ANALYSIS OF TIME-FREQUENCY FEATURES FOR CLASSIFICATION OF ASTHMA SEVERITY LEVEL USING COMPUTERIZED WHEEZE SOUNDS

by

Fizza Ghulam Nabi
(1540611923)

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy

School of Mechatronic Engineering
UNIVERSITI MALAYSIA PERLIS

2019

ACKNOWLEDGMENT

In the name of Almighty Allah, The Most Beneficent, The Most Merciful, who give me the opportunity of PhD. Also, give me courage to resolve all issues one by one during whole period of PhD. I would say that without the help of Allah this challenging journey was impossible. I would like to extend my gratitude to the beloved Holy Prophet Muhammad (PBUH).

I would like to grant my premier thanks to my co-supervisor Prof. Ir. Dr. Kenneth Sundaraj, UTeM for his continuous inspirations for working and publishing papers, constructive criticism, and sincere willingness to solve problems have made this work to be completed within expected time. I would also like to thank my main-supervisor Dr. Lam Chee Kiang for his meaningful support during this work.

I would like to give special thanks to UniMAP, UTeM and Malaysian government. I extend my gratitude to the hospitals – Al-Mustafa Chest Clinic at Wazirabad, Pakistan and District Headquarters Teaching Hospital at Gujranwala, Pakistan for their admirable support for data collection. I would like to convey special thanks to the physicians and all the subjects.

It is my fortune to gratefully acknowledge the moral support and encouragement from my family members especially my beloved husband Ghulam Sajjad Ali Akbar and little prince Ghulam Muhammad Asghar Awan (4.5-year-old) for their equal contribution still being 2787 miles away from me. I would say thanks to my husband for his love, prayers, understanding, encouragement and assistance in this work.

The words and emotions fail me to express my deepest appreciation to my beloved mother-in-law Maqsooda Fatima and mother Asia Bibi for their love and prayer. I am also grateful to my father-in-law Malik Akhter Hussain advocate and father Mlik Ghulam Nabi Awan. I would like to thank to my grandparents Zulifqar Ali, Maqbool Bibi and Bashiran Bibi for their prayers and encouragement. Furthermore, I would say thanks to my sisters, brothers, brothers-in-law, uncles and aunties for their continuous encouragement and support during this work.

I am also grateful to all my laboratory colleagues and specially team members, for their continuous friendly supports, which encourage me to do work attentively. I thank you all very much and proud to be part of your friends. I am thanks full to Muhammad Shahid Iqbal and Rizwan Sadiq (Department of Electrical and Electronic Engineering KOC University Istanbul Turkey) and Muhammad Saeed for their valuable ideas, suggestion and discussion that helped me through this work.

@This item is protected by original copyright

TABLE OF CONTENTS

@This item is protected by original copyright

@This item is protected by original copyright

@This item is protected by original copyright

@This item is protected by original copyright

@This item is protected by original copyright

LIST OF TABLES

@This item is protected by original copyright

@This item is protected by original copyright

@This item is protected by original copyright

LIST OF FIGURES

@This item is protected by original copyright

@This item is protected by original copyright

@This item is protected by original copyright

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ANN	Artificial Neural Network
AP	Average Power
AR	Auto-Regressive
ATS	American Thoracic Society
COPD	Chronic Obstructive Pulmonary Disease
CORSA	Computerized Respiratory Sound Analysis
DB	Doubtful
DFT	Discrete Fourier Transform
EMMD	Empirical Mode Decomposition
FEV_1	Forced Expiratory Volume in one Second
FEV_1%	Forced Expiratory Volume in one Second Percentage
FFT	Fast-Fourier-Transform
FP	False Positive
FPGA	Field Programmable Gate Array
FVC	Force Vital Capacity
GMM	Gaussian Mixture Model
HS	Hilbert Spectrum
IMFCC	Inverted MFCC
IP	Integrated Power
KNN	k-Nearest-Neighbour
LAWDA	Local Adaptive Wheeze Detection Algorithm
LLB	Lower Lung Base
LLB-Expir	LLB Expiratory
LLB-Inspir	LLB Inspiratory
LFCC	Linear Frequency Cepstral Coefficients
LPC	Linear Prediction Coding
LPCC	Linear Prediction Cepstral Coefficients
MANOVA	Multivariate Analysis
MARS	Marburg Respiratory Rounds
MF	Mean Frequency
MFCC	Mel Frequency Cepstrum Coefficient

@This item is protected by original copyright

PCA	Principal Component Analysis
PLPCC	Perceptual Linear Prediction Cepstral Coefficients
PPV	Positive Predictive Value
SBC	Sub-Band based Cepstral Parameters
SI	Spectral Integrated
SNR	Signal-to-Noise Ratio
SSCP	Square Sum of Cross Product
STFT	Short-Time-Fourier-Transform
SVM	Support Vector Machine
TDR	Total Detectability Rate
TF-WD	Time-Frequency Wheeze Detection
TP	True Positive
T-Expir	Trachea Expiratory
T-Inspir	Trachea Inspiratory
VQ	Vector Quantization
WISE	Wireless Digital Stethoscope
WPD	Wavelet Packet Decomposition
WPT	Wavelet Packet Transform
WT	Wavelet Transform

@This item is protected by original copyright

LIST OF SYMBOLS

N	Number of subjects
$\eta 2$	Effect Size
%	Percentage
±	Plus/minus
F-value	Critical value of F-distribution

@This item is protected by original copyright

Analisis Ciri-ciri Frekuensi Masa untuk Klasifikasi Tahap Keparahan Asma menggunakan Bunyi Wheeze Berkomputer

ABSTRAK

Bunyi wheeze dihasilkan kerana terdapat halangan dalam bunyi peparu di kalangan pesakit asma. Sebarang erubatan atau pengurusan pesakit dilakukan mengikut tahap keparahan pesakit asma iaitu ringan, sederhana dan parah. Hasil kajian lepas menunjukkan bahawa analisis dan klasifikasi bunyi wheeze mengikut tahap keparahan pesakit asma menggunakan ciri frekuensi masa untuk dataset yang berbeza mengikut lokasi dan fasa belum pernah dikaji. Objektif kajian ini adalah untuk mengkaji dan mengklasifikasikan bunyi wheeze mengikut tahap keparahan (ringan, sederhana dan parah) pesakit asma menggunakan ciri frekuensi masa. Kajian ini memberi tumpuan kepada pemantauan dan pengurusan diri pesakit asma menggunakan pernafasan pasang surut. Bunyi wheeze yang disegmenkan dan disahkan dikumpul dari trakea dan asas peparu rendah (LLB) daripada 111 pesakit asma semasa pernafasan pasang surut. Data yang dikumpul dibahagikan kepada 9 dataset berdasarkan lokasi auskultasi, dan/atau fasa pernafasan. Bagi setiap segmen, ciri frekuensi, spektral bersepadu (*SI*) dan kuasa (*IP*) bersepadu ditentukan. Selepas itu, analisis statistik univariat dan multivariate dilakukan untuk menyiasat perbezaan ciri-ciri secara terperinci. Klasifikasi kemudiannya dilakukan menggunakan ensemble, sokongan vektor mesin (SVM) dan kaedah k-terdekat tetangga (KNN). Di samping itu, dua kerangka klasifikasi diperkenalkan untuk mengenal pasti klasifikasi yang paling efektif berdasarkan tahap keparahan. Antara sebahagian besar ciri-ciri individu dan frekuensi vector, SI, IP diperhatikan menunjukkan perbezaan ketara ($p < 0.05$) dalam kebanyakan dataset. Secara keseluruhan, *PPV* terbaik diperolehi dengan ciri *IP* untuk sampel ringan, sederhana dan parah adalah 100% (KKN), 92% (SVM) and 94% (ensemble). Nilai μ (SD) untuk ciri-ciri yang dipilih tidak menunjukkan sebarang tingkah laku khusus dan berterusan untuk tahap keparahan dalam kesemua sembilan dataset. Penemuan penyelidikan menggambarkan bahawa pengagihan frekuensi dan tenaga spektrum dalam isyarat yang direkodkan berbeza bergantung kepada lokasi auskultasi (trakea dan LLB), fasa (inspirasi dan expirasi) dan tahap keparahan (ringan, sederhana dan parah). Jika dipertimbangkan dari segi lokasi auskultasi, trakea menghasilkan saiz kesan yang lebih tinggi daripada dataset yang berkaitan dengan LLB. Bagi ciri-ciri SI dan IP, dalam kebanyakan perbandingan, pengelas ensemble menghasilkan prestasi terbaik dari segi kepekaan, spesifikasi dan nilai ramalan positif (PPV). Walau bagaimanapun, ciri berasaskan frekuensi menunjukkan prestasi tertinggi dengan pengelas KNN dan SVM. Sampel dataset trakea menghasilkan prestasi klasifikasi tertinggi berbanding dengan semua dataset lain dalam semua jenis gabungan. Keputusan pengesahan juga didapati pada paras atas secara purata. Pada masa hadapan, teknik pengelasan pembelajaran mendalam dan pengoptimuman ciri boleh dilakukan atas ciri-ciri akustik.

@This item is protected by original copyright

Analysis of Time-Frequency Features for Classification of Asthma Severity Level using Computerized Wheeze Sounds

ABSTRACT

In asthma patients wheeze sounds are produced due to obstruction in lung sounds. Any medication or management of patients is done according mild, moderate and severe condition of asthma patients. Literature review indicates that analysis and classification of wheeze sounds according to severity levels of asthma patients using time-frequency features in different datasets according to location and phase is required to explore more. The objective of this study is to investigate and classify wheeze sounds according to the severity levels (mild, moderate and severe) of asthma patients using time-frequency features. This study focusses on the self-monitoring and self-management of asthma patients using tidal breathing. Segmented and validated wheeze sounds were collected from the trachea and lower lung base (LLB) of 111 asthmatic patients during tidal breathing. The collected data was split into 9 datasets based on the auscultation location, and/or breath phases. For every segment, the frequency-based, spectral integrated (*SI*) and integrated power (*IP*) features were computed. Subsequently, a univariate and multivariate statistical analysis were performed on the features to investigate the significant difference of features in details. Classification was then performed using the ensemble, support vector machine (SVM) and k-nearest neighbor (KNN) methods. In addition, two classification frameworks introduced to identify most effective classification of severity levels. Most of the selected individual features and feature vectors frequency-based, *SI*, *IP* observed indicated significant difference ($p < 0.05$) in majority of datasets. Overall, the best *PPV* for the mild, moderate and severe samples were found to be 100% (KNN), 92% (SVM) and 94% (ensemble) respectively were obtained with *IP* features. The $\mu(SD)$ values of features have not indicated any specific and continues behavior with respect to severity level in all nine datasets. The findings of research illustrate that the distribution of frequency and spectral energy in the recorded signal varies depending on the auscultation location (trachea and LLB), phase (inspiratory and expiratory) and severity levels (mild, moderate and severe). With the consideration of auscultation location trachea-related datasets produce higher effect size than that of LLB-related datasets. For *SI* and *IP* features in most comparisons, the ensemble classifier produced best performance in terms of *sensitivity*, *specificity* and positive predictive value (*PPV*). However, frequency-based features indicated highest performance with the KNN and SVM classifier. Trachea-related datasets samples produced the highest classification performance than all other datasets in all type of combinations. The results of validation also have been found above on average. In future acoustic features, deep learning classification technique and feature optimization can be implemented.

@This item is protected by original copyright

CHAPTER 1 : INTRODUCTION

1.1 Research Background

Auscultation is a clinical procedure that involves listening to the sounds in a human body with a stethoscope and was first developed in 1816 with the invention of the stethoscope (Oud, 2003). Physicians in the field of pulmonary medicine use a stethoscope to listen to respiratory sound with the goal of diagnosing respiratory disorders and abnormalities. Acoustic signals are generated in the lung due to oscillation of the bronchi walls, which produce turbulent air flow during breathing (Mazić, Mirjana, & Džaja, 2015), and these respiratory sounds constitute a powerful source of information regarding lung conditions (Oud, 2003). In the field of pulmonary medicine, respiratory sounds are closely related to pulmonary pathology (Skalicky et al., 2017).

Stethoscope is economical, non-invasive and less time consuming. Further, for the assessment of airway obstruction of asthma patient's spirometer and peak expiratory flow meters are used. These devices work with the principle of forced expiratory breathing manoeuvre. But these methods cannot be used for continues monitoring of patients due to its breathing manoeuvre. It also cannot be used in most severe condition of patient. These methods also depend on the experience of professional to recognize disorder in underline pathology, respiratory sounds of patients also cannot be saved to maintain history of patients. Auscultation physician and patient relation also contributes for making decision about patient conditions, personal management of respiratory sounds at home cannot be done at home. To overcome these drawbacks computerized respiratory

@This item is protected by original copyright

sound analysis started from 1980 (Palaniappan, Sundaraj K, Ahamed, Arjunan, & Sundaraj S, 2013).

1.2 Motivation of the Work

According to the World Health Organization (WHO), 235 million individuals are suffering from asthma (WHO, 2013). These statistics (COPD, 2015; Pneumonia, 2015; WHO, 2013) have driven researchers towards the development of computerized devices for the self-monitoring and self-management of asthma, which are becoming increasingly more necessary and important. To address this effect, physician-assisted devices currently being used are spirometers and peak flow metres. But these devices are predominantly utilized during supervised forced respiratory manoeuvres, which could pose a problem when dealing with children, when performing long-term manipulation and continuous observation of patients, when treating very severe asthmatic conditions and during unsupervised sessions. In addition, wheezing during forced exhalation is not always correlated to the degree of airway obstruction in asthmatic patients, which reveals that force expiratory volume in one second (FEV_1) values obtained using spirometry might not always correlate with the acuteness of asthma (Fiz et al., 1999). Although lung function values provide an indication of the state of the lungs, they are not always related to airway obstruction in asthmatic patients. In addition, spirometers are very costly and not user friendly.

In general, three conditions are considered during the treatment of asthmatic patients: mild, moderate and severe. The management of asthmatic patients requires prior knowledge of the patient condition. Any medicine or treatment provided must suit the

@This item is protected by original copyright

condition of the patient (Education, 2007). It is required to classify wheeze sound of asthma patients according to severity levels.

1.3 Problem Statement

In the field of computerized wheeze sounds analysis mostly database is related to only wheeze and non-wheeze sounds. Furthermore, wheeze sounds analysis most of the authors focused on only the wheeze and non-wheeze classification (Pramono, Bowyer, & Rodriguez-Villegas, 2017). However, in general, three conditions are considered during the medicine or treatment of asthmatic patients: mild, moderate and severe (Education, 2007). One author classified breath sounds according to asthma severity level (Shaharum, Sundaraj, Aniza, Palaniappan, & Helmy, 2016; Shaharum, Sundaraj, Aniza, Palaniappan, & Helmy, 2018). However, in those studies recordings were made from only one auscultation location (LLB), full recordings were used for analysis and wheeze sounds segments (most reliable signal) were not extracted out for classification. Furthermore, inspiratory and expiratory wheeze sounds were not analysed (Shaharum et al., 2016; Shaharum et al., 2018). Studies dealing with the correlation of respiratory sounds and spirometer values [force expiratory volume in one second FEV_1(%), force vital capacity (FVC)] are collecting breathing sounds with forced expiratory manoeuvres (Fiz et al., 2006). Also, in few studies, subjects are with induced asthma or after using medication (Oud, 2003; Oud, Dooijes, & van der Zee, 2000). These conditions change the behaviour of wheeze sounds. Further, these studies performed analysis on full breathing cycle and collected data from only trachea (Oud, 2003; Oud et al., 2000). Furthermore, in literature authors used different features but did not make any relation of features to physiology which is required according to medical point of view as well. In

@This item is protected by original copyright

addition, it is also required to explore characteristics of wheeze sound segments according to severity level obtained from lower lung base (LLB) and further inspiratory and expiratory phases of breath sounds for different medical applications. In fact, analysis of wheeze sounds is a challenging task as the source of wheeze sounds manifestation is same.

It was observed in the literature that most of the studies implemented single-stage classification. However, for multi-class data, single-stage classification performance may be upper bound limited as compared to multi-stage classification. This is because, the complex behaviour of multi-class data can overburden the single-stage classifier. In this aspect, two studies that deal with multi-stage (architecture) classification (I. Mazić et al., 2015; Palaniappan et al., 2015) have been observed, but these studies did not provide the basis for their design of classifier architecture. Nevertheless, classifier system design is essential from an engineering perspective (Catherine et al., 2019). In order to overcome this drawback, additional information from the behaviour of data must be exploited for the purpose as a basis to build the architecture of a multi-stage classifier.

1.4 Research Questions

1) What are suitable features to identify severity levels in asthma patients?
2) What is the behaviour of wheeze sounds according to mild, moderate and severe condition of asthmatic patients using time-frequency features with tidal breathing?
3) What are the univariate and multivariate statics of selected features according to severity level?

@This item is protected by original copyright

4) Is it possible to classify wheeze sounds according to severity level of asthma patients?
5) What are the statistical and classification results of wheeze sound according to severity levels in different datasets (especially at LLB)?
6) What are the validation results of system?

1.5 Objectives of the Study

1) To extract suitable time-frequency features from wheeze segments which are related to complex human respiratory physiology and can explore all datasets related to location, phase and/or combinations.
2) To investigate the characteristics of wheeze sounds from individual features (Univariate) and feature vectors (multivariate analysis of the variance (MANOVA) according to asthmatic patient's severity levels (mild, moderate and severe) in datasets related to location, phase and/or combinations.
3) To conduct two-stage classification based on a multi-stage architecture which is designed from the statistical behaviour of time-frequency features obtained from datasets of wheeze sound recordings of asthmatic patients with mild, moderate and severe severity levels, grouped according to location, phase and/or their combinations.

@This item is protected by original copyright

1.6 Research Scope

1) The protocol of this research has been designed after a detail review of literature and CORSA standard to overcome the shortcomings of previous studies. Protocol of study has been designed to obtain a system which can be used for continues monitoring, self-management and un-supervised management of asthma patients. Data has been collected from patients suffering from asthma patients.
2) Segmented and validated wheeze sounds segments have been subdivided into only nine datasets, based on auscultation location of data collection and/or breathe sounds. Data has been collected from only two auscultation locations trachea and LLB.
3) For this research data has been collected two times in different span of periods, which produce – Database 1 (used for testing and training of machine learning models) and Database 2 (used for the validation of the system/models).
4) Only three time-frequency based features frequency-based, *SI* and *IP* features have been used for analysis. For classification three classifiers Support vector machine (SVM), k-nearest neighbour (KNN) and ensemble have been used to classify wheeze sounds into three categories mild, moderate and severe condition of patients. In addition, two Frameworks have been used for the classification.
5) All pre-processing and filtration have been done by MATLAB® (version 2017a, Math Works, USA). Feature selection was also performed with the MATLAB. For statistical analysis IBM SPSS Statistics (version 20, IBM Corporation, USA) was chosen for analysis. Furthermore, all type of classification done through MATLAB. Same tools were used in order to provide same source for comparison to maintain quality.

@This item is protected by original copyright

1.7 Organization of Thesis

Chapter 1 (current chapter) provides the introduction of respiratory pathology and a background of topic of interest. The limitations of existing methods, the significance of the current study, problem statement, objective of this study, research questions, research scope and the organization of this thesis.

Chapter 2 presents the overview of human respiratory system, lung sounds, wheeze sounds characteristics, sensors and devices for data collection, sensors selection, wheeze sound databases, location for data collection, computerized wheeze sound analysis, wheeze detection or classification (logic-based algorithm, machine learning algorithms, topological method), wheeze characterization (determination of spectral parameters, correlation of airway obstruction and spectra) and research gape.

Chapter 3 describes the design of protocol for data collection, pre-processing of data, and segmentation of wheeze sounds, feature extraction techniques, statistical analysis techniques and machine learning techniques used for this thesis. It discuss all methodology in details.

Chapter 4 elaborates the univariate and multivariate statistical analysis results using selected features on nine datasets. Further it describes the results obtained using machine learning techniques. In addition, it also elaborates the discussion part of the results.

@This item is protected by original copyright

Chapter 5 summarizes the findings of the study, research contribution, research limitation and recommendations for future work.

@This item is protected by original copyright

CHAPTER 2 : LITERATURE REVIEW

This chapter present conventional method used for investigation of respiratory sound (wheeze), disadvantages of existing methods and benefits of using computerized wheeze detection. The main objectives of proposed research are also described in this section.

This chapter presents overview of human respiratory system, lung sounds, wheeze characteristics, respiratory sounds data acquisition (sensors and devices for data collection, wheeze sounds data bases, environment and conditions for data collection, respiratory sounds pre-processing), wheeze detection or classification (logic-based algorithm, machine learning algorithm and topological method), wheeze sound characterization (Determination of spectral parameters, correlation of airway obstruction and spectra), research gap and summary. Asthma patients produce wheeze during airway obstruction. For literature review we have reviewed in detail all studies working with wheeze sounds.

2.1 Overview of Human Respiratory System

The primary function of respiratory system is the exchange of gasses by supplying oxygen to blood cells and getting rid of carbon dioxide. Respiratory track is divided in two major parts; upper respiratory track and lower respiratory track as shown in Figure 2.1 upper airway consist of nasal, nasal cavity to larynx and lower airway consist of trachea and lung. Taking oxygen inside body is known as inhalation and extraction of carbon dioxide is known as exhalation. One respiratory cycle consists of inhalation and

@This item is protected by original copyright

exhalation. Oxygen moves from upper airway to lower airway and then travel to blood through alveoli. There are two lungs: left and right, right lung consists of three lobes (upper, middle and lower) and left lung made of two lobes (upper and lower).

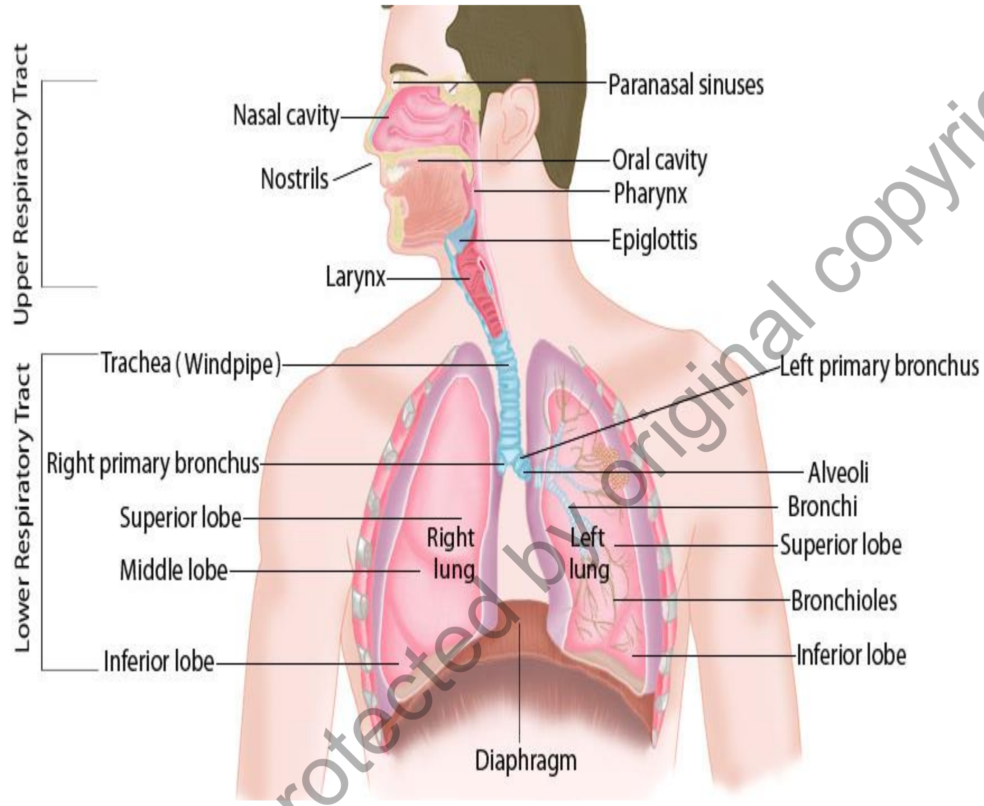

Figure 2.1 Human Respiratory system ("The Respiratory System," 2017)

During breathing acoustic signals are produced in lungs due to turbulent flow which oscillate bronchi walls. Respiratory acoustic signals have meaningful information about lung condition. During normal condition of lungs normal breath sounds are produced while any disorder or obstruction returns abnormal sounds. Wheeze sounds are produced due to airway obstruction in asthmatic patients. A group of researchers calculated airway thickness (wall area) in normal, mild, moderate and severe asthmatic

@This item is protected by original copyright

subjects by computed tomography (Niimi et al., 2000). It was noticed that wall thickness increase by increasing severity level of asthmatic patients (Niimi et al., 2000).

2.2 Lung Sounds

Auscultation is a clinical procedure that involves listening to the sounds in the human body and was first developed in year 1816 with the invention of the stethoscope (Oud, 2003). Physicians in the field of pulmonary medicine use a stethoscope and listen to respiratory sounds with the goal of diagnosing respiratory disorders and abnormalities. Acoustic signals are generated in the lungs due to the oscillation of bronchi walls, which produce turbulent air flow during breathing (Mazić, Bonković, & Džaja, 2015). These respiratory sounds constitute a powerful source of information regarding lung conditions (Oud, 2003). In the field of pulmonary medicine, respiratory sounds are closely related to pulmonary pathology. Accounts of computerized respiratory sound analysis has been appearing in the literature as early as 1980 (Palaniappan, Sundaraj, & Ahamed, 2013).

Respiratory sounds are divided into normal and abnormal (adventitious) sounds depending on their acoustic properties (Palaniappan et al., 2013). A description of lung sounds is also given in Figure 2.2. Normal respiratory sounds are produced by healthy subjects with normal airway and tracheal pathologies, whereas adventitious sounds are produced as a result of airway obstruction and can be subdivided into continuous and discontinuous lung sounds. In addition, continuous sounds are further subdivided into rhonchi and wheezes, and discontinuous respiratory sounds are further subdivided into fine and course crackles (Bahoura, 2009; Palaniappan, et al., 2013).

@This item is protected by original copyright

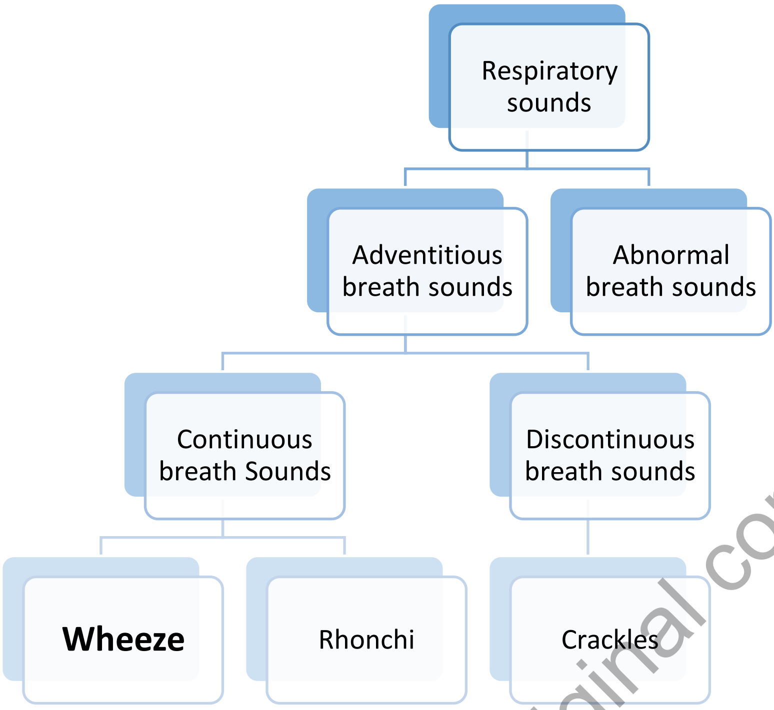

Figure 2.2 Description of types of lung sounds (Pramono et al., 2017)

@This item is protected by original copyright

2.3 Wheeze Sounds Characteristics

Wheeze sounds are produced due to airway obstruction in the airways of respiratory system. These sounds are most frequently heard in patients with asthma but are also detected in patients with several other diseases, such as chronic obstructive pulmonary disease (COPD), pneumonia, pulmonary oedema, and bronchomalacia (Meslier, Charbonneau, & Racineux, 1995). These sounds have also been observed in healthy subjects during forced respiratory manoeuvres (Fiz et al., 2002; Meslier et al., 1995; Shabtai-Musih, Grotberg, & Gavriely, 1992).

Wheezes are continuous sounds that are superimposed on normal breath sounds. American Thoracic Society (ATS) defines the duration of wheeze to be longer than 250 ms and indicates that the dominant frequency of wheeze as 400 Hz (Bahoura, 2009;

Meslier et al., 1995). However, according to the developed Computerized Respiratory Sound Analysis (CORSA) standards, the duration of wheeze is longer than 100 ms and their dominant frequency is greater than 100 Hz (Sovijärvi et al., 2000). CORSA also reports that no wheeze has been found with frequency greater than 1600 Hz (Charbonneau et al., 2000). Wheezes constitute melodic tones with a distinguishable high pitch, and their waveform has a sinusoidal shape (Meslier et al., 1995). Monophonic wheezes have a single pitch, whereas polyphonic wheezes have multiple frequencies (Meslier et al., 1995). It also has been observed that wheezes consist of at most four harmonics (Taplidou & Hadjileontiadis, 2007).

2.4 Respiratory Sounds Data Acquisition

Respiratory data acquisition consists of different steps which includes sensors for data collection, environmental conditions for data collection etc. Details are given below.

2.4.1 Sensors and Devices for Data Collection

Generally, the literature reveals that microphones are used for data acquisition of respiratory sounds. There are two major approaches when microphones are used – kinematic and acoustic. Irrespective of the approach, mechanical vibrations are converted into electric signals through a condenser or piezoelectric sensor. Condenser microphones are normally attached to the skin through couplers known as air couplers while piezoelectric microphones are directly attached to skin surface, for the collection of respiratory sounds (Vannuccini et al., 2000). Few studies have utilized piezoelectric contact-microphones (Fiz et al., 2002; Fiz Jose Antonio et al., 1999; Homs-Corbera, Fiz,

@This item is protected by original copyright

Morera, & Jane, 2004; Shabtai-Musih et al., 1992) while other have used air-coupled microphones (Styliani A. Taplidou, 2004; Taplidou , Hadjileontiadis, Penzel, Gross, & Panas, 2003). The RALE and MARS databases consist of data collected from an accelerometer (EMT-25C) and air-coupled microphone (ECM77) respectively.

Some of studies used custom-made devices to collect respiratory sounds. Two studies acquired respiratory sounds using a device that utilized an electret microphone (EK-3024) and an accelerometer (BU-3173) (Mazić et al., 2015; Mazić et al., 2003). Another group of researchers introduced a device which consists of a condenser sensor (TS-6022A) embedded in a stethoscope bell (Li, Lin, Tsai, Yang, & Lin, 2017). This device was attached to the body through a wearable belt and data was transferred to a computer wirelessly. In another study, diaphragm was attached to a polymer based soft chamber to transmit lung sounds through a microphone (Yu, Tsai, Huang, & Lin, 2013). The authors focused on designing a soft stethoscope for children that is small and flexible (non-rigid). Another study, cut one section of a y-shaped hose of a traditional stethoscope and placed a condenser microphone at this detached end which served as the output for recording signals (Chen, Huang, Tan, Chang, & Chang, 2015).

Air flow measurement devices (spirometers or pneumotachograph) have also been used by researchers to maintain constant air flow rate and to identify or determine the phases (inspiratory, expiratory) of a respiratory cycle. Simultaneous measurement and monitoring of air flow rate with respiratory sound data collection was observed in (Fenton, Pasterkamp, Tal, & Chernick, 1985; Fiz et al., 2002; Homs-Corbera et al., 2004; Styliani A. Taplidou, 2004). Lung function values like Force Vital Capacity (FVC), Forced Expiratory Volume, FEV_1 and FEV_1% in one second and percentage respectively

@This item is protected by original copyright

in (Fenton et al., 1985; Fiz et al., 2002; Homs-Corbera et al., 2004; Lozano, Fiz, & Jané, 2016a; Oud, 2003; Oud & Dooijes, 1996; Oud, Dooijes, & Van der Veen, 1998; Oud et al., 2000; Rietveld, Oud, & Dooijes, 1999; Styliani et al., 2004; Taplidou & Hadjileontiadis, 2007a, 2007b, 2010) have also been parameters of interest by researchers in the quest to observe or correlate wheeze with the degree of airway obstruction.

The aforementioned parameters have also been determined using a domestic device less approach (Palaniappan, et al., 2017). Few studies indicated severity of airway obstruction being determined solely by a physician (Mazić et al., 2003; Oud, 1996; Wisniewski & Zielinski, 2015). Similarly, phases of respiratory cycles have also been segmented solely by physicians through an audio-visual inspection of the recorded respiratory sounds in (Mazić et al., 2015).

Sensor selection for reliable data collection is also an important issue in this field. Two type of sensors air-coupled or contact sensors (accelerometers or piezoelectric) are used to acquire lung sounds. For selection of sensor, its signal-to-noise ratio (SNR) is an important issue and it also has been concluded that there is no difference in SNR of both type of sensors (Pasterkamp, Kraman, Defrain, & Wodicka, 1993). But the flatness of frequency response of air-coupled microphone is effected by width, depth and shape of chamber (Kraman, Wodicka, Oh, & Pasterkamp, 1995; Wodicka, Kraman, Zenk, & Pasterkamp, 1994). Frequencies less than 500 Hz are not affected by depth of chamber. Recommended coupler size is conical shape 10-15 mm diameter and 2.5-5 mm depth. Air-coupled sensors also have the option of vent for proper propagation of vibration, but it also added ambient noise to microphone chamber. Contact sensors are directly attached

@This item is protected by original copyright

to skin so they are free of any type of distortion due to coupling chamber (Vannuccini et al., 2000).

2.5 Wheeze Sound Database

In literature, researchers have used recordings of respiratory sounds (wheezy and normal) from various sources such as:

- Databases – RALE (Database), Marburg Respiratory Rounds (MARS) (Gross, Hadjileontiadis, Penzel, Koehler, & Vogelmeier, 2003), ASTRA (Bahoura, 2009). However, RALE and MARS are two prominent databases.
- Public internet data (Alic, Lackovic, Bilas, Sersic, & Magjarevic, 2007; Riella, Nohama, & Maia, 2009; Wisniewski et al., 2015).
- Accompanying teaching and learning materials intended to train medical students in auscultation (Wilkins, Hodgkin, & Lopez, 2004; Williams & Wilkins, 2009; Wrigley, 2011).
- Self-collected data using different devices.

However, the RALE database is the only commercially available database for respiratory sound analysis. It comprises more than 50 recordings covering all age groups in healthy subjects and those with diseases e.g. pneumonia, asthma and COPD etc. This database which is used to train medical students also provides information about respiratory patterns in young children (Rossi et al., 2000). In contrast, the MARS database

@This item is protected by original copyright

is accompanied by lung function parameters and information on the respiratory phases, was only used by (Taplidou et al., 2007a, 2007b, 2010). MARS consists of data that were collected according to the CORSA standards from mostly normal subjects and patients with asthma, COPD and pneumonia. Currently none of the free or commercially available databases are available according to severity level of asthma patients.

2.5.1 Environment and Conditions for Data Collection

Selection of age group is an important issue, in the literature, mostly author's deal with adult age group or children. As changes in respiratory physiology cannot be totally ignored within age groups (adults and children) (Fiz, Jané, Lozano, Gómez, & Ruiz, 2014; Gross, Dittmar, Penzel, Schuttler, & Wichert, 2000).

Background environment is also an important concern while data collection, according to CORSA recommendation, data should be collected in a sound proof room with environmental noise < 30 dB (Rossi et al., 2000).

Posture of body is also a question while data collection, subjects should be in an easy environment, as movement of body can affect the recordings. Respiratory sound data can be collected in supine posture of body or in sitting position. For long term recording supine position is recommended while for short time recording sitting position is recommended (Rossi et al., 2000). In literature, most of the studies collected data with short term recoding.

@This item is protected by original copyright

It can been noticed in literature that most of the selected studies collected data from only the trachea, three studies from chest (exact location unknown), three studies from LLB, one upper and LLB, two right upper chest anterior, one near mouth and 11 studies from trachea and different combinations of LLB, axilla or upper lung lobe and one study from upper and LLB. Trachea and LLB are the positions selected by 96% of the authors for wheeze data collection. Hence, it can be deduced from the literature that the trachea and LLB are the most frequently chosen positions for wheeze sound data collection. These positions are optimal for the recording of high frequencies (wheezes) with maximum possible information about the underlying pathologies and conditions of the patients. For the implementation and validation of a meaningful real-time computer-based system, it is recommended that as much as possible, reliable and valid data be initially collected and thereafter used in the development and testing stages.

CORSA standard has suggested two breathing manoeuvres for data collection – Tidal breathing and forced expiratory manoeuvres (Rossi et al., 2000). But few studies have proved that wheeze can also be induced in normal subjects with forced expiratory manoeuvres (Fiz et al., 2002; Homs-Corbera et al., 2004; Shabtai-Musih et al., 1992). Which indicates that during forced exhalation wheeze is not always related to air way obstruction (Fiz et al., 1999). In literature few studies did not reported breathing manoeuvres, but it is an important issue.

@This item is protected by original copyright

2.5.2 Respiratory Sounds Pre-processing Techniques

Pre-processing is also an important task for computerized wheeze sound analysis which involves the reduction of noise to obtain only respiratory sounds. The recordings obtained from subjects contain various sounds such as arteffects, hart sounds, environmental sounds like sound of fan, air condition etc. In literature (APENDIX A) most of the authors have used Butterworth filter for the purpose of filtering but selection of order of filter is different. Most of the researchers used high pass filter with the cut off frequency 30-150 Hz and low pass filter with the cut off 1600-3000 Hz. CORSA standard also has provided recommendations related to pre-processing of breath sounds.

2.6 Computerized Wheeze Sound Analysis

In the area of computerized wheeze sound analysis researchers are putting effort for wheeze detection or classification, and wheeze characterization. In it can be noted most of the studies has identified wheeze by three methods – 1) logic-based algorithm, 2) machine learning techniques and 3) topological methods. Furthermore, wheeze characterization consists of – 1) determination of spectral parameters and 2) correlation of airway obstruction and spectra.

2.6.1 Wheeze Detection or Classification

This part of literature includes the identification of manifestation of wheeze sounds in the breath sounds for the diagnosing of patients. Wheeze detection is searching set of peaks using logic-based algorithm to identify wheeze sounds. However, in wheeze classification machine learning algorithm has been used. Details of the methods are given below.

@This item is protected by original copyright

2.6.1.1 Logic-Based Algorithms

Logic-based algorithm involves the search for sets of peaks in the sound spectrum that meet the properties of wheezes. Some articles address the issue using logic-based algorithms (Alic et al., 2007; Fenton et al., 1985; Homs-Corbera et al., 2004; Mazić, 2003; Qiu, Whittaker, Lucas, & Anderson, 2005; Shabtai-Musih et al., 1992; Taplidou et al., 2004; Taplidou et al., 2007b; 2003; Waris, Helistö, Haltsonen, Saarinen, & Sovijärvi, 1998; Yu et al., 2013) as given in Table 2.1. Generally, in this approach researchers attempt to smooth the spectrum, set threshold values and search for cluster(s) of peaks that meet some pre-defined criteria with the aim of increasing the *accuracy*, *sensitivity*, and *specificity* of this process. Smoothing the spectrum generally focuses on window selection and averaging the periodogram. Thresholds are normally determined from the standard deviation of the peaks. A higher threshold value does not detect weak wheezes, and a lower threshold value detects false positives (Qiu et al., 2005). In addition, other criteria have been used by authors to characterize wheeze like behaviour, such as frequency > 100 Hz, duration > 100 ms, wheeze harmonics ≤ 4.

A group of researchers developed algorithm which detects wheezes by searching for peaks in the power spectrum that meet a set of defined rules (Shabtai-Musih et al., 1992). The authors applied their algorithm in an experiment where five healthy non-smoking males performed forced expiratory vital capacity manoeuvres. This algorithm used a logic-based threshold criteria (peaks defined as more than 3.5 times of the values normalized to the variance) and identified those peaks that formed a cluster (a set of nearby peaks) as wheezes, and separated them from the background (Shabtai-Musih et al., 1992). However, the authors have not made any inference to the inspiratory and

@This item is protected by original copyright

expiratory phases. Furthermore, there was no mention of a minimum duration criteria applied to the determined clusters.

Another group of researchers, introduced a logic-based scoring algorithm to ensure the minimum wheeze duration criteria (Alic et al., 2007). The resulting algorithm was applied to wheezy sounds of 26 asthmatic children which were seven from self-recordings and 19 obtained from RALE and online sources. Performance of the system for wheeze detection is given for recorded and online data separately – a) Recorded data – 40 out of 43 true positive (TP), 5 out of 7 doubtful (DB), and 5 false positive (FP) and b) Online Data – true positive 22 out of 22 , doubtful 3 out of 3 and false positive 4 (Alic et al., 2007). It was noted that in this study, a uniform breathing manoeuvre could not be determined as the data used was obtained from various sources. Breathing manoeuvre plays a vital role in the characteristics of the generated wheeze (Rossi et al., 2000). It should be considered in the data collection process.

A group of researchers Homs-Corbera et al., (2004), developed a local adaptive wheeze detection algorithm (LAWDA) that introduces a grouping technique and a modification to the smoothing and threshold used in the algorithm developed by (Shabtai-Musih et al., 1992). Wheezy sounds were collected from the trachea of 16 asthmatic patients, and normal sounds were obtained from 15 healthy (control) subjects. All subjects underwent forced respiratory manoeuvres. The study focused on the exhalation airflow interval of 1.2-0.2 l/s. All wheeze segments were validated by a physician. After the initial peaks search, the algorithm attempts to group each peak with other nearest peaks by eliminating those with duration less than 80 ms and only accepting peaks more than 100 Hz. The LAWDA exhibits *sensitivity* of 100%, 87% and 71% with airflow rates

@This item is protected by original copyright

of 1.2-0.4, 0.4-0.2 and 0.2-0.0 l/s respectively. *Sensitivity* of algorithm was found to be highest at maximum flow rate attained in the expiratory phase and vice-versa. However, this study only considered the expiratory phase in the analysis.

Another study developed a time-frequency wheeze detection algorithm (TF-WD) that introduces a peak coexistence criterion, i.e., the number of peaks at the same instance should not be greater than four (Taplidou et al., 2007b). This algorithm was used in a pilot study consisting of 13 wheezing patients with COPD, asthma and pneumonia. The data from three patients (one with COPD, one with asthma, and one with pneumonia) were used to select appropriate threshold for algorithm (training), and the data from 10 patients were used for analysis (testing) of the algorithm. The time-frequency representation was obtained by Short-Time-Fourier-Transform (STFT) of the magnitude (dB) spectrum. A smoothing procedure was performed using Box filtering (average or mean filter). For threshold setting, the frequency band between 100-1000 Hz was divided into four sub-bands (100-300, 300-500, 500-800, 800-1000 Hz), and the magnitude threshold for each band was determined. A pre-defined duration of 150 ms was used to characterize wheeze. Those identified peaks meeting all the other criteria were defined as wheezes, and all other non-wheezing peaks were discarded (Taplidou et al., 2007b). The TF-WD algorithm exhibits a performance with a total detectability rate of 99.7±1%, a *sensitivity* of 95.5±4.8% and a *specificity* of 93.7±9.3%. The robustness of this algorithm to noise, its accuracy and low computational complexity has been reported (Taplidou et al., 2007b). This work can be improved by increasing the number of frequency sub-bands, changing the wheeze duration criteria from 150 ms to 100 ms as per CORSA standards (Sovijärvi et al., 2000). It was observed that only three subjects

@This item is protected by original copyright

with different pathologies were engaged for setting the threshold, so this work can be improved by recruiting more subjects for the determination of threshold.

Recently group of researchers introduce wheeze detection using compressive sensing Short time Fourier transform using orthogonal matching pursuit (Oletic & Bilas, 2017). Experiment was performed on the data obtained from online sources and collected. Algorithm yielded *sensitivity* 96.28% and accuracy 94.91%. Furthermore, recent work applied an Hidden Markov Model wheeze detection algorithm based on the identification of instantaneous frequency (Oletic et al., 2018). This algorithm tracks the start and end of instantaneous frequency lines in the time-frequency decomposition of the recordings.

@This item is protected by original copyright

Table 2.1 Summary of studies dealing with Logic-Based algorithm

Authors	Subject details	Data collection environment+ ground truth identification	Data collection position(s) / point(s)	Method used for analysis and its purpose	Indication of the severity of bronchial obstruction	Indication of Inspiratory and Expiratory phase (Yes/No) + source of determination	Limitations / Observations
(Fenton et al., 1985)	• 7 children (age 10-16) • 5 with asthma • 2 normal	• Data was self-recorded • Tidal breathing • Wheeze segments were validated by physician using audio-visual inspection of recordings and spectrogram	Trachea and right upper lung lobe	Peak detection by setting threshold within power spectrum obtained using FFT of recordings for wheeze identification	Asthma (4 severe, 1 moderate)	Yes + Flow meter	• Validation of system not reported
Key Findings: 1) The comparison of power spectrums indicate that the trachea is a superior location for wheeze sound recording while the chest wall filters some of the high frequencies. 2) Percentage of wheeze in expiratory and inspiratory phases provides valuable information on the dynamics of wheezing and follows change in lung function. 3) False positive rate < 2% and false negative rate < 2% between normal and asthmatic subjects.							
(Shabtai-Musih et al., 1992)	• 5 male adults • All normal	• Data was self-recorded • Forced respiratory manoeuvres • Validation of wheeze segments not reported	Trachea	Peak detection by logic-based algorithm on the power spectrum obtained using FFT of recordings for wheeze identification	Not applicable	Yes + Flow meter	• Validation of system not reported • No inference was made to the inspiratory and expiratory phases
Key Findings: 1) Wheeze frequency range and spectrogram did not change as a function of the gas density, experiment done with the mixture of air, He, SF_6 and O_2.							
(Waris et al., 1998)	• 8 subjects • 4 asthma patients • 4 normal subjects	• Data was self-recorded • Wheeze segments were validated by physician using audio-visual inspection of recording and time-expanded sinusidal signal waveforms	LLB	Pattern recognition using image processing conducted on spectrogram obtained from FFT of recordings for wheeze identification	Not reported	Yes + Flow meter	• Age range of subjects not reported • Breathing manoeuvre not reported • Validation of system not reported • No inference made to the inspiratory and expiratory phases
Key Findings: 1) Accuracy of system in terms of false positive varies from 1%-9% with respect to the duration of wheezes.							
(Mazić.I, 2003)	• 26 children • 10 with asthma (age 3-7) • 16 normal (age 1-7)	• Data was self-recorded • Tidal breathing and force respiratory manoeuvres • Validation of wheeze identification done by physician using only audio inspection of recordings	Trachea and LLB	Power spectral analysis obtained using FFT of recordings for wheeze identification	Mild to severe asthma identified by physician	Not reported	• Analysis conducted on frequency > 200 Hz which does not fulfil the suggested frequency range of wheeze as suggested by CORSA • Validation of system not reported
Key Findings: 1) Wheeze sound analysis with forced exhalation can be performed at home in asthmatic infants. 2) Accuracy of system reported as 70%.							
(Taplidou et al., 2003)	• 14 subjects • All with asthma	• Data was self-recorded • Tidal breathing	Trachea, upper lung lobe, LLB	Peak detection by logic-based algorithm on the power spectrum obtained using DFT of	Not reported	Not reported	• Age range of subjects not reported

@This item is protected by original copyright