFs and a residual component in decreasing order of frequency. Although the IEM method is superficially similar to the EMD method in that processing of each iteration begins with an estimation of the upper and lower envelopes of a signal based on the local extrema, which is then subtracted from the input signal, there are distinct differences between the two methods. In particular, the EMD method continues each iteration until the output meets the strict criteria that define an IMF [19] whereas the IEM method ceases when the stopping criterion (3) reduces to the required precision.

2.2. Fractal dimension technique

The IEM method makes a good estimate of the stationary and non-stationary parts of the lung sound signal as can be seen from Fig. 2(b), however it is not usually sufficient by itself to ensure optimal separation. To minimise the remaining elements of normal breath sounds in the non-stationary signal estimate ($NSTS(n)$) the FD technique is applied following the same steps as described [17]. For our study the sampling frequency of the non-stationary signal estimate is $f_s = 44{,}100$ Hz and using a multiplication factor $m_f = 0.006$ following [17], the length of the fractal dimension window $W_{FD} = 264$ samples.

2.3. Iterative envelope mean fractal dimension filter

The IEM FD filter is implemented using the IEM method followed by the FD technique. Two iteration loops are used in the IEM-FD filter: a loop related to IEM method $q = 1, 2, \ldots, Q$, and a loop for the combination of the IEM- and FD processes $k = 1, 2, \ldots, K$.

The IEM-FD filter working process is shown in Fig. 2. First the IEM method estimates the non-stationary ($NSTS(n)$) (see Fig. 2(b)) and stationary ($STS(n)$) parts of the lung sound signal. Next point to point FD values of the estimated non-stationary output are calculated and the fractal dimension peak peeling ($FDPP$) algorithm [17] is applied to automatically detect those peaks of the estimated $FD(n)$ which may correspond to crackles (Fig. 2(c)). Then, using the estimated FDPP sequence, two binary thresholds are calculated: the non-stationary binary threshold:

$$NBTH^k(n) = \left\{ \begin{array}{l} 1 \; if \; FDPP^k(n) \neq 1 \\ 0 \; if \; FDPP^k(n) = 1 \end{array} \right\} \tag{6}$$

and the stationary binary threshold:

$$SBTH^k(n) = \left[1 - NBTH^k(n)\right] \tag{7}$$

as displayed in Fig. 2(d) and (e), respectively.

Table 3
Evaluation of separation of IEM-FD filter in the case of fine and coarse crackles.

Cases		BN	D_g	SNR (dB)	N_R	NOTS	IEM-FD K (min-max)	IEM-FD Q (min-max)	IEM-FD N_E (min-max)	IEM-FD $\bar{D}_R(SD)$ (%)	WT-FD K (min-max)	WT-FD N_E (min-max)	WT-FD $\bar{D}_R(SD)$ (%)
	A_F	BR_N	NA	−1	10	501	1–1	2–2	10–10	100 (0)	1–1	10–10	100 (0)
SFC	H_F	BR_N	NA	−1	10	501	1–1	1–2	10–10	100 (0)	1–1	10–10	100 (0)
	C_F	BR_N	NA	−1	10	501	1–1	2–2	10–10	100 (0)	1–1	10–10	100 (0)
RFC		BR_N	IPF	−1	10	501	1–1	1–1	10–10	100 (0)	1–1	9–10	99.48 (2.22)
RBFC		NBS	IPF	ND	32	1	1	1	21	65.63	1	20	62.50
	A_C	BR_N	NA	1	10	501	1–1	2–2	10–10	100 (0)	1–1	10–10	100 (0)
SCC	H_C	BR_N	NA	1	10	501	1–1	2–2	10–10	100 (0)	1–1	10–10	100 (0)
	C_C	BR_N	NA	1	10	501	1–1	2–2	10–10	100 (0)	1–1	10–10	100 (0)
RCC		BR_N	B_r	1	10	501	1–1	1–1	9–10	99.18 (2.74)	1–1	8–10	96.73 (5.18)
RBCC		NBS	B_r	ND	6	1	1	1	6	100	1	6	100
Total Performance (TD_R^{XX}, $SD_{TD_R}^{XX}$)		Fine crackles (SFC, RFC and RBFC)					$TD_R^{FC} = 99.98$ %, $SD_{TD_R}^{FC} = 0.77$ %				$TD_R^{FC} = 99.85$ %, $SD_{TD_R}^{FC} = 1.43$ %		
		Coarse crackles (SCC, RCC and RBCC)					$TD_R^{CC} = 99.80$%, $SD_{TD_R}^{CC} = 1.42$ %				$TD_R^{CC} = 99.18$ %, $SD_{TD_R}^{CC} = 2.93$ %		

SFC: Simulated fine crackles; A_F: IDW = 0.7 ms & 2CD = 5 ms [5]; H_F: IDW = 0.5 ms & 2CD = 3.3 ms [45]; C_F: IDW = 0.9 ms & 2CD = 6 ms [46]; RFC: Real fine crackles; SCC: Simulated coarse crackles; A_C: IDW = 1.5 ms & 2CD = 10 ms [5]; H_C: IDW = 1 ms & 2CD = 5.1 ms [45]; C_C: IDW = 1.25 ms & 2CD = 9.5 ms [46]; RCC: Real coarse crackles; RBFC: Real breath sound with fine crackles; RBCC: Real breath sound with coarse crackles; BN: Background noise; BR_N: Breath noise; NBS: Normal breath sound; IPF: Idiopathic pulmonary fibrosis; B_r: Bronchiectasis; D_g: Diagnosis; SNR: Signal to noise ratio; NOTS: number of test samples; D_R :Mean of Rate of Detectability; *SD*: Standard deviation; NA: Not applicable; ND: Not defined; N_R: Real number of crackles; N_E: Separated crackles; TD_R^{XX}: Total Performance; $SD_{TD_R}^{XX}$: Standard deviation; *XX* stands *FC* for fine crackles and *CC* for coarse crackles; *K* and *Q*: number of iterations; min: Minimum value; max: Maximum value; In all cases number of samples (N) = *32,768*.

The non-stationary output of the IEM method is multiplied by non-stationary binary threshold $NBTH(n)$ to get the refined non stationary estimate $NST(n)$ and the non stationary estimate of the IEM method is multiplied by $SBTH(n)$ to obtain the remaining normal breath sound signal $SSR(n)$ from the $NSTS(n)$:

The summation of the STS (5) and the SSR gives the estimate of the stationary output, $SSF^k(n)$ of the IEM FD filter at iteration k.

To end the IEM-FD filter, a stopping criterion based on the stationary output can be calculated and compared with accuracy level ($\beta 2$).

The stopping criterion:

$$STC2^k = \left| E\left\{ \left(SSF^{k-1}\right)^2(n) \right\} - E\left\{ \left(SSF^k\right)^2(n) \right\} \right| \tag{8}$$

where $E\{\}$ is the expected value and the initial value of $SSF^{k-1} = 0$. $STC2^k$ is compared with accuracy level $\beta 2$, where, $1 > \beta 2 > 0$. If $STC2^k \geq \beta 2$, input signal $y^{k+1}(n) = SSF^k(n)$, otherwise $k = K$ and the iterative loop ends. Here the value of the $\beta 2$ is empirically set equal to 0.1 and K represents the maximum iteration level. In the final step, the non-stationary and stationary parts of the signal are calculated when $k = K$, using:

$$NS(n) = \sum_{k=1}^{K} NST_k(n) \tag{9}$$

$$ST(n) = SSF_{(at\ k=K)}(n) \tag{10}$$

The non-stationary and stationary outputs of the IEM-FD method are shown in Fig. 2(f) and (g) respectively.

3. Analysis

In this section, the test data set is described and quantitative measures for evaluating how well each algorithm separates the crackles from the breath sounds are discussed.

3.1. Dataset and test samples

A previously published test dataset (Table 2) [44] is used for evaluating the crackle separation. The dataset consists of: simulated fine (SFC) and coarse (SCC) crackles for which the IDW and 2CD may be selected; real fine (RFC) and coarse (RCC) crackles with a variety of IDW and 2CD values extracted from recorded lung sound signals; noise simulated to have the spectral characteristics of breath noise (referred to hereafter as 'breath noise') (BR_N) two real lung sound signals, one with predominantly fine crackles (RBFC) from a patient with idiopathic pulmonary fibrosis (IPF) and one with mostly coarse crackles (RBCC) from a patient with Bronchiectasis (Br), both recorded with an electronic stethoscope. Further details of the dataset may be found in [44] and the data used here is summarized in Table 2.

To explore the robustness of the separation process to noise, test samples were generated by burying 10 simulated or 10 real crackles within a simulated breath noise sample (BR_N). Average SNR was varied from -10 to 10 dB in steps of 1 dB.

The local SNR for any given crackle in a test signal varies randomly, which may affect the separation, therefore for each set of crackles and each average SNR, 501 test samples were generated each with its own sample of BR_N. The use of 501 test samples at each SNR is justified in section 4.1. All test signals are sampled at 44,100 Hz.

3.2. Quantitative evaluators

Successful separation must meet three criteria: extracting all the embedded crackles, minimizing the inclusion of non-crackle components and preserving crackle morphology after separation. We refer to the failure to extract all crackles or loss of some portion of the crackle in the output signal as under estimation, and the inclusion of non crackle components as over-estimation.

The separation of the IEM-FD was evaluated against reference test signals and against the separation of the WT-FD filter [15,16], chosen for its excellent accuracy in separating both fine and coarse crackles [16]. For the synthesized test signals, the time series of the crackles in the absence of breath noise was used as a reference signal. In the test samples measured in patients (RBFC and RBCC), an experienced pulmonary acoustics researcher had previously marked the location of each crackle.

Separation was evaluated using several different metrics: To measure the similarity between the estimated non stationary output of the IEM FD separation process and the crackle reference signal, the cross-correlation index (CCI) was used; the outcome of crackle separation

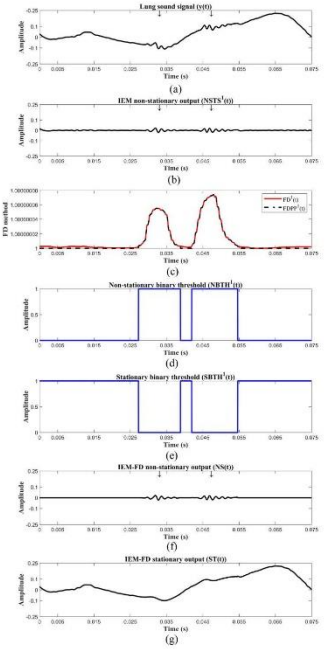


Fig. 2. A worked example of the proposed IEM FD filter: (a) A time section of 0.075 s lung sound data recorded from a patient with idiopathic pulmonary fibrosis ($y(t)$), where location of the crackles is marked with arrowheads; (b) Non-stationary output of the IEM method ($NSTS^1(t)$); (c) The FD of the IEM method non-stationary output ($FD^1(t)$) and the FDPP algorithm for estimating FD valid peaks ($FDPP^1(t)$); (d) The non-stationary binary threshold ($NBTH^1(t)$); (e) The stationary binary threshold ($SBTH^1(t)$); (f) The non-stationary output of the IEM-FD filter ($NS(t)$); (g) The stationary output of the IEM-FD filter ($ST(t)$).

was assessed by two quantitative evaluators proposed by Hadjileontiadis et al. [13]: Rate of Detectability (D_R), and Total Performance (TD_R); To quantify over- or under-estimation in the separation, the Quality Factors (QFs) proposed by Hadjileontiadis et al. [13] were adapted to benefit from the existence in our test data of the reference signals; To evaluate the ability of the separation process to preserve crackle morphology the 2CD percentage error (PE_{2CD}) was calculated.

The process for calculating each metric is as follows:

3.2.1. Cross correlation index (CCI)

The cross-correlation index (CCI) indicates the ability of the IEM-FD and the WT-FD filters to separate all the crackles into the non-stationary signal estimate for a given SNR. The CCI was calculated using Pearson's correlation coefficient:

$$CCI = \frac{\sum_{n=1}^{N}\left(R_C(n)-\overline{R}_C\right)\left(NS(n)-\overline{NS}\right)}{\sqrt{\sum_{n=1}^{N}\left(R_C(n)-\overline{R}_C\right)^2}\sqrt{\sum_{n=1}^{N}\left(NS(n)-\overline{NS}\right)^2}} \times 100\% \quad (11)$$

where $R_C(n)$ is the crackle reference signal, $NS(n)$ is the non stationary output of the separation method, n is the sample index (n=1, 2…N), and $\overline{R}_C$ and $\overline{NS}$ are the average values of the crackle reference signal and non stationary output, respectively.

3.2.2. Rate of detectability

The Rate of Detectability (D_R) measures the ability of the IEM-FD and WT-FD filters to separate the correct number of crackles at the correct locations into their non stationary outputs. D_R was calculated using (12).

$$D_R = \frac{N_E}{N_R} \times 100\ \% \quad (12)$$

where N_E is the number of crackles in the non-stationary output of a separation algorithm and N_R is the number of crackles in the input signal. For test signals: SFC, RFC, SCC and RCC, where a reference signal exists, each the non stationary output was correlated with the reference signal and only those crackles with CCI $\geq$ 0.5, where the cut-off CCI value was empirically selected, were counted as contributing to N_E. For RBFC and RBCC where crackle reference signals do not exist, the number of crackles in the output was counted manually by comparing their location with the marked crackles in the input signal. D_R was calculated for each test signal. The mean value and standard deviation over all tests are reported.

3.2.3. Total performance (TD_R)

The Total Performance (TD_R) also measures accuracy in terms of number and temporal location of crackles in the non-stationary output by the IEM FD and the WT FD filters. TD_R is the D_R calculated separately for all fine and for all coarse crackles (TD_R^{FC} and TD_R^{CC}). As for D_R the mean and standard deviation over all test signals (SFC, RFC and RBFC for fine crackles and SCC, RCC and RBCC for coarse crackles) is reported.

3.2.4. Quality factors (QFs)

Quality Factors measure over- and under-estimation in the non-stationary output signal. Building on [13,16] but noting that our test data set provides us with reference signals, we define four Quality Factors: a Reference Quality Factor (R_{QF_U}) for under-estimation, an Estimated Quality Factor (E_{QF_U}) for under-estimation, a Reference Quality Factor (R_{QF_O}) for over-estimation and an Estimated Quality Factor (E_{QF_O}) for over estimation. To calculate the QFs, firstly two thresholds are defined:

$$TH_1(n) = \left\{ \begin{matrix} 1 \text{ if } R_C(n) \neq 0 \\ 0 \text{ if } R_C(n) - 0 \end{matrix} \right\} \quad (13)$$

$$TH_2(n) = [1 - TH_1(n)] \quad (14)$$

where $R_C(n)$ is the crackle reference signal and n is the sample index with $n = 1, 2, \ldots, N$. Secondly, the threshold $TH_2(n)$ is multiplied by the input signal $y(n)$, to calculate a background noise reference signal $R_{BN}(n)$.

Thirdly, the non-stationary output of the chosen separation filter $NS(n)$ is divided into two parts: non-stationary signal with only crackle portion ($NS_C(n)$) and remaining non-stationary part ($NS_R(n)$), according

to (15) and (16).

$$NS_C(n) = NS(n)\ TH_1(n) \tag{15}$$

$$NS_R(n) = NS(n)\ TH_2(n) \tag{16}$$

Next QFs for under estimation are calculated using the area under the input signal $y(n)$, area under the crackle reference signal $R_C(n)$ and area under the crackle portion of the non-stationary signal $NS_C(n)$. Similarly, the two Quality Factors for over-estimation are evaluated using the area under the input signal $y(n)$, area under the background noise reference signal $R_{BN}(n)$ and area under the remaining non-stationary part $NS_R(n)$.

$$R_{QF_U} = \frac{\sum_{i=1}^{N}|y(n)|\Delta n - \sum_{i}^{N}|R_C(n)|\Delta n}{\sum_{1}^{N}|y(n)|\Delta n} \tag{17}$$

$$E_{QF_U} = \frac{\sum_{i=1}^{N}|y(n)|\Delta n - \sum_{i}^{N}|NS_C(n)|\Delta n}{\sum_{1}^{N}|y(n)|\Delta n} \tag{18}$$

$$R_{QF_O} = \frac{\sum_{i=1}^{N}|y(n)|\Delta n - \sum_{i}^{N}|R_{BN}(n)|\Delta n}{\sum_{1}^{N}|y(n)|\Delta n} \tag{19}$$

$$E_{QF_O} = \frac{\sum_{1}^{N}|y(n)|\Delta n - \sum_{i}^{N}|NS_R(n)|\Delta n}{\sum_{1}^{N}|y(n)|\Delta n} \tag{20}$$

where $||$ represents the absolute value and Δn is the sample period in seconds. The maximum value of area under $|NS_C(n)|$ was set equal to the area under $|R_C(n)|$. E_{QF_U} in the range $R_{QF_U} < E_{QF_U} \leq 1$ represents under-estimation and a value close to 1 indicates high under estimation. Similarly E_{QF_O} in the range $R_{QF_O} \leq E_{QF_O} < 1$ represents over-estimation and a value close to R_{QF_O} shows high over-estimation. Hence, a value of E_{QF_U} close to R_{QF_U} with a value of E_{QF_O} near to 1 represents crackle separation without either high under- or high over estimation.

3.2.5. 2CD percentage error (PE_{2CD})

The 2CD percentage error assesses the ability of an algorithm to preserve crackle morphology after separation from background noise. The percentage error in 2CD following separation is calculated using:

$$PE_{2CD} = \left|\frac{AC_{2CD} - EC_{2CD}}{AC_{2CD}}\right| \times 100\ \% \tag{21}$$

where $||$ represents the absolute value, AC_{2CD} is the actual crackle 2CD calculated from the crackle reference signal $R_C(n)$ and EC_{2CD} is the estimated crackle 2CD calculated from the non stationary filter output ($NS(n)$). The 2CD was calculated using the first five zero crossings of the crackle.

For the RBFC and the RBCC signals, where a crackle reference signal does not exist, the crackle separation of the two filters was not evaluated using the QFs or the PE_{2CD}.

4. Filter parameters

The parameters used for the new IEM-FD filter and the previously published WT-FD filter [15,16] are shown in Table 4.

4.1. Selection of number of test samples

As mentioned in Section 3.1, 501 test samples are generated for each simulated test signal to account for the effect of random variation of the local SNR around any given crackle.

For each SNR, the same 10 crackles embedded in 501 unique noise signal samples are passed through each separation filter and the resulting CCIs are averaged to get one CCI value for each SNR point for each filter.

Fig. 3 shows the average CCI for the IEM-FD filter for real fine crackles (RFC) at an SNR of 1 dB and real coarse crackles (RCC) at an SNR of 1 dB when the number of test samples is increased from 1 to 2001 in steps of 1.

The choice of SNR for these plots is discussed in section 5. We note that for more than 501 samples, the CCI is approximately constant in both cases. The selected number of 501 test samples is marked on Fig. 3.

Table 4
Parameters used for different separating methods.

Parameters		IEM-FD	WT-FD
Number of samples (N)		32,768	32,768
Number of decomposition levels (M)		NA	$M = 1$ [16]
Type of wavelet		NA	*DB4* [16]
Sampling frequency (f_s)		44,100 Hz	44,100 Hz
Accuracy level (β)	β_1	0.01	NA
	β_2	0.1	0.1
	β_3	0.01 [17]	0.01 [16]
Multiplication factor		$m_f = 0.006$ [17]	$m_f = 0.006$ [16]

NA: not applicable; *DB4*: Daubechies quadrature mirror filters (QMFs) of eight coefficients; M: Number of WT decomposition levels; m_f : Multiplication factor; β_1: Accuracy level for the iterative envelope mean method; β_2: Accuracy level for desired stationary and non-stationary outputs; β_3: Accuracy level for the fractal dimension peak peeling algorithm.

5. Experimental results

This section presents the results obtained using the IEM-FD filter and provides systematic comparison with the WT-FD filter [15,16]. Both the separation techniques are implemented using the Matlab (R2019a) programming language.

5.1. Separation by the IEM-FD filter

Fig. 4 shows plots of CCI averaged over all 501 test signals against SNR using i) the IEM-FD filter and ii) the WT-FD filter for the separation. Plots labelled (a) show curves for simulated fine crackles with three different values of IDW/2CD and real fine crackles; plots labelled (b) show curves for simulated coarse crackles with three different values of IDW/2CD, and real coarse crackles.

Taking CCI $> = 0.8$ to indicate strong correlation between the separated signal and the test signal, strong correlation occurs for all fine crackle test signals with SNR greater than 1 dB except for SFC with 2CD = 6 ms for both IEM-FD and WT-FD. For SFC with 2CD = 6 ms the CCI is just below 0.8 at SNR = - 1 dB but is above at SNR = 0 dB. For coarse crackles strong correlation occurs for SNR $> = 1$ dB. The plots therefore

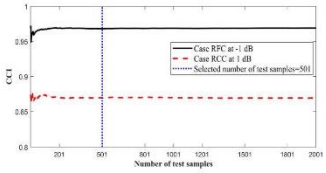


Fig. 3. Selection of number of test samples to eliminate random variation of the IEM-FD filter crackle separation due to variations in local SNR using real fine crackles (RFC) case at 1 dB and real coarse crackles (RCC) case at 1 dB.

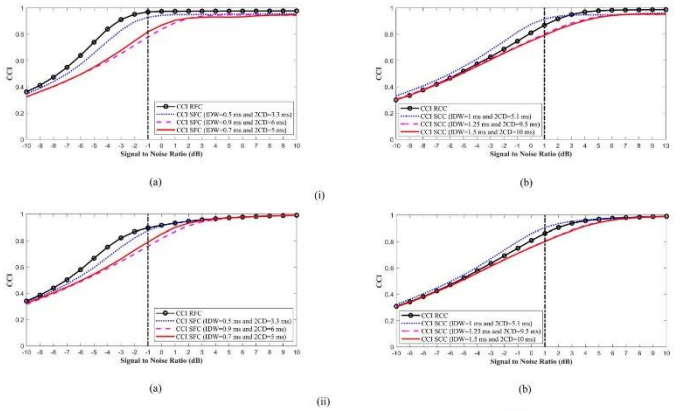


Fig. 4. Cross correlation index (CCI) plots for breath noise cases with a signal to noise ratio in the range of -10 to 10 dB (Table 2). (i) The IEM-FD filter, (a) RFC (real fine crackles) and SFC (simulated fine crackles) cases; (b) RCC (real coarse crackles) and SCC (simulated coarse crackles) cases. (ii) The WT-FD filter, (a) RFC (real fine crackles) and SFC (simulated fine crackles) cases; (b) RCC (real coarse crackles) and SCC (simulated coarse crackles) cases.

suggest two threshold SNRs above which good separation can be achieved by both filters: SNR $= -1$ dB for fine crackles and SNR $= 1$ dB for coarse crackles.

Comparative evaluation of the IEM-FD filter and the WT-FD filter [15,16] was made using the synthesized signals at these threshold SNR values supplemented by the real breath sound with fine crackles (RBFC) and the real breath sound with coarse crackles (RBCC) using the quantitative evaluators described in section 3.2.2 to 3.2.5.

Separation by the IEM-FD filter into non-stationary and stationary estimates for test signals using two randomly selected, real crackle cases at the threshold SNRs, one for fine and one for coarse crackles (Table 3, cases RFC and RCC) are shown in Fig. 5 i (a–c) and ii (a–c), respectively. The location of the crackles before inserting into the background noise is marked with arrowheads. Fig. 5 (a) displays the input signals. The non-stationary and stationary signal estimates after applying the IEM-FD filter are shown in Fig. 5 (b) and (c), respectively. Comparing these with the input signal, we can clearly see that for both fine and coarse crackle samples, all the fine and coarse crackles are separated from breath noise into the non-stationary signal estimate with their time duration and morphology preserved. For both fine and coarse crackles, the breath noise is retained in the stationary estimate with its proper shape and amplitude.

5.2. Comparison of the IEM-FD filter with the WT-FD filter

The separation of the IEM-FD filter was compared with the previously published WT-FD filter [15,16] in terms of Rate of Detectability (D_R), Total Performance ($TD_R^{\Sigma X}$), Quality Factors for of crackle separation (over- or under- estimation), 2CD percentage error (PE_{2CD}) and computational complexity.

5.2.1. Rate of detectability (D_R) and Total Performance ($TD_R^{\Sigma X}$)

Table 3 shows the Rate of Detectability (D_R) and the Total Performance ($TD_R^{\Sigma X}$) for the IEM-FD filter and the WT-FD filter for test samples at SNR $= -1$ dB for real and simulated fine crackles, and at SNR $= 1$ dB for real and simulated coarse crackles and for a real breath sound with fine crackles and a real breath sound with coarse crackles. We note that for both methods (D_R) is the same except for RFC, RCC and RBFC where the IEM-FD give higher values than the WT-FD leading to an overall higher ($TD_R^{\Sigma X}$) for the IEM-FD. For RBFC both filters show a lower D_R than for other signals and this is due to crackles remaining in the stationary signal. This can be rectified either by changing the FDPP algorithm accuracy level (β_3) or by changing the accuracy level (β_2) for stopping iteration of the IEM-FD, but only at the cost of increasing over-estimation.

5.2.2. Quality factors

Table 5 shows the Quality Factors for crackle separation of the IEM-FD filter and the WT-FD filter in terms of over- or under-estimation. We observe that the average under-estimation Quality Factor ($\bar{F}_{QF_U}$) for both filters is either very close or equal to the average reference under-estimation Quality Factor ($\bar{R}_{QF_U}$) indicating that there is very little under-estimation. For over-estimation, we observe that the average over- estimation Quality Factor ($\bar{F}_{QF_O}$) of the IEM-FD filter is generally much closer to 1 compared to that for the WT-FD filter in all cases of fine and coarse crackles, indicating less over estimation in the proposed IEM-FD filter compared to the WT-FD filter. A comparison between the outputs of the two filters is shown in Fig. 6 for a 0.743 second section of the RBCC signal. The location of the crackles was audio visually identified by an experienced pulmonary acoustics researcher and marked with arrowheads. The non-stationary and stationary parts after applying the IEM-FD filter are shown in Fig. 6 ii (a) and (b), respectively.

Appendix A

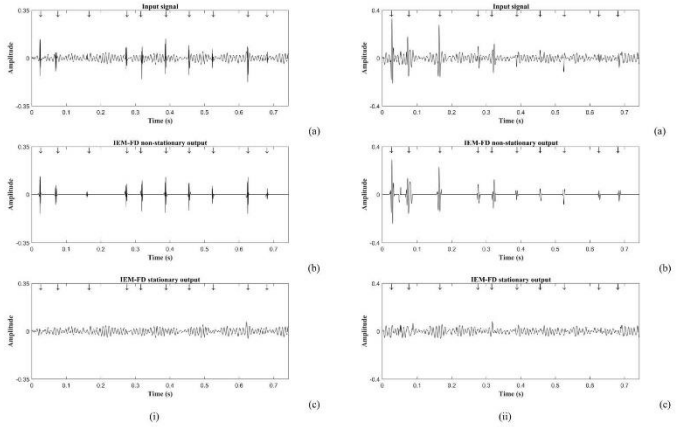


Fig. 5. (i): (a) Input signal with real fine crackles (Table 3, Case RFC one test sample out of 501); (b) IEM FD filter non stationary output; (c) IEM FD filter stationary output. (ii): (a) Input signal with real coarse crackles (Table 3, Case RCC one sample out of 501); (b) IEM-FD filter non-stationary output; (c) IEM-FD filter stationary output.

Table 5
Quality Factor for crackle separation by the IEM FD filter.

Cases		BN	D_g	SNR (dB)	NOTS	$\overline{R}_{QF_S}(ND)$	$\overline{R}_{QF_S}(ND)$	IEM-FD $\overline{E}_{QF_S}(ND)$	IEM-FD $\overline{E}_{QF_S}(ND)$	WT-FD $\overline{E}_{QF_C}(ND)$	WT-FD $\overline{E}_{QF_C}(ND)$
	A_F	BR_N	NA	−1	501	0.817 (0.004)	0.198 (0.005)	0.817 (0.004)	0.888 (0.032)	0.817 (0.004)	0.720 (0.036)
SFC	H_F	BR_N	NA	−1	501	0.848 (0.003)	0.161 (0.004)	0.848 (0.003)	0.958 (0.015)	0.848 (0.003)	0.847 (0.032)
	C_F	BR_N	NA	−1	501	0.801 (0.004)	0.217 (0.006)	0.801 (0.004)	0.828 (0.036)	0.801 (0.004)	0.664 (0.033)
RFC		BR_N	IPF	−1	501	0.828 (0.003)	0.177 (0.005)	0.839 (0.003)	0.959 (0.008)	0.829 (0.003)	0.872 (0.014)
	A_C	BR_N	NA	1	501	0.704 (0.006)	0.325 (0.012)	0.705 (0.006)	0.770 (0.036)	0.704 (0.006)	0.650 (0.038)
SCC	H_C	BR_N	NA	1	501	0.777 (0.004)	0.236 (0.008)	0.779 (0.005)	0.932 (0.025)	0.777 (0.004)	0.854 (0.031)
	C_C	BR_N	NA	1	501	0.710 (0.005)	0.316 (0.010)	0.711 (0.006)	0.790 (0.033)	0.710 (0.005)	0.653 (0.036)
RCC		BR_N	B_r	1	501	0.731 (0.006)	0.292 (0.010)	0.747 (0.008)	0.871 (0.033)	0.732 (0.006)	0.803 (0.041)

SFC: Simulated fine crackles; A_F: IDW = 0.7 ms & 2CD = 5 ms [5]; H_F: IDW = 0.5 ms & 2CD = 3.3 ms [45]; C_F: IDW = 0.9 ms & 2CD = 6 ms [46]; RFC: Real fine crackles; SCC: Simulated coarse crackles; A_C: IDW = 1.5 ms & 2CD = 10 ms [5]; H_C: IDW = 1 ms & 2CD = 5.1 ms [45]; C_C: IDW = 1.25 ms & 2CD = 9.5 ms [46]; RCC: Real coarse crackles; BN: Background noise; BR_N: Breath noise; IPF: Idiopathic pulmonary fibrosis; B_r: Bronchiectasis; D_g: Diagnosis; SNR: Signal to noise ratio; NOTS: number of test samples; R_{QF_S}: Mean of reference under estimation Quality Factor; E_{QF_S}: Mean of estimated under estimation Quality Factor; R_{QF_C}: Mean of reference over-estimation Quality Factor; E_{QF_C}: Mean of estimated over-estimation Quality Factor; *SD*: Standard deviation; NA: Not applicable; In all cases number of samples (N) = 32,768.

Comparing these results with the input signal, we note that all the coarse crackles are separated from normal breath sound into the non-stationary signal estimate with their time duration and morphology preserved, and that the normal breath sound is retained in the stationary estimate with its proper shape and amplitude. For the same input signal, the non-stationary and stationary outputs of the WT-FD filter are displayed in Fig. 6 iii (a) and (b), respectively. Here we see that the non stationary output (Fig. 6-iii (a)) of the WT-FD filter contains the crackles but also a part of the normal breath sound due to over-estimation. Moreover, in the stationary output (Fig. 6-iii (b)) at the location of crackles, normal breath sound segments are missing. This occurs due to domination in magnitude of WT coefficients related to crackles over the WT coefficients corresponding to normal breath sounds [16]. Although over-estimation does not alter the number of crackles in the non-stationary signal, it can affect the morphology of the crackles and therefore is not ideal if crackle characteristics (IDW, 2CD), rather than just number, are important.

5.2.3. 2CD percentage error (PE_{2CD})

Table 6 shows the separation outcomes of the IEM-FD and the WT-FD

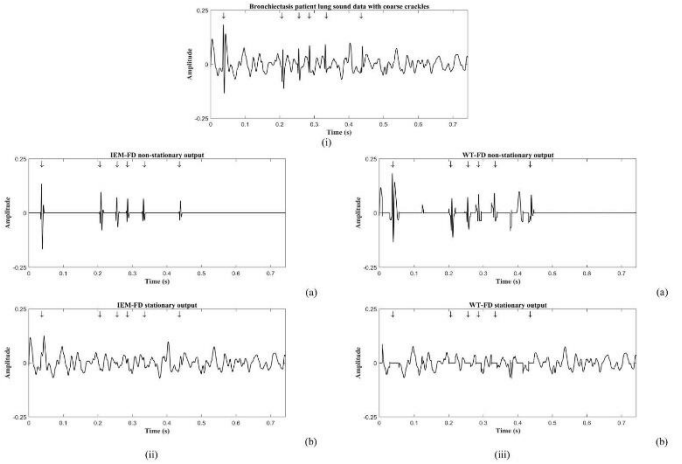


Fig. 6. Comparison between crackle separation of the IEM-FD filter and the WT-FD filter; (i) Time section of 0.743 s of real breath sound with coarse crackles (Table 3, Case RBCC) recorded from a patient with bronchiectasis. (ii): (a) IEM FD filter non stationary output; (b) IEM FD filter stationary output. (iii): (a) WT FD filter non-stationary output; (b) WT-FD filter stationary output.

Table 6
Evaluation of IEM-FD filter separation in terms of 2CD percentage error.

Cases		BN	D_x	SNR (dB)	NOTS	NOC	$\overline{AC}_{2CD}(SD)$ (ms)	IEM-FD		WT-FD	
								$\overline{EC}_{2CD}(SD)$ (ms)	$\overline{PE}_{2CD}(SD)$ (%)	$\overline{EC}_{2CD}(SD)$ (ms)	$\overline{PE}_{2CD}(SD)$ (%)
SFC	A_F	BR_N	NA	−1	501	5010	5 (0)	5.580 (0.510)	12 (10.058)	7.410 (2.254)	58.088 (31.972)
	H_F	BR_N	NA	−1	501	5010	3.3 (0)	4.159 (0.612)	26.496 (18.601)	5.871 (2.748)	87.678 (73.919)
	C_F	BR_N	NA	−1	501	5010	6 (0)	6.419 (0.404)	7.910 (5.940)	8.427 (1.883)	46.733 (21.604)
RFC		BR_N	IPF	−1	501	5010	3.519 (0.103)	3.446 (0.085)	2.375 (0.865)	5.637 (2.555)	63.853 (69.716)
	A_C	BR_N	NA	1	501	5010	10 (0)	9.297 (0.505)	7.824 (3.308)	11.425 (2.602)	21.078 (21.120)
SCC	H_C	BR_N	NA	1	501	5010	5.1 (0)	5.594 (0.512)	10.160 (10.051)	7.719 (2.095)	58.953 (30.445)
	C_C	BR_N	NA	1	501	5010	9.5 (0)	8.907 (0.394)	6.754 (2.828)	10.886 (2.352)	19.552 (21.153)
RCC		BR_N	B_t	1	501	5010	8.406 (1.532)	8.316 (1.427)	10.068 (7.702)	11.107 (2.682)	36.707 (32.675)

SFC: Simulated fine crackles; A_F: IDW = 0.7 ms & 2CD = 5 ms [5]; H_F: IDW = 0.5 ms & 2CD = 3.3 ms [45]; C_F: IDW = 0.9 ms & 2CD = 6 ms [46]; RFC: Real fine crackles; SCC: Simulated coarse crackles; A_C: IDW = 1.5 ms & 2CD = 10 ms [5]; H_C: IDW = 1 ms & 2CD = 5.1 ms [45]; C_C: IDW = 1.25 ms & 2CD = 9.5 ms [46]; RCC: Real coarse crackles; BN: Background noise; BR_N: Breath noise; IPF: Idiopathic pulmonary fibrosis; B_t: Bronchiectasis; D_x: Diagnosis; SNR: Signal to noise ratio; NOTS: number of test samples; NOC: Number of crackles (10 crackles in each test sample); $\overline{AC}_{2CD}$: Mean of actual crackles 2CD; $\overline{EC}_{2CD}$: Mean of estimated crackles 2CD; $\overline{PE}_{2CD}$: Mean of 2CD percentage error; *SD*: Standard deviation; NA: Not applicable; In all cases number of samples (N) = *32,768*.

filters in terms of 2CD percentage error. For the IEM FD filter, the average 2CD percentage error is no more than 27 % for fine crackles (SFC and RFC) and less than 11 % for coarse crackles (SCC and RCC). For the WT-FD filter, the average 2CD percentage error is between 46 % and 88 % for fine crackles (SFC and RFC) and between 19 % and 59 % for coarse crackles (SCC and RCC). The increased error in the output from the WT FD filter relates to uncertainties introduced by the over estimation.

5.2.4. Computational cost

The computational cost for the FD technique is $2(N - W_{FD} + 1)[2(W_{FD} + L1) + 1] + 4L1 + 1$ additions and $2(N - W_{FD} + 1)(W_{FD} + L1 + 2) + 8L1$ multiplication [16,17], where, N is the number of samples in the input signal, W_{FD} is the fractal dimension window length and $L1$ is the maximum number of peeling levels in the FDPP algorithm. The IEM method requires at least $O(QN)$ operations for number of iterations Q and signal length of N. On the other hand, the MRD-MRR procedure in

Appendix A

the WT-FD filter requires $O(KN \log N)$ [16] operations for the number of iterations K and signal length of N. In our analysis for number of samples $N = 32,768$ and maximum number of iterations $Q_{max} = 2$ the IEM method requires O (65,536) operations, whereas for the same number of samples and for maximum number of iterations $K_{max} = 1$, the WT in the WT-FD filter requires O (147,962) operations.

6. Discussion

Separation of crackles from normal breath sounds is an initial processing stage towards good estimation of number of crackles and their time domain features. The separation of crackles from normal breath sounds with low over- or under-estimation aids accurate measurement of crackle time domain features which can help to differentiate between cardio- pulmonary diseases with high sensitivity and specificity. For example, Flietstra et al. [35] showed that based on their different crackle features (number of crackles in inspiratory phase, number of zero-line crossings, and first half period of the crackle) idiopathic pulmonary fibrosis patients can be differentiated from patients with pneumonia and congestive heart failure. Munakata et al. [6] identified that baseline drift over the duration of a crackle (a consequence of over-estimation) may introduce errors when calculating 1DW and 2CD leading to incorrect classification of crackle type (fine or coarse) and increasing the potential for misdiagnosis.

The automatic separation of crackles from lung sounds using the IEM FD filter shows good potential for a single channel, computer based separation of crackles from breath sounds with low over- and under-estimation, high Rate of Detectability, good robustness to noise above SNRs of -1 dB for fine crackles and 1 dB for coarse crackles, well preserved morphology and high processing speed.

In comparison with the established WT-FD filter, [15,16] the IEM-FD filter has an equally high Rate of Detectability for both fine and coarse crackles except for test signals RFC, RCC and RBFC where the IEM outperforms WT FD leading to a significantly better Total Performance for both fine ($t(8016) = 5.12$, $p < 0.000$) and coarse crackles ($t(8016) = 11.99$, $p < 0.000$). In addition, the IEM-FD had less over-estimation, a lower 2CD percentage error and lower computational complexity. Further, the IEM FD filter has fewer data dependent optimization parameters than the WT-FD filter making it generally applicable to signals recorded from cardiopulmonary patients with different diagnoses without the need for data dependent customization to optimize the separation.

Recently Garcia et al. [25] have showed that ICA coupled with a statistical measure offers a promising crackle separation technique and, coupled with spectral analysis, allows the number of crackles present to be identified. However, ICA requires simultaneous measurements from at least as many separate recording channels as there are independent sound generation mechanisms, and therefore requires a bespoke recording system, whereas our analysis can be carried out on a single recording channel for example from an electronic stethoscope.

However, the IEM-FD filter, has several limitations: First, the selection of the smoothing filter (Savizky Golay) parameters in the IEM method is not adaptive which may mean that, in the presence of high frequency background noise, the envelope mean value is not properly estimated; Second, the dependency of the IEM-FD filter stopping criteria on three non- adaptive accuracy levels : β_1 for the IEM method, β_2 for the IEM-FD filter and β_3 for the FDPP algorithm may affect whether all crackles are separated and whether there is over- or under-estimation.

7. Conclusions

This paper presented an automatic technique for separating pulmonary crackles from breath sounds: the IEM-FD filter. The IEM-FD filter was evaluated using a publicly available dataset for systematic testing of crackle separation techniques and compared with the previously published WT FD filter. Key findings of this study were: (1) The IEM FD filter can achieve high accuracy for the number of crackles detected with 99.98 % of fine crackles and 99.80 % of coarse crackles detected in our test samples; (2) The IEM-FD filter has low computational cost compared to the established WT FD filter; (3) The IEM FD filter can provide crackle separation with less over-estimation compared to the WT-FD filter and (4) The IEM-FD filter can better preserve crackle morphology after separation compared to the WT-FD filter in both fine and coarse crackle test signals.

We concluded that the IEM-FD filter would be suitable for use in a clinical context for estimating number of crackles or as a first step in classifying crackles (fine or coarse) on the basis of their time domain features. This in turn can assist with diagnosing lung diseases and in monitoring disease progression. Future research will focus on developing filter parameters that are fully adaptive and on evaluating the operation of the IEM-FD on a more diverse dataset recorded from cardiopulmonary patients, which can further test its ability to detect crackles in different pulmonary conditions.

CRediT authorship contribution statement

Ravi Pal: Methodology, Software, Formal analysis, Writing - original draft. **Anna Barney:** Supervision, Funding acquisition, Writing - original draft, Writing - review & editing.

Acknowledgments

This work was supported by the NIHR Southampton Biomedical Research Centre, the Engineering and Physical Sciences Research Council (EPSRC), and the AAIR Charity.

Declaration of Competing Interest

The authors report no declarations of interest.

References

[1] P. Forgacs, Crackles and wheezes, Lancet 290 (7508) (1967) 203–205.
[2] A. Vyshedskiy, R.M. Alhashem, R. Paciej, M. Ebril, I. Rudman, J.J. Fredberg, R. Murphy, Mechanism of inspiratory and expiratory crackles, Chest. 135 (1) (2009) 156–164.
[3] B.Y. Lu, M. Liu, M.L. Hsueh, P.A. Liu, X.W. Li, H.D. Wu, Preliminary study on production of coarse and fine crackles in respiration using a model of communication theory, in: Proc. 22nd Int. Conf. on Adv. Commun. Tech. (ICACT), PyeongChang, Korea (South), 2020, pp. 461–464.
[4] A.R.A. Sovijarvi, L.P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmasso, C. Sacco, M. Rossi, J.E. Earis, Characteristics of breath sounds and adventitious respiratory sounds, Eur. Respir. Rev. 10 (77) (2000) 591–596.
[5] G. Charbonneau, E. Ademovic, B.M.G. Cheetham, L.P. Malmberg, J. Vanderschoot, A.R.A. Sovijarvi, Basic techniques for respiratory sound analysis, Eur. Respir. Rev. 10 (77) (2000) 625–635.
[6] M. Munakata, H. Ukita, I. Doi, Y. Ohtsuka, Y. Masaki, Y. Homma, Y. Kawakami, Spectral and waveform characteristics of fine and coarse crackles, Thorax. 46 (9) (1991) 651–657.
[7] R.X.A. Pramono, S. Bowyer, E. Rodriguez-Villegas, Automatic adventitious respiratory sound analysis: a systematic review, PLoS One 12 (5) (2017).
[8] P. Piirila, A.R.A. Sovijarvi, Crackles: recording, analysis and clinical significance, Eur. Respir. J. 8 (12) (1995) 2139–2148.
[9] B.A. Reyes, N. Olvera- Montes, S. Charleston- Villalobos, R. Gonzalez- Camarena, M. Mejia-Avila, T. Aljama-Corrales, A smartphone-based system for automated bedside detection of crackle sounds in diffuse interstitial pneumonia patients, Sensors (Basel) 18 (11) (2018).
[10] R.L.H. Murphy, S.K. Holford, W.C. Knowler, Visual lung-sound characterization by time-expanded wave- form analysis, N. Engl. J. Med. 296 (1977) 968–971.
[11] Y.A. Tolias, L.J. Hadjileontiadis, S.M. Panas, Real-time separation of discontinuous adventitious sounds from vesicular sounds using a fuzzy rule-based filter, IEEE Trans. Inf. Technol. Biomed. 2 (3) (1998) 204–215.
[12] T. Kaisla, P. Piirila, K. Kallio, E. Paajanen, T. Rosqvist, A.R. Sovijärvi, Original waveform of lung sound crackles: a case study of the effect of high-pass filtration, J. Appl. Physiol. 71 (6) (1991) 2173–2177.
[13] L.J. Hadjileontiadis, S.M. Panas, Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter, IEEE Trans. Biomed. Eng. 44 (12) (1997) 1269–1281.
[14] G. Kandiloglannakis, P. Mastorocostas, Neurofuzzy modelling of lung sounds, Cont. Eng. Sci. 11 (98) (2018) 4879–4890.

[15] L.J. Hadjileontiadis, Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding - Part I: methodology, IEEE Trans. Biomed. Eng. 52 (6) (2005) 1143–1148.
[16] L.J. Hadjileontiadis, Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding - Part II: application results, IEEE Trans. Biomed. Eng. 52 (6) (2005) 1050–1064.
[17] L.J. Hadjileontiadis, I.T. Rekanos, Detection of explosive lung and bowel sounds by means of fractal dimension, IEEE Signal Process. Lett. 10 (10) (2003) 311–314.
[18] L.J. Hadjileontiadis, Empirical mode decomposition and fractal dimension filter: a novel technique for denoising explosive lung sounds, IEEE Eng. Med. Biol. Mag. 26 (1) (2007) 30–39.
[19] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis, Proc. R. Soc. Lond. A 454 (1971) (1998) 903–995.
[20] M. Ono, K. Arakawa, M. Mori, T. Sugimoto, H. Harashima, Separation of fine crackles from vesicular sounds by a nonlinear digital filter, IEEE Trans. Biomed. Eng. 36 (2) (1989) 286–291.
[21] L.J. Hadjileontiadis, S.M. Panas, Nonlinear separation of crackles and squawks from vesicular sounds using third-order statistics, in: Amsterdam, The NetherlandsProc. IEEE 18th EMBS Conf. (EMBS'96), 5, 1996, pp. 2217–2219.
[22] Y.A. Tolias, L.J. Hadjileontiadis, S.M. Panas, A fuzzy rule-based system for real-time separation of crackles from vesicular sounds, in: Proc. 19th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., EMBS'97, Chicago, IL., 1997, pp. 1115–1118.
[23] P.A. Mastorocostas, Y.A. Tolias, J.B. Theocharis, L.J. Hadjileontiadis, S.M. Panas, An orthogonal least squares-based fuzzy filter for real-time analysis of lung sounds, IEEE Trans. Biomed. Eng. 47 (9) (2000) 1165–1176.
[24] X. Lu, M. Bahoura, An integrated automated system for crackles extraction and classification, Biomed. Signal Process. Control 3 (3) (2008) 244–254.
[25] M.E.R. García, S.C. Villalobos, N.C. Villa, A.J. Gonzalez, R.G. Camarena, T. A. Corrales, Automated extraction of fine and coarse crackles by independent component analysis, Health Technol. 10 (2020) 459–463.
[26] D. Bardou, K. Zhang, S.M. Ahmad, Lung sounds classification using convolutional neural networks, Artif. Intell. Med. 88 (2018) 58–69.
[27] L. Aris, E.H.T. Lim, F.M. van de Ven, L. Heunks, F.R. Tuinman, The diagnostic accuracy of lung auscultation in adult patients with acute pulmonary pathologies: a meta- analysis, Sci. Rep. 10 (1) (2020) 7347.
[28] T. Fukumitsu, Y. Obase, Y. Ishimatsu, S. Nakashima, H. Ishimoto, N. Sakamoto, K. Nishitsuji, S. Shiwa, T. Sakai, S. Miyahara, K. Ashizawa, H. Mukae, R. Kozu, The acoustic characteristics of fine crackles predict honeycombing on high-resolution computed tomography, BMC Pulm. Med. 19 (153) (2019).
[29] H. Katsoka, O. Matsuno, Age-related pulmonary crackles (Rales) in asymptomatic cardiovascular patients, Ann. Fam. Med. 6 (3) (2008) 239–245.
[30] S. Gaydos, Clinical auscultation in noisy environments, J. Emerg. Med. 43 (3) (2012) 492–493.
[31] D. Emmanouilidou, M. Elhilali, Characterization of noise contaminations in lung Sound recordings, in: Proc. 35th Annu. Int. Conf. IEEE EMBS, Osaka, Japan, 2013, pp. 2551–2554.
[32] G.-C. Chang, Y.-F. Lai, Performance evaluation and enhancement of lung sound recognition system in two real noisy environments, Comput. Methods Progr. Biomed. 97 (2) (2010) 141–150.
[33] National Institute for Health Research, Diagnostic Technology: automated lung sound analysis for asthma, Horizon Scan Rep. 0016 (2011).
[34] A.J. Young, New technologies and general practice, Br. J. Gen. Pract. 66 (653) (2016) 601–602.
[35] B. Flietstra, N. Markuzon, A. Vyshedskiy, R. Murphy, Automated analysis of crackles in patients with interstitial pulmonary fibrosis, Pulm. Med. 2011 (9) (2010) 1–7.
[36] G. Sgalla, S.L.F. Walsh, N. Sverzellati, S. Fletcher, S. Cerri, B. Dimitrov, D. Nikolic, A. Barney, F. Pancaldi, L. Larcher, F. Luppi, M.G. Jones, D. Davies, L. Richeldi, "Velcro-type" crackles predict specific radiologic features of fibrotic interstitial lung disease, BMC Pulm. Med. 18 (103) (2018).
[37] P. Pitrils, Changes in crackle characteristics during the clinical course of pneumonia, Chest 102 (1) (1992) 176–183.
[38] A. Marquez, A. Bruton, A. Barney, The reliability of lung crackle characteristics in cystic fibrosis and bronchiectasis patients in a clinical setting, Physiol. Meas. 30 (9) (2009) 903–912.
[39] K. Jafarian, M. Aminesiami, K. Hassani, M. Navidbakhsh, M.N. Lahiji, D.J. Doyle, A multi-channel acoustics monitor for perioperative respiratory monitoring : preliminary data, J. Clin. Monit. Comput. 30 (2016) 107–116.
[40] K. Jafarian, K. Hassani, D.J. Doyle, M.N. Lahiji, O.M. Moghaddam, A. Saket, M. Majidi, F. Ladi, Color spectrographic respiratory monitoring from the external ear canal, Clin. Scien. 132 (24) (2018) 2599–2607.
[41] R.L. Murphy, A. Vyshedskiy, V.-A. Power, D. Bana, P. Marinelli, A.W. Tee, R. Paciej, Automated lung sound analysis in patients with pneumonia, Chest 124 (4) (2003), 1908.
[42] M.A. Islam, I. Bandyopadhyaya, P. Bhattacharyya, G. Saha, Multichannel lung sound analysis for asthma detection, Comput. Methods Progr. Biomed. 159 (2018) 111–123.
[43] L. Vannuccini, M. Rossi, C. Pasquali, A new method to detect crackles in respiratory sounds, Technol. Health Care 6 (1) (1998) 75–79.
[44] R. Pal, A. Barney, A dataset for systematic testing of crackle separation techniques, in: Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.- EMBS'19, Berlin, 2019, pp. 4690–4693.
[45] J. Hoevers, R.G. Loudon, Measuring crackles, Chest 98 (5) (1990) 1240–1243.
[46] A. Cohen, Signal processing methods for upper airway and pulmonary dysfunction diagnosis, IEEE Eng. Med. Biol. Mag. 9 (1) (1990) 72–75.

Bibliography

Aiello, M., Bertorelli, G., Bocchino, M., Chetta, A., Fiore-donati, A., Fois, A., … & Sanduzzi, A. (2017). The earlier, the better : Impact of early diagnosis on clinical outcome in idiopathic pulmonary fi brosis. *Pulmonary Pharmacology & Therapeutics*, *44*, 7–15.

Alvarado, A., & Arce, I. (2016). The Art and Science of the Medical Auscultation: Respiratory Sounds. *British Journal of Medicine and Medical Research*, *17* (12), 1–10.

Andres, E., Gass, R., Charloux, A., Brandt, C., & Hentzler, A. (2018). Respiratory sound analysis in the era of evidence-based medicine and the world of medicine 2.0. *Journal of Medicine and Life, 11* (2), 89–106.

Bajaj, V., & Pachori, R. B. (2012). Classification of Seizure and Nonseizure EEG Signals Using Empirical Mode Decomposition. *IEEE Transactions on Information Technology in Biomedicine*, *16* (6), 1135–1142.

Baughman, R.P., Shipley, R. T., Loudon, R. G., & Lower, E. E. (1991). Crackles in interstitial lung disease. Comparison of sarcoidosis and fibrosing alveolitis. *Chest*, *100*(1), 96–101.

Betiencourt, P. E., Bono, E. A. D., Spiegelman, D., Hertzmark, E., & Murphy, R. L. H. (1994). Clinical Utility of Chest Auscultation in Common Pulmonary Diseases. *AmJ Respir Crit Care Med*, *150*, 1291–1297.

Bohadana, A., Izbicki, G., & Kraman, S. S. (2014). Fundamentals of Lung Auscultation. *New England Journal of Medicine, 370 (8)* 744-751.

Bois, R. M. D. (2012). An earlier and more confident diagnosis of idiopathic pulmonary fibrosis. *Eur Respir Rev*, *21*(124), 141–146.

Charbonneau, G., Ademovic, E., Cheetham, B.M.G., Malmberg, L. P., Vanderschoot, J., & Sovijarvi, A.R.A. (2000). Basic techniques for respiratory sound analysis. *Eur Respir Rev*, *10* (77), 625-635.

Charleston-Villalobos, S., González-Camarena, R., Chi-Lem, G., & Aljama-Corrales, T. (2007). Crackle sounds analysis by empirical mode decomposition. Nonlinear and nonstationary signal analysis for distinction of crackles in lung sounds. *IEEE Engineering in Medicine and Biology Magazine*, *26* (1), 40–47.

Chauhan, K., Ng, H. C. H., & Marusic, I. (2010). Empirical mode decomposition and Hilbert transforms for analysis of oil-film interferograms. *Meas. Sci. Technol, 21* (105405)*,* 1-13.

Cheetham, B.M.G., Charbonneau, G., Giordano, A., Helisto, P., & Vanderschoot, J. (2000). Digitization of data for respiratory sound recordings. *Eur Respir Rev*, *10* (77), 621-624.

Chen, X., Shao, J., Long, Y., Que, C., & Zhang, J. (2014). Identification of Velcro rales based on Hilbert – Huang transform. *Physica A*, *401* (8), 34–44.

Chuah, J. S., & Moussavi, Z. K. (2000). Automated Respiratory Phase Detection by Acoustical Means. *In Proc. Systems, Cybernetics & Informatics (SCI) Conf.,* 228–231.

Cohen, A. (1990). Signal Processing Methods for Upper Airway and Pulmonary Dysfunction Diagnosis. *IEEE Engineering in Medicine and Biology Magazine*, *9* (1), 72–75.

Cordier J.-F., & Cottin, V. (2013). Neglected evidence in idiopathic pulmonary fibrosis : from

history to earlier diagnosis. *European Respiratory Journal,* 42(4), 916–923.

Cottin, V., & Cordier, J. F. (2012). Velcro crackles: The key for early diagnosis of idiopathic pulmonary fibrosis? *European Respiratory Journal*, *40* (3), 519–521.

Cottin, V., & Richeldi, L. (2014). Neglected evidence in idiopathic pulmonary fibrosis and the importance of early diagnosis and treatment. *Eur Respir Rev*, *23* (131), 106–110.

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets, Commun. *Pure Appl. Math.*, *41*, 909–996.

Dellweg, D., Haidl, P., Siemon, K., Appelhans, P., & Kohler, D. (2008). Impact of breathing pattern on work of breathing in healthy subjects and patients with COPD. *Respiratory Physiology & Neurobiology*, *161*, 197–200.

Dinis, J., Campos, G., Rodrigues, J. M., & Marques, A. (2012). Respiratory Sound Annotation Software. *In Proceedings of the International Conference on Health Informatics*, 183-188.

Du, M., Chan, F.H.Y., Lam, F. K., & Sun, J. (1997). Crackle detection and classification Based on matched wavelet analysis. *In Proceedings 19th International Conference - IEEE/EMBS*, 1638-1641.

Earis, J. (1992). Lung sounds. *Thorax*, *47*, 671-672.

Ellington, L. E., Gilman, R. H., Tielsch, J. M., Steinhoff, M., Figueroa, D., Rodriguez, S., … & Checkley, W. (2012). Computerised lung sound analysis to improve the specificity of paediatric pneumonia diagnosis in resource-poor settings : protocol and methods for an observational study. *BMJ Open*, 1–7.

Emmanouilidou, D., Mccollum, E. D., Park, D. E., & Elhilali, M. (2018). Computerized Lung Sound Screening for Pediatric Auscultation in Noisy Field Environments. *IEEE Transactions on Biomedical Engineering*, *65* (7), 1564–1574.

Epler, G. R., Carrington, C. B., & Gaensler, E. A. (1978). Crackles (rales) in the interstitial pulmonary diseases. *Chest*, *73* (3), 333-339.

Esteller, R., Vachtsevanos, G., Echauz, J., …& Litt, B. (1999). Fractal dimension characterizes seizure onset in epileptic patients. *In Proc. 1999 IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 2343–2346.

Esteller, R., Vachtsevanos, G., Echauz, J., & Litt, B. (2001). A Comparison of Waveform Fractal Dimension Algorithms. *IEEE Transactions on Circuits and Systems—I: Fundamental Theory and Applications, 48* (2), 177–183.

Flach, P. (2010). ROC analysis. *In: Sammut, C., Webb, G.J. (Eds.), Encyclopedia of Machine Leaming. Springer*, 869-875.

Flietstra, B., Markuzon, N., Vyshedskiy, A., & Murphy, R. (2011). Automated Analysis of Crackles in Patients with Interstitial Pulmonary Fibrosis. *Pulmonary Medicine*, *2011*.

García, M. E. R., Villalobos, S. C., Villa, N. C., González, A. J., Camarena, R. G., & Corrales, T. A. (2020). Automated extraction of fine and coarse crackles by independent component analysis. *Health and Technology*, *10*, 459–463.

Goh, C., Hamadicharef, B., Henderson, G., & Ifeachor, E. (2005). Comparison of Fractal Dimension Algorithms for the Computation of EEG Biomarkers for Dementia. *In 2nd Int. Conf. on Computational Intelligence in Medicine and Healthcare CIMED2005*, 464-471.

Grzywalski, T., Piecuch, M., Szajek, M., Bręborowicz, A., Hafke-dys, H., Kociński, J., Pastusiak, A., & Belluzzo, R. (2019). Practical implementation of artificial intelligence algorithms in pulmonary auscultation examination. *European Journal of Pediatrics,* 883–890.

Guler, I., Polat, H., & Ergun, U. (2005). Combining Neural Network and Genetic Algorithm for Prediction of Lung Sounds. *Journal of Medical Systems, 29 (3), 217-231.*

Gurung, A., Scrafford, C. G., Tielsch, J. M., Levine, O. S., & Checkley, W. (2011). Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: A systematic review and meta-analysis. *Respiratory Medicine*, *105* (9), 1396–1403.

Hadjileontiadis, L. J. (2005). Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding - Part I: Methodology. *IEEE Transactions on Biomedical Engineering*, *52* (6), 11143–11148.

Hadjileontiadis, L. J. (2005). Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding - Part II: Application results. *IEEE Transactions on Biomedical Engineering*, *52* (6), 1050–1064.

Hadjileontiadis, L. J., & Panas, S. M. (1996). Nonlinear separation of crackles and squawks from vesicular sounds using third-order statistics. *18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2217–2219.

Hadjileontiadis, L. J., & Panas, S. M. (1997). Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter. *IEEE Transactions on Biomedical Engineering*, *44* (12), 1269–1281.

Hadjileontiadis, L. J., & Rekanos, I. T. (2003). Detection of Explosive Lung and Bowel Sounds by Means of Fractal Dimension. *IEEE Signal Processing Letters*, *10* (10), 311–314.

Hadjileontiadis, L. J. (2007). Empirical Mode Decomposition and Fractal Dimension Filter: A Novel Techinque for Denoising Explosive Lung Sounds. *IEEE Engineering in Medicine and Biology Magazine*, *26* (1), 30–39.

Hafke-Dys, H., Bręborowicz, A., Kleka, P., Kociński, J., & Biniakowski, A. (2019). The accuracy of lung auscultation in the practice of physicians and medical students. *PLoS ONE, 14* (8), *1-16.*

Henry, B., & Royston, T. J. (2018). Localization of adventitious respiratory sounds. *J. Acoust. Soc. Am.*, *143* (3), 1297–1307.

Hoevers, J., & Loudon, R. G. (1990). Measuring crackles. *Chest*, *98* (5), 1240-1243.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. *Proc. R. Soc. Lond. A., 454* (1971), 903–995.

Huq, S., & Moussavi, Z. (2012). Acoustic breath-phase detection using tracheal breath sounds. *Med Biol Eng Comput, 50,* 297–308.

Jacome, C., & Marques A. (2017). Computerized Respiratory Sounds : Novel Outcomes for Pulmonary Rehabilitation in COPD. *Respiratory Care, 62* (2), 199–208.

Jacome, C., & Marques, A. (2015). Computerized Respiratory Sounds in Patients with COPD : A Systematic Review. *COPD: Journal of Chronic Obstructive Pulmonary Disease, 12* (1), *104-112.*

Jacome, C., Ravn, J., Holsbo, E., Aviles-Solis, J. C., Melbye, H., & Bongo, L. A. (2019). Convolutional Neural Network for Breathing Phase Detection in Lung Sounds. *Sensors, 19, 1798, 1-10.*

Jang, J. S. R. (1993). ANFIS : Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on Systems, Man, and Cybernetics, 23* (3), 665–685.

Kaisla, T., Sovijarvi, A., Pirrila, P., Rajala, H. M., Haltsonen, S., & Rosqvist, T. (1991). Validated method for automatic detection of lung sound crackles. *Medical & Biological Engineering & Computing*, *29* (5), 517-521.

Kandaswamy, A., Sathish Kumar, C., Ramanathan, Rm. PI., Jayaraman, S., & Malmurugan, N. (2004). Neural classification of lung sounds using wavelet coefficients. *Computers in Biology and Medicine*, 34 (6), 523-537.

Kandilogiannakis, G., & Mastorocostas, P. (2018). Neurofuzzy Modelling of Lung Sounds. *Cont. Engi. Sci.*, *11* (98), 4879–4890.

Katila, T., Piirilä, P., Kallio, K., Paajanen, E., Rosqvist, T., & Sovijärvi, A. R. (1991). Original waveform of lung sound crackles: a case study of the effect of high-pass filtration. *Journal of Applied Physiology, 71* (6), 2173–2177.

Katz, M. (1988). Fractals and the analysis of waveforms. *Comput. Biol. Med.*, 18 (3),145–156.

King Jr, T. E., Pardo, A., & Selman, M. (2011). Idiopathic pulmonary fibrosis. *Lancet, 378*, 1949–1961.

Kiyokawa, H., Greenberg, M., Shirota, K., & Pasterkamp, H. (2001). Auditory detection of simulated crackles in breath sounds. *Chest*, *119* (6), 1886–1892.

Kompis, M., Pasterkamp, H., Oh, Y., & Wodicka, G. R. (1997). Distribution of inspiratory and expiratory respiratory. *In Proceedings - 19th International Conference - IEEE/EMBS*, 2047–2050.

Kompis, M., Pasterkamp, H., & Wodicka, G. R. (2001). Acoustic Imaging of the Human Chest. *Chest, 120*, 1309-1321.

Kraman, S. S. (1993). Lung sounds: An Introduction to the Interpretation of the Auscultatory Finding,. *Northbrook, IL: Amer*. *College of Chest Physician*, 19-21.

Lamas, D. J., Kawut, S. M., Bagiella, E., Philip, N., Arcasoy, S.M., & Lederer, D. J. (2011). Delayed Access and Survival in Idiopathic Pulmonary Fibrosis A Cohort Study. *Am J Respir Crit Care Med*, *184*, 842-847.

Lartillot, O., & Toiviainen, P. (2007). A matlab toolbox for musical feature extraction from audio. *In Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07)*, 1–8.

Lei, Y., Lin, J., He, Z., & Zuo, M. J. (2013). A review on empirical mode decomposition in fault diagnosis of rotating machinery. *Mechanical Systems and Signal Processing*, *35* (1), 108-126.

Leng, S., Tan, R. S., Chai, K. T. C., Wang, C., Ghista, D., & Zhong, L. (2015). The electronic stethoscope. *BioMedical Engineering OnLine*, *14* (66), 1–37.

Loudon, R., & Murphy, R. L. H. (1984). Lung sounds. Am. *Rev. Respir. Dis.*, *130* (4),663–673.

Lu, X., & Bahoura, M. (2008). An integrated automated system for crackles extraction and classification. *Biomedical Signal Processing and Control*, *3* (3), 244–254.

Malik, B., Eya, N., Migdadi, H., Ngala, M. J., Abd-Alhameed, R. A., & Noras, J. M. (2017). Design and Development of an Electronic. *In 2017 Internet Technologies and Applications, ITA 2017 - Proceedings of the 7th International Conference*, 324–328.

Marques, A., Bruton, A., & Barney, A. (2009). The reliability of lung crackle characteristics in cystic fibrosis and bronchiectasis patients in a clinical setting. *Physiological Measurement*, *30*, 903–912.

Mastorocostas, P. A., Tolias, Y. A., Theocharis, J. B., Hadjileontiadis, L. J., & Panas, S. M. (2000). An Orthogonal Least Squares-Based Fuzzy Filter for Real-Time Analysis of Lung Sounds. *IEEE Transactions on BiomedicaL Engineering, 47* (9), 1165–1176.

Meltzer, E. B., & Noble, P. W. (2008). Idiopathic pulmonary fibrosis. *Orphanet Journal of Rare Diseases, 3* (8), 1–15.

Meslier, N., Charbonneau, G., & Racineux, J. L. (1995). Wheezes. *Eur Respir J*, *8*, 1942–1948.

Moussavi, Z. K., Leopando, M. T., & Rempe, G. R. (1998). Automated detection of respiratory phases by acoustical means. *In Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 20* (1), 21–24.

Mukaka, M. M. (2012). Statistics Corner: A guide to appropriate use of Correlation coefficient in medical research. *Malawi Medical Journal, 24* (3), 69-71.

Munakata, M., Ukita, H., Doi, I., Ohtsuka, Y., Masaki, Y., Homma, Y., & Kawakami, Y. (1991). Spectral and waveform characteristics of fine and coarse crackles.*Thorax, 46,* 651–657.

Murphy, R. L. H., Bono, E. A. D., & Davidson, F. (1989). Validation of an Automatic Crackle (Rale) Counter. *Am Rev Respir Dis, 140,* 1017–1020.

Murphy, R. L. H., Holford, S. K., & Knowler, W. C. (1977). Visual Lung-Sound Characterization by Time-Expanded Wave-Form Analysis. *New England Journal of Medicine*, *296* (17), 968–971.

Murphy, R. L., Vyshedskiy, A., Power-Charnitsky, V.-A., Bana, D. S., Marinelli, P. M., Wong-Tse, A., & Paciej, R. (2004). Automated Lung Sound Analysis in Patients With Pneumonia. *Respiratory Care*, *49* (12), 1490–1497.

Nagasaka, Y. (2012). Lung sounds in bronchial asthma. *Allergol Int. 61*, 353-363.

Nath, A. R., & Capel, L. H. (1974). Inspiratory crackles-early and late. *Thorax*, *29*, 223-227.

Obuchowsk, N. A. (2005). Fundamentals of Clinical Research for Radiologists. *Am. J.Roentgenal*. *184*, 364–372.

Oldham, J. M., & Noth, I. (2014). Idiopathic pulmonary fibrosis : Early detection and referral. *Respiratory Medicine*, *108* (6), 819–829.

Oliveira A., & Marques , A. (2014). Respiratory sounds in healthy people: a systematic review. *Respiratory Medicine*, *108* (4), 550-570.

Omitaomu, O. A., Protopopescu, V. A., & Ganguly, A. R. (2011). Empirical Mode Decomposition Technique With Conditional Mutual Information for Denoising Operational Sensor Data. *IEEE Sensors Journal*, *11* (10), 2565–2575.

Ono, M., Arakawa, K., Mori, M., Sugimoto, T., & Harashima, H. (1989), Separation of Fine Crackles from Vesicular Sounds by a Nonlinear Digital Filter. *IEEE Transactions on Biomedical Engineering*, *36* (2), 286-291.

Pasterkamp, H., Brand, P. L. P., Everard, M., Garcia-marcos, L., Melbye, H., & Priftis, K. N. (2016). Towards the standardisation of lung sound nomenclature. *Eur Respir J*, *47*, 724–732.

Piirilä, P., & Sovijärvi, A. R. A. (1995). Crackles : recording , analysis and clinical significance. *Eur*

Respir J, *8*, 2139–2148.

Piirila, P., Sovijarvi, A. R. A., Kaisla, T., Rajala, H. M., & Katila, T. (1991). Crackles in patients with fibrosing alveolitis, bronchiectasis, COPD, and heart failure. *Chest*, *99* (5), 1076–1083.

Pinho, C., Oliveira, A., Jacome, C., Rodrigues, J., & Marques, A. (2015). Automatic Crackle Detection Algorithm Based on Fractal Dimension and Box Filtering. *Procedia Computer Science*, *64*, 705–712.

Polat, H., & Guler, I. (2004). A Simple Computer-Based Measurement and Analysis System of Pulmonary Auscultation Sounds. *Journal of Medical Systems, 28 (6), 665-672.*

Ponte, D. F., Moraes, R., Hizume, D. C., & Alencar, A. M. (2013). Characterization of crackles from patients with fibrosis, heart failure and pneumonia. *Medical Engineering and Physics*, *35* (4), 448-456.

Pramono, R. X. A., Bowyer, S., & Rodriguez-Villegas, E. (2017). Automatic adventitious respiratory sound analysis: A systematic review. *PLoS ONE*, *12* (5), 1-43.

Pramono, R. X. A., Imtiaz, S. A., & Rodriguez-Villegas, E. (2019). Evaluation of features for classification of wheezes and normal respiratory sounds. *PLOS ONE*, *14* (3), 1-21.

Purokivi, M., Hodgson, U., Myllärniemi, M., Salomaa, E.- R., & Kaarteenaho, R. (2017). Are physicians in primary health care able to recognize pulmonary fibrosis ?. *European Clinical Respiratory Journal*, *4* (1), 1-8.

Rao, A., Huynh, E., Royston, T. J., Kornblith, A., & Roy, S. (2019). Acoustic Methods for Pulmonary Diagnosis. *IEEE Reviews in Biomedical Engineering*, *12*, 221–239.

Reichert, S., Gass, R., Brandt, C., & Andres, E. (2008). Analysis of Respiratory Sounds : State of the Art. *Clinical Medicine: Circulatory, Respiratory and Pulmonary Medicine, 2,* 45–58.

Reyes, B. A., Olvera-Montes, N., Charleston-Villalobos, S., González-Camarena, R., Mejía-Ávila, M., & Aljama-Corrales, T. (2018). A Smartphone-Based System for Automated Bedside Detection of Crackle Sounds in Diffuse Interstitial Pneumonia Patients. *Sensors*, *18* (11), 3813, 1-21.

Reyes, B. A., Reljin, N., Kong, Y., Nam, Y., Ha, S., & Chon, K., H. (2016). Towards the Development of a Mobile Phonopneumogram: Automatic Breath-Phase Classification Using Smartphones. *Annals of Biomedical Engineering*, *44* (9), 2746–2759.

Richeldi, L. (2016). How we will diagnose IPF in the future. *QJM: An International Journal of Medicine*, 581–583.

Rocha, B. M., Filos, D.,……& Carvalho, P. D. (2019). An open access database for the evaluation of respiratory sound classification algorithms. *Physiol. Meas., 40* (3), 1-17.

Sarkar, M., Madabhavi, I., Niranjan, N., & Dogra, M. (2015). Auscultation of the respiratory system. *Ann Thorac Med., 10* (3), 158–168.

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem., 36* (8), 1627–1639.

Sellares, J., Hernandez-Gonzalez, F., ……& Xaubet, A. (2016). Auscultation of Velcro Crackles is Associated With Usual Interstitial Pneumonia*. Medicine*, *95* (5), 1-5.

Serbes, G., Sakar, C. O., Kahya, Y. P., & Aydin, N. (2013). Pulmonary crackle detection using time – frequency and time – scale analysis. *Digital Signal Processing*, *23*(3), 1012–1021.

Sevcik, C. (2010). A procedure to Estimate the Fractal Dimension of Waveforms. *Complexity International*, *5*, 1-19.

Sgalla, G. (2017), Characterization of lung sounds for early identification and monitoring of fibrotic Interstitial Lung Disease, University of Southampton.

Sgalla, G., Larici, A. R., …… & Richeldi, L. (2019). Quantitative analysis of lung sounds for monitoring idiopathic pulmonary fibrosis: a prospective pilot study. *European Respiratory Journal*, *53*, 1-4.

Sgalla, G., Walsh, S. L. F., …… & Richeldi, L. (2018). "Velcro- type" crackles predict specific radiologic features of fibrotic interstitial lung disease. *BMC Pulmonary Medicine*, *18* (1):103.

Sovijarvi, A. R. A.,……Earis, J. E. (2000a). Characteristics of breath sounds and adventitious respiratory sounds. *Eur. Respir. Rev., 10* (77), 591-596.

Sovijarvi, A. R. A.,…… & Stoneman, S. A. T. (2000b). Definition of terms for applications of respiratory sounds. *Eur. Respir. Rev., 10* (77), 597-610.

Sovijarvi, A. R. A.,……& Earis, J. E. (2000c). Computerized respiratory sound analysis (CORSA): recommended standards for terms and techniques. ERS Task Force Report. *Eur. Respir. Rev., 10* (77), 585-649.

Speranza, C. G., Ponte, D. F. D., Rocha, C. A. F. D., & Moraes, R. (2020). Blind Equalization of Lung Crackle Sounds to Compensate Chest Attenuation. *IEEE Journal of Biomedical and Health Informatics*, *24* (6), 1796 - 1804.

Spieth, P. M., & Zhang, H. (2011). Analyzing lung crackle sounds : stethoscopes and beyond. *Intensive Care Med*, 37, 1238–1239.

Taplidou, S. A., & Hadjileontiadis, L. J. (2007). Wheeze detection based on time-frequency analysis of breath sounds. *Computers in Biology and Medicine, 37*, 1073–1083.

Tarrant, S. C., Ellis, R. E., Flack, F. C., & Selley, W. G. (1997). Comparative Review of Techniques for Recording Respiratory Events at Rest and during Deglutition. *Dysphagia, 12*, 24–38.

Todd, S., Walsted, E. S., Grillo, L., Livingston, R., Menzies-Gow, A., & Hull, J. H. (2018). Novel assessment tool to detect breathing pattern disorder in patients with refractory asthma. *Respirology*, *23*, 284–290.

Tolias, Y. A., Hadjileontiadis, L. J., & Panas, S. M. (1997). A fuzzy rule-based system for real-time separation of crackles from vesicular sounds. *In Proceedings - 19th International Conference - IEEE/EMBS*, 1115–1118.

Tolias, Y. A., Hadjileontiadis, L. J., & Panas, S. M. (1998). Real-time separation of discontinuous adventitious sounds from vesicular sounds using a fuzzy rule-based filter. *IEEE Transactions on Information Technology in Biomedicine, 2* (3), 204–215.

Vannuccini, L., Rossi, M., & Pasquali, G. (1998). A new method to detect crackles in respiratory sounds. *Technol. Health Care*, *6* (1), 75-79.

Vyshedskiy, A., & Murphy, R. (2012). Crackle Pitch Rises Progressively during Inspiration in Pneumonia , CHF, and IPF Patients. *Pulmonary Medicine, 2012*.

Widrow, B., Mccool, J. M., Larimore, M. G., & Johnson, C. R. (1976). Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter. *Proceedings of the IEEE*, *64* (8), 1151-1162.

Wuyts, W., Antin-ozerkis, D., Huggins, J. T., Lacamera, P. P., Spagnolo, P., Vasakova, M., … & Scholand, M. B. (2019). Serious adverse events in patients with idiopathic pulmonary fibrosis in the placebo arms of 6 clinical trials. *Respiratory Medicine*, *150*, 120–125.

Yeginer, M., & Kahya, Y. P. (2005). Modeling of Pulmonary Crackles Using Wavelet Networks. *In Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, 1-4.

Yeginer, M., & Kahya, Y. P. (2008). Elimination of vesicular sounds from pulmonary crackle waveforms. *Computer Methods and Programs in Biomedicine*, *89* (1), 1-13.

Zhang, Y., Zhang, S., & Ji, X. (2018). EEG-based classification of emotions using empirical mode decomposition and autoregressive model. *Multimed Tools Appl*, 77:26697–26710.