The University Of Sheffield.

A Portable Low-power Electronic Adherence Monitoring System for Cystic Fibrosis

Fei HE

Department of Electronic & Electrical Engineering

University of Sheffield

This thesis is submitted for the degree of

Doctor of Philosophy

October 2019

To my parents and your unconditional love.

Acknowledgements

I would like to acknowledge the help and supervision of Dr Mohammed Benaissa for his superb guidance, support, and understanding during my PhD studies. I would like to express my sincere appreciation to him for the great opportunity to study in this research and the enhancement of my knowledge.

Also, I would like to send my genuine gratitude to Dr Tim Good for helping me building the prototype and helping me during the entire research. Without his continuous support and invaluable direction, achieving this work would have been impossible.

I would like to thank my lab mates, Zheng Hui and Mohammad Eissa, who have helped me with proofreading of this thesis. Thanks for all the advice on the technology of both software and hardware, and important information during studies.

I would also like to thank my parents, my aunt, and my uncle for their unstinting support and love.

My special thanks go to my best friends, PX Zhang, XF Li, YX Li, SR Huang. I was fortunate to have intelligent and diligent peers in my life.

My thanks also go to the department for their kind help during my postgraduate studies.

Abstract

In the healthcare sector, the importance of adherence monitoring has been continuously growing in decades, with the increased need for a reliable convenient electronic system able to detect and monitor administering routine of consumed doses during home treatment. Patients with cystic fibrosis (CF), a severe chronic respiratory disease, require continuous treatment to relieve symptoms. However, the adherence to medical treatment of CF is reported low. The non-compliance may cause low utilisation of drugs, increased hospitalisation, and progressive deterioration. A robust monitoring system becomes crucial to address adherence to the prescribed regimen, predict exacerbation, and evaluate the efficacy of dosed medicine.

This thesis presents a developed electronic adherence monitoring system with a systematic study of inhalation detecting using a novel low power disposable electronic sensor module with limited memory, which shows advantages regarding computational demands and the accuracy of detection. The selected features are proven to be reliable with high accuracy and low computational complexity.

The designed prototype is low-power and capable to work continuously during a 28-day treatment supplied by a thin bendable battery. The portable device can be attached to the target dry powder inhaler and provide effective adherence monitoring with the functionality of audio processing, motion detection, data storage, data transmitting, and a direct feedback system. With real-time recognition and fast decision-making system, it can recognise and record administering events, the date and time of occurrence, and the duration of inhalations with high robustness against the noisy environment.

According to the result of user testing, the device is initially proved to be accurate on detection and duration measurement, and the volunteers described it as convenient to use. It shows potential to improve the quality of inhalation and adherence to medical treatment of respiratory diseases.

Table of contents

List of figures

List of tables

Nomenclature

Acronyms / Abbreviations

FEV_1 Forced Expiratory Volume in 1 second

I^2C Inter-Integrated Circuit

I^2S Inter-IC Sound

3D   3-Dimensional

ACF   Autocorrelation Function

ADC   Analogue-to-Digital Converter

AMDF   Average Magnitude Difference Function

ASR   Automatic Speech Recognition

BIC   Bayesian Information Criterion

BLE   Bluetooth Low Energy

CAMDF   Circular AMDF

CCF   Cross-Correlation Function

CCS   Code Composer Studio

CF   Cystic Fibrosis

CPU   Central Processing Unit

dB   Decibel

dBA   A-weighted Decibels

dBFS   Decibels relative to Full Scale

DFE   Direct time-domain fundamental Frequency Estimation

DIP   Dual In-line Package

DL   Duration Length

DNN   Deep Neural Network

DPI   Dry Powder Inhalers

EAMDF   Extended AMDF

EEPROM   Electrically Erasable Programmable Read-Only Memory

EPROM   Erasable Programmable Read-Only Memory

FeRAM   Ferroelectric Random Access Memory

FFT   Fast Fourier Transform

FLAC   Free Lossless Audio Codec

FNN   Fuzzy Neural Network

FP   False Positive

FRAM   Ferroelectric Random Access Memory

FS   Full Scale

GFCC   Gammatone Frequency Cepstral Coefficients

GMM   Gaussian Mixture Model

GSM   Global System for Mobile

HMM   Hidden Markov Model

HRAMDF   High-Resolution AMDF

Hz   Hertz

I/O   Input and Output

IDE   Integrated Development Environment

IEC   International Electrotechnical Commission

IF   Instantaneous Frequency

IIR   Infinite Impulse Response

IMF   Intrinsic Mode Function

ISM   Industrial, Scientific and Medical

ISO   International Organization for Standardization

KiB   Kibibyte

LED   Light-Emitting Diode

LPC   Linear Predictive Coding

LSB   Least Significant Bit

LSI   Large-Scale Integration

LSP   Line Spectral Pair

MEMS   Micro-Electro-Mechanical Systems

MFCC   Mel-Frequency Cepstral Coefficients

MISO   Master In Slave Out

MUX   Multiplexer

NFC   Near-field communication

NHS   National Health Service

ODR   Output Data Rate

PC   Personal Computer

PCB   Printed Circuit Boards

PDM   Pulse-Density Modulation

PL   Pulse Length

PLP   Perceptual Linear Predictive

pMDI   The pressurised Metered Dose Inhaler

RAM   Random-Access Memory

ReRAM   Resistive Random Access Memory

RF   Radio Frequency

RFID   Radio-Frequency Identification

RGB   Red, Green, and Blue

RMS   Root Mean Square

RNN   Recurrent Neural Network

RTC   Real-Time Clock

SA   Simulated Annealing

SMT   Surface-mount Technology

SNR   Signal-to-Noise Ratio

SPI   Serial Peripheral Interface

SPL   Sound Pressure Level

SPS   Sample Per Second

SRAM   Static Random Access Memory

STE   Short-Time Energy

STFT   Short-Time Fourier Transform

STMD   Short-Time Magnitude Difference

SVM   Support Vector Machine

TD   Time Difference

TESPAR   Time-Encoded Signal Processing and Recognition

TI   Texas Instruments

TP   Ture Positive

USB   Universal Serial Bus

WAV   Waveform audio file format

WPAN   Wireless Personal Area Network

ZCR   Zero-Crossing Rate

Greek letters

μ Mean value

$\psi(t)$ Mother wavelet

ρ_{XY} Pearson Correlation of X and Y

σ_X Standard deviation of X

τ Lag index number

Symbols

p_m Corresponding probability density

Z_n The ZCR of a signal

D AMDF

$E(n)$ Short-time energy

H Entropy

$I(X, Y)$ Information gain between X and Y

$kurt$ Kurtosis

N The total frame shift number or the length of the signal

n   The *n*th frame of the signal

R   ACF

sgn[(n)]   The sign of the *n*th sample

skew   Skewness

U(X,Y)   Symmetric uncertainty of *X* and *Y*

var   Variance

w[n]   The windowing function

WT(a,b)   Wavelet Transform

x(n)   The signal

Chapter 1

Introduction

This chapter introduces a genetic respiratory disorder named cystic fibrosis and dry powder inhalers for symptom-relieving, including an introduction of the target TOBI Podhaler. The motivation for the research project is sequentially clarified with the aims and objectives presented. The key contributions are then reviewed. In the final part of the chapter, the structure of this thesis is outlined.

1.1 Background

1.1.1 The introduction of Cystic Fibrosis

As one of the most common genetically inherited disorder, cystic fibrosis (CF) is currently affecting more than 10,800 people in the United Kingdom [1] and being a common condition in North West European descent [2]. According to an estimation, about 1 in every 25 people in the UK are carriers of cystic fibrosis [3]. With the high mortality rate of newborn babies, the symptoms of CF commonly start in early childhood and continuously impact adolescents and adults. The severe condition affects mostly the respiratory system but also

impairs pancreas, liver or kidneys. The clinical manifestations include chronic endobronchial infection, exaggerated inflammatory response [4], sinus infections, pancreatic deficiencies with attendant malabsorption [5], diarrhoea, frequent coughs, wheezing [3], and clubbing of fingers and toes [6]. Associated with possible diabetes, liver disease, or bone disease, the condition of CF may become even complicated [7].

The CF is described as an inexorable disease and will eventually deprive the health of a patient, and it may be fatal with serious infections or losing functionality of lungs. The reported national median age of survival was only 25 years old in 1993 [2]. With the development of medical science and treatment technology, there are significant advances in the management of infection and visceral insufficiency [7]. Around half of the patients with CF have an improved life expectancy of more than 40 years, and the median survival age was 47 years old in 2016 [1].

1.1.2   Dry powder inhalers

There is no cure for CF at the moment, whilst various palliative treatments exist for symptom-easing, which can prevent or reduce long-term damage caused by repeated chronic infections. Those treatments include antibiotics, physiotherapy, aspiration, and bronchodilators, which are usually combined with regular use of an inhaled mucus-clearing agent [8].

One existing treatment to ease symptoms and relieve pain is delivering medicines directly to the affected lung. As drug delivery devices, both nebulisers and inhalers are commonly used in administering medication. Nebulisers are in the form of the inhaled mist or various respiratory diseases, and medicines are broken into small aerosol droplets or mists by compressed air or ultrasonic power [9]. Nebulisers are mostly placed in hospitals or houses due to the equipment size, thereby bringing inconvenience by its immobility. Furthermore,

the nebulisers have been reported to be relatively inefficient with higher power wasting and time consumption [4].

The inhalers show their advantages compared to the conventional nebulisers, which are handier and with wider design space. The pressurised Metered Dose Inhaler (pMDI) is the most commonly prescribed device and the standard drug delivering mechanism with small airways, which is low-cost, compact, easy to use, independent of inspiratory flow, and has no contamination risk [10]. However, the usage of pMDIs requires the higher coordinating ability of inspiration to achieve correct inhalation and proper deposition of medicine [11]. It is reported that many patients are not using the inhalers properly even after repeated tuition sessions [11–13]. As a result, instead of being delivered to the lungs, the aerosolised drug may impact the back of the throat [10].

Among different types of inhalers, dry powder inhalers (DPI) require less coordination than metered-dose inhalers. The DPIs are also more user-friendly without producing Cold Freon effect, which is a throat reaction to the cold blast [14], as the DPIs contain no propellant gases [11]. In the current market, a variety of DPIs have been provided and widely used in asthma and other chronic obstructive pulmonary diseases.

1.1.3   TOBI podhaler

The target device of this research is TOBI Podhaler [15]. TOBI Podhaler capsules store the functional powder, which is delivered through oral inhalation. An antibacterial aminoglycoside named tobramycin is contained in the capsules. It is the first and only dry powder inhaled antibiotic and prescription medicine for an appropriate cohort of patients with CF and those who are affected by *Pseudomonas Aeruginosa*, which is a bacterium causing serious health risks and living in the thick mucus in the lungs of patients. The chance of having

Pseudomonas Aeruginosa in lungs increases with age, reaching its peak with around an 80% infection rate by 25 years old [15].
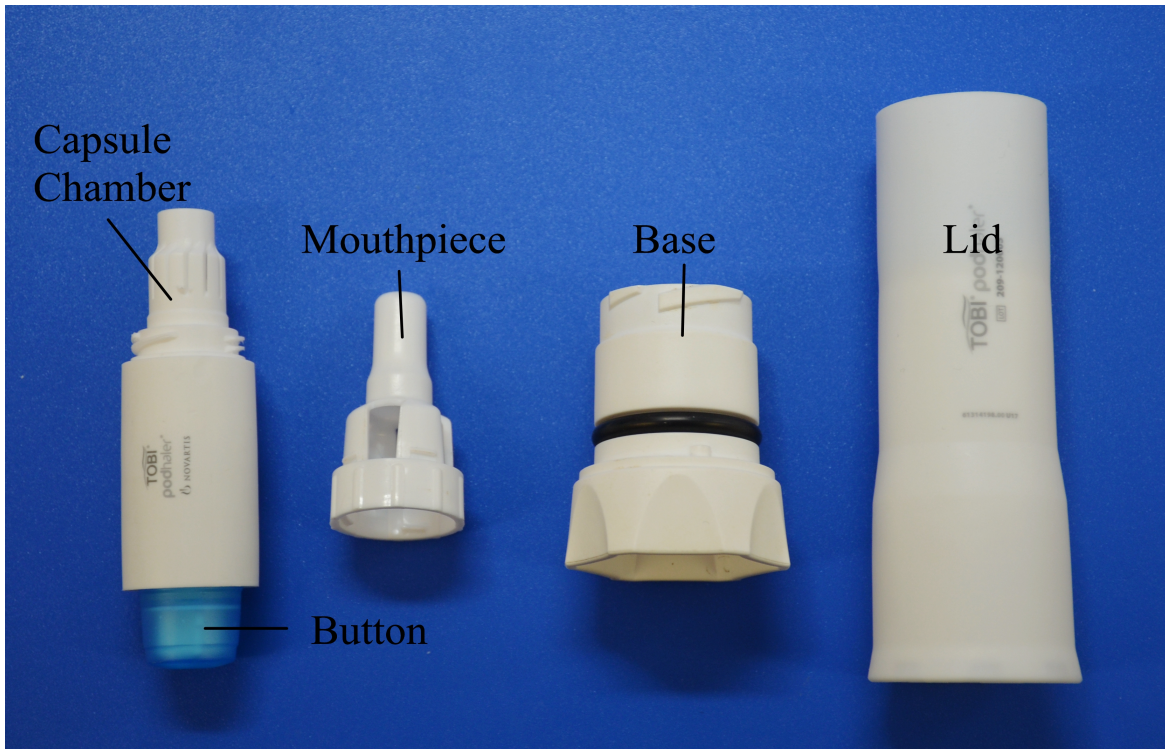
Fig. 1.1 The TOBI Podhaler

Fig. 1.1 shows the structure of a Podhaler. The device comes in a two-piece case, consists of a base piece and a lid. The inhaler itself has a removable mouthpiece, a capsule chamber, and a button for capsule piecing [16].

TOBI Podhaler is compact and lightweight, which is portable and convenient to be carried. It is also easier to be cleaned. The result of a clinical trial showed an improvement of lung function measured by forced expiratory volume in 1 second (FEV1) [15], which means the exhaled amount of air during the first second of a forced expiratory, with 64% fewer respiratory-related hospitalisations and declined need of IV antipseudomonal antibiotics after 28 days of treatment. The measured value of FEV1 is compared to the predicted normal

FEV1 and expressed as a percentage of it. After a full treatment, the FEV1 improved by 12.54% [15].

The device is available as three supply packages for a 28-day, 7-day, or 1-day treatment [16]. The longest treatment follows a 28-day on and 28-day off routine. One administering using TOBI Podhaler takes from 2 to 7 minutes [15]. 4 capsules need to be consumed to receive a full dose, and two doses will be taken each day. There will be 2 or more inhalations from each capsule to empty it. Therefore, approximately 400 - 500 intakes will occur during a full treatment. The price of 8 capsules is from $386.45 [17] according to the current market.

1.2   Research motivation

Chronic respiratory system diseases like CF and asthma are often under medication monitoring as the intensity of the symptoms varies frequently. CF patients usually need a large number of medicines to relieve symptoms and disease progression and require continuous treatment from an early age as CF imposes a demand for full regimens [18]. Around 50% of CF patients are receiving pulmonary medication treatment [19], which has improved the prognosis system significantly.

The primary factors may impact the efficacy of inhaled medications are devices, inhaling technique and patient adherence [20], whereas patient adherence was emphasised as the vital one. Adherence is the term for the success of a patient to follow the instructions regarding correct therapeutic usage or treatment. It is defined as the voluntary and collaborative involvement of patients in the mutually acceptance of behaviour to bring about a preventative or therapeutic result [21], which is similar to patient compliance and demonstrates the progress of pulmonary therapeutic treatment to achieve better management of CF [22].

Poor adherence or non-adherence of treatment has been considered to be a main factor of exacerbation [23], whereas the knowledge of the efficacy of adherence is still limited. Remote monitoring devices can acquire relevant objective information during home treatment [23]. To analyse medication intakes, address medication adherence to the prescribed regimen, predict exacerbation, and evaluate the efficacy of selected medicine, an improved adherence monitoring system, which can detect, identify and measure inhalation, is still in need. Furthermore, additional factors of adherence, including the time and date of administering medications, can improve the predictive accuracy and effect of treatment [22–24]. The system needs to be implemented with a big population of CF patients. For CF, the data of inhalation duration length may provide indispensable information to the clinics. There is still a lack of self-evaluation or feedback system to allow the patients to hone in the new sensing device, especially in nascent periods of usage.

The designed electronic sensing system aims to detect and demarcate inhalation, measure intake duration, record usage frequency, and save the information of adherence, which provides the clinical information of consumed doses. The recorded data should be stored in the device and transmitted to computers or mobiles for onward assessments. Therefore, the device may aid in regulating the using of inhalers and furthermore providing a fitted treatment regimen to individuals.

This work builds upon the knowledge of sound recognition, motion detection, real-time processing, and low-power hardware system. The sound recognition system is selected for inhalation detection and duration measurement. The traditional microphones that have been used in speech analysis or speaker identification system are good-quality high-performance products with inner filters, whereas their financial cost and power consumption are higher. With the development of Micro-Electro-Mechanical Systems (MEMS) devices, sound sensors are getting smaller in size with lower power consumption. Similar to respiratory measure-ments, the developed sound detecting system focuses on inhalation, not the estimated result

from intake movements or unconsumed doses, thus provides more accurate information of administering. Filtering and transformations are vital in pre-processing and identification stage for detecting inhalation from a mixed background sound. Building up a model consists of feature threshold value helps to compose fast and accurate identification.

1.3 Research aim and objectives

The hypothesis of this thesis was: would a technology-based approach to measuring adherence to CF treatment via dry powder inhaler lead to improved health outcomes, quality of life and reduced cost? The aim of this research is to design a low-cost, low-power, lightweight, small-size reliable electronic device to detect and measure inhalation precisely and achieve an efficient adherence monitoring of a CF long-term home treatment, which is only the first monitoring detector of a dry powder inhaler.

The technology is developed to assist in administering, monitoring and help the patient in maintaining the required dosage regime using dry powder inhalers. The ideal inhalation detector should feature efficient detection, accurate measurement, optimised data processing and transmission, high portability, and low energy consumption. It also benefits if the users receive direct feedback of inhalation evaluation or usage guide from the system. The challenges of the research include pre-processing collected data, analysing feature candidates, identifying features which can discriminate proper or improper usages, building up the prototype system to provide the technological underpinnings for a clinical investigation and more involved future work.

The main aims and objectives of this thesis include:

1. Discuss the importance and the existing technologies of adherence monitoring.

2. Review the conventional sound processing technologies and widely-used features of sound detection and recognition.

3. Analyse the collected inhaling data with further feature extraction, comparison, and selection.

4. Describe and explore an accurate sound detecting method with duration measurement based on simple time-domain features.

5. Construct a portable prototype for a demonstration of the proposed methodology with selected algorithms.

6. Design a complete monitoring system with a motion-detecting wake-up system and a light-emitting diode (LED) direct feedback system.

7. Perform a relative evaluation of the constructed system to explore its accuracy and reliability on inhalation detection and adherence monitoring compared to existing technologies.

8. Organise initial user testing experiments with volunteers who have restricted lung functions and have been using inhalers.

9. Incorporate recommendations for a future design for a powered clinical trial.

1.4 Key contributions

1) A low processor-memory solution

In terms of audio processing, the difficulty of real-time inhalation identification under a noisy environment with a limited-sized random-access memory (RAM) is one main challenge in the research. A novel real-time audio processing architecture with a challenging 1 kibibyte (KiB)

processing random-access memory (RAM) is composed. The methodology of abandoning processed raw audio signal from a MEMS microphone reduces the size needed for data storage. Comparing to the frequency-domain counterparts, the calculation of time-domain features requires less memory with higher speed [25, 26]. The reduced processing memory enables the real-time detection and classification system to be implemented on the selected simple low-end processors with good performance.

2) A detection system against noisy environment

Comparing to the existing time-domain approaches based on the zero-crossing rate (ZCR) and energy, the composed novel inhalation detection system is more resistant to the noisy environment. The optimal feature-set has been selected from the candidates based on their performances to present inhalation, the robustness against noises, the orthogonality between features, the computational complexity, and the needed memory space for matrix calculation. The system is capable to detect inhalation from noisy environment of 10 dB signal-to-noise ratio (SNR) or at even higher noise levels. White noise shares high similarity with the target inhalation in both frequency-domain and time-domain as they are both low-volume broadband signals. A novel simple time-domain feature, the Short-Time Magnitude Difference (STMD) is used instead of energy to reduce interference from white noise. The size of the processing window and the abandoning criteria enable an efficient extraction from the contamination of loud sound impulse.

3) An accurate duration length measurement

The study accomplishes an accurate measurement of inhalation duration length as an improve-ment of the existing event detection. The novel rapid decision-making system recognises inhalation in milliseconds and starts counting samples for a later time-length calculation. Instead of measuring the time gap between the opening and closing gestures of container

cover and estimating time length of usage, the duration of actual administering is measured by sound detection, and the duration difference has been declined.

4) Motion detection for the wake-up system

Motion recognition has been introduced into the system for power saving. The simple hand gestures are detected and recognised through a 3-axis accelerometer to trigger the microphone, and the system stays in the sleep mode until specific motion and gesture are detected, as MEMS microphones are costing much more energy than accelerometers. The wake-up system is two-phase, and horizontal alignment of the inhaler is detected in the second phase.

5) Storing capability

The system is capable to store and transmit the information of adherence of a full treatment course. To avoid data loss caused by unexpected instability of power supply, the result of inhalation detection and measurement is stored in the 0.5 KiB non-volatile Ferroelectric RAM (FeRAM) of the near-field communication (NFC) chip. The non-volatile random-access memory (NVRAM) has been used to store pre-set threshold values, parameters, variables, and constant numbers, which contribute to the implementation. The adherence information within a treatment, such as a short treatment of 7 days or a long treatment of 28 days, can be fitted into the limited memory with developed highly-compressed storing solutions. The new simple fast coding algorithms are employed before data storing. A C# PC application has been developed for data transmitting from the device to a PC via an NFC/RFID contactless reader, and the encoding process is conducted after data reading.

6) A low-power portable hardware prototype

A low-power finger-tip hardware prototype to achieve audio and movement processing is constructed in this study. The applied power supply is an internal 43-mAh flat flexible rechargeable Li-Ion battery. To extend the period of use, the power consumption is kept

low during processing and stays in the extra-low mode when it is sleeping or being idle. The average working current is less than 30 uA, allowing a 20-mAh battery operating for a month, which also necessitates a low operational duty cycle less than 1:100 and single-digit micro-amp standby current.

7) A direct feedback system

A direct feedback system is present with the potential to improve the quality of medicine intakes. An attached red/green light-emitting diode (LED) can be observed externally through the shell. The sign of detected inhalation allows the patients to have an early and fast evaluation of their inhaling quality. The transitions between different working modes are also informed through the lights. Single red light flashing, double flashing, long-period flashing or other flashing modes are endowed with different meanings and providing multiple feedback to users.

1.5   Thesis outline

The first chapter of the thesis introduces the cystic fibrosis and its main symptoms. It also describes the popular conventional symptom-easing equipment, including nebulisers and various inhalers. It emphasises the TOBI Podhaler as it is the target DPI. After a brief introduction of adherence to CF, the motivation to build a low-power portable real-time detection and identification device for inhalation is presented. The aim and objectives are stated sequentially with key contributions of the project outlined.

The second chapter of the thesis reviews the importance of adherence monitoring for the treatment of CF. With a literature review of existing adherence or healthcare monitoring system, it emphasises electronic monitoring technology as it is the state-of-art method and has

played an important role in the modern health care system. It describes various technologies that have contributed to adherence monitoring.

The third chapter introduces different sound feature extraction algorithms, including time-domain features, frequency-domain features, and wavelet extraction. The advantages and disadvantages of each extraction method have been discussed and compared. Sequentially, it presents the knowledge of the audio collection, pre-processing, analysis of recorded inhalation, and the filtering algorithm, followed by the feature extraction of target inhalation and the evaluation of the environment. The feature selection for the real-time device also has been discussed in this chapter. The optimal feature set consists of three features, which are easy to be calculated by a limited-sized processor and have the capability to distinguish inhalation exclusively from the surrounding noises. The efficacy of each acoustic feature is analysed separately.

In the fourth chapter, the NFC chip used for data storing and transmission have been considered for data storage and transmission. Instead of the original 16-bit inhalation length, two highly-compressing methodologies are developed to store information of hundreds of inhaling events in the limited-size non-volatile memory safely.

The sixth chapter discusses the experiments of inhalation detection and the result comprehensively. An observed two-phase initial user testing study has been completed with two groups of volunteers to evaluate the performance of the prototype monitoring system. The participants used an inert inhaler without drug delivery. The findings indicate that the device is easy to use with more than 90% accuracy as a self-evaluation system.

The conclusions and major contributions of the thesis are outlined in chapter seven. The suggestion of future work is then proposed.

1.6 Research Publications

The following publications to be submitted based on the research in this thesis is presented below:

1. The portable low-power electronic adherence monitoring system for cystic fibrosis, IEEE Transactions on Biomedical Engineering.

2. Design of low-power real-time audio processing system of monitoring devices for respiratory disease, IEEE Transactions on Biomedical Circuits and Systems.

Chapter 2

The Adherence Monitoring in the Healthcare System

With an introduction of medication adherence for cystic fibrosis, the chapter emphasises the importance of adherence monitoring during a long-term treatment and reviews existing monitoring technologies. Electronic monitoring technology is the state-of-art method and has played an important role in the health care system. Different technologies of adherence monitoring have also been introduced in this chapter.

2.1 The importance of adherence monitoring

Adherence to the treatment plan is one of the recognised common issues among patients with cystic fibrosis (CF). The inappropriate inhaler usage and poor adherence normally result in lower utilisation of drugs, inefficient disease control, and higher consumption of powder [27]. The consequences also include progressive deterioration, excessive prescription of drugs to achieve enough potency, extended duration and the increased cost of hospitalisation [27],

increased outpatient visits, the erroneous conclusion about the efficacy [2], weaken baseline lung function [19], and even a failed treatment [7].

The mean adherence is reported low for long-term regimens and treatments. Notwithstanding that patient and clinics make efforts on the adherence to the treatment, the reinforcement is limited to individual pulmonary medication among patients with cystic fibrosis [7]. The National Health Service (NHS) spends around £30 million annually on inhaled therapy whereas the shown median adherence is only 36% [28]. Inadequate collection of less than 50% of the medication during inhalation may increase the possibility of unscheduled emergency care and additional hospital admission, thus cost the healthcare system significantly compared to those who collect 80% or above [24].

The inefficient inhaling technique is considered one of the inadequacies of daily medical treatment and the main causes of incorrect medicine delivery, alongside with incorrect priming the powder and slow inspiration [12]. The usages of seven inhaler devices, including the Easi-Breathe (Allen & Hanburys), the Autohaler (3M Healthcare), and the most commonly prescribed pMDI, have been recorded and analysed among 100 instructed users by Lenney, Innes, and Crompton in 2000. Some users have used the inhalers effectively, whereas the rest failed to follow expert instruction and the rate of valid drug delivery dropped among them. Only 79% usages of the conventional pMDIs were effective. This assessment graded poor techniques indicating the scant drug delivery with partial, little or no powder received by users. Poor co-ordination, triggering at the wrong stage of inspiration, nose inspiration, induced coughing, incorrect priming or no priming, and other faulty behaviours would lead to a failed inhalation [12]. Besides inaccurate inhaler operating, the inhaling quality also affects the delivery of powder. Also, inspiratory flow, tidal capacity, and breathing rate differ from patient to patient [29]. An inadequate inhalation indisposes treatment quality due to the reduced amount of committed dose. As a consequence of low inspiratory rate, few portions of prescribed medication could finally act on the target organ.

Without quantifiable information of actual usage profile, the medical team is unable to accurately assess adherence and thus may be uncertain in prescribing changes. Consequently, the treatment may be compromised with the inevitable increases in financial burden. To address the problem of incorrect inhalation, adherence monitoring has been introduced to demonstrate whether the users are receiving the correct amount of medicine as advised thus reduces drug abuse or waste.

The reason for measuring failure includes the lack of valid and reliable adherence measurement, no clear delimitation between adherence and non-adherence, and the varied adherence due to the different regimen of each individual [18]. The situation varies between each patient as their age, personality, gender, knowledge of the disease, the frequency of clinic visits, the severity of symptoms, family, or social environment is different [2, 18]. Without a reliable measurement of adherence, there is not enough accurate information about the prevalence of incomplete adherence, therefore the consensus about the importance of each treatment method is still absent among the health care researchers.

Proper measurement of inhalation adherence is imperative in the study of CF, whereas non-compliance is difficult to be recognised without a detailed demonstration against the varieties of patterns. The reasons for non-compliance might be complicated. When the patients dose themselves in the home environment, their adherence may vary over time [30]. The patient may not be aware of the importance of adherence, especially when the symptomatic relief fails to occur immediately and lead to a decline in the belief and confidence in the treatment. The complexity, demands, and duration of a long-term treatment also dispel the belief of patients [18]. The emotional resistance may affect adherence when a patient with depression or anxiety refuses daily doses. The accessibility of medical device is also considered important, the doses may not be taken continuously if the device is difficult to operate.

For adult patients, the forgetfulness is cited as a frequent reason for poor adherence [30]. The patient may miss regular doses during the busy daily life and have not been aware of the negligence. The strong concerns about potential adverse effect may hinder the progress of adherence [7]. Only when the patients are worrying about the disease and believing that the treatment is beneficial and provides positive reinforcement, the adherence is managed successfully [2].

Children suffering from CF require special medical care to maintain health, and their adherence was recorded to be notably higher than older patient groups [22]. The parents may start to transfer treatment responsibility to children with the growth of their age and independence ability. Thus, the adherence of older children, teenagers, and young adults are declined in this transitional period due to the difference of beliefs about treatments between adults and adolescents [31]. With the increased complexity of administering treatment, even adults may find it challenging to understand the efficacy and importance of their own regimen. The uncertainty may transfer from parents to children during an explanation of adherence [7].

2.2 Adherence monitoring technologies

2.2.1 Traditional health diaries and direct measurements

The use of health diaries became a common adherence monitoring method in the early days. The patients were asked to take records of their diary medical regimens, the remaining amount of medicine, and how the symptoms vary during treatment. Health diaries provide a comprehensive view of the patient's condition [32] and make early detection of worsening trends to introduce therapeutic measures in time [33]. To record, assess and properly control CF inhaler adherence, self-report or daily diaries are utilised by the clinicians in existing cases. The University of Minnesota CF Centre aimed to develop a home monitoring model

based on home-monitoring data to evaluate the therapeutic process [34] and implement it into an expert system shell in 1989 [32]. The considered parameters in the program included respiratory rate, vital capacity, resting pulse, and weight. The project introduced the idea that the specific parameters could represent the health condition of patients and had demonstrated an example of a computer-based prototype system.

The novel monitoring and direct measurement of adherence based on self-reported health diaries, including the retrospective analysis of therapeutic response, the usage of tracer substance, and inspection of blood, urinary excretion, or body metabolites [7] have been developed. With a higher requirement of equipment and more human actions involved, the frequent tests may increase both the financial and time cost and be considered as impractical.

Self-report adherence is known to be higher than the measured adherence through electronic monitoring devices, as the former may introduce positive measurement bias [35] or insertional manipulation. The complexity of interpreting a large amount of time-based data increases the difficulty of further analysis and the uncertainty of human-recorded data may lead the statistical errors in the final evaluation [36]. According to statistics, self-reports by independently monitored patients showed low accuracy and low reliability without precise information on usage times or duration for a long study period. Daily diary report from the clinician side showed objectivity, which was considered more promising and accurate, whereas the cost and time consumption would also increase as long-term treatment progressed. The usage of questionnaires can also collect direct or subjective information from the patients [37].

The communication between the patients and the healthcare teams benefits the treatment especially when frequent conversations can be achieved. Despite the important perceived information from patients, physicians may still make a misguided assumption about how good the patient will comply with the adherence following the instructions [7]. As a consequence,

the patients and the clinicians might not be aware of the non-compliance and incorrect usages, causing exacerbation of symptoms eventually.

2.2.2   Electronic adherence monitoring technologies

Electronic monitors are considered as one of two primary objective adherence measurement alongside with pharmacy refill records [19]. Automatic electronic monitoring through dedicated devices is regarded as the optimum method with lower cost, higher accuracy [8], and the capability of providing a date/time stamp of medication administering [19]. The integrated monitored devices measure and record relevant information of usage rather than depend on self-reporting, without placing a time or financial burden on the clinicians. By sensing the correct usage and storing accurate records electronically, the system provides supporting information to demonstrate treatment progress more effectively.

The cost of each plastic inhaler is relatively low comparing to the price of medicine, which costs around $386.45 for 8 capsules [17], thus inhaler can be replaced frequently once it is not in the good condition. The target monitoring device should be designed to be low-cost with an approximate cost lower than 10 pounds. In this way, the electronic device can be replaced and deposed frequently with the used inhaler.

To sense, monitor, and diagnose CF, diverse electronic technologies have been developed in recent years. Comparing to the traditional monitoring methods with more human activities involved, the electronic methods are more objective and independent alongside with the following advantages:

- The electronic methodology has higher accuracy and reliability [38].

- The devices are non-invasive [7] and relatively safer to the human body.

- It reduces the frequency of clinic visiting with continuous data collected in the relatively long-period treatment.

The electronic monitoring for CF and other diseases can be based on various technologies, such as the respiratory measurement, motion detecting, sound detection or perspiration detection [39]. With the development of smartphones, healthcare mobile applications capture the market and become an emerging technology for adherence monitoring. Existing applications include the function of self-report dairies, medication reminders [40] or condition monitors have been used for diseases like diabetes and high blood pressure, which need daily monitoring. Without extra devices, the patients can download applications into their own phones. The time of administering training might be reduced by recorded demonstration audios, animation or videos. The functional processor implanted in smartphones can provide advanced signal processing and solve the transmission issue with Wi-Fi or Bluetooth technologies. As the current applications are more dependent on the action of patients, the accuracy and reliability of reported data are declined as the recorded data are varied by their own medical briefs, which is similar to the situation of self-reports.

Electronic monitoring devices for respiratory diseases can be attached to the nebulisers, aerosol dispensers, and inhalers to measure and record the adherence. Developed electronic technologies include:

1. Ingestible sensors

Besides extracorporeal monitoring methods, ingestible integrated micro-sensors with in-vivo communication system has been designed, optimised, and embedded inside medication for daily adherence monitoring. The functional sensing circuit can be developed in food-particle millimetre-size. The contacting with gastric fluid activates and powers the integrated circuit, providing the electrolytic solution without additional batteries [41]. The adherence

information, alongside with other possible measured properties, can be transmitted from the sensor to an outside receiver, such as an electronic patch [42], a compliant cloud-based web server [43], or a mobile application. An efficient signal processing including channel sensing, frame synchronisation with decimation, auto-correlation and frequency estimation may be introduced and implemented on a low-power platform for further data analysis [42].

The ingestible integrated sensors are small in size, low in power consumption and financial cost and safe to the human body. Direct information about adherence and non-adherence will be detected and recorded with less recall bias or distortion. The identification and transmission system can provide real-time processing with feedback, which can guide the users in nascent periods of treatment to adapt the therapy.

In the case of a monitoring system for inhalation, ingestible sensors are not suitable as the powder will not be through the digestive system and they are not considered in this research.

2. Motion detection

Motion detection is one main approach for adherence monitoring as accelerometers have been embedded in most smart-phones, smart watches or other emerging electronic devices. The signal from real-time movements directly responses to the action with high sensibility. To detect and identify the movements of adherence, the accelerometers sensors can be attached to the pill containers, mobile phones or the human body.

Existing research studies attach importance to upper-limb movements as it can present the action of the medicine administering after proper data training. Zhang and Xu attached 3D accelerometers to a wrist strap and the information of upper-limb movement was collected and processed to detect real-time human motions by normalisation and interpolation [44]. The data were transmitted to a PC for processing through a USB interface.

A research group from Washington State University has developed a learning-based methodology to detect the medicine-taking action of patients. Wrist motions were detected and collected by a wearable sensor on a wristband device. The collected data were transmitted to a computer for signal segmentation, feature extraction, model training and final execution [45]. The accuracy rate of detection achieved 78.3% by a single sensor. Another detection system to recognise a hand-to-mouth gesture has been developed for cigarette-intake identification [46]. The specific gesture is similar to medicine intakes and can inspire adherence monitoring technology of inhalers.

The duration of inhalation can be estimated by the movements of patients. AdhereTech [47] is built in the same size as a pill bottle. It can collect adherence information and obtain valuable insights. A battery-powered tablet-dispenser to monitor the preventive therapy for tuberculosis is developed in 1999 [38]. Instead of measuring the inhalation, the lid of the container is monitored by an electromagnetic detecting system. The date and time, and the duration of the removal of the lid will be detected and recorded into the memory. An interface box is used for data transmission from the memory to a computer. The battery and memory have the capability to continue for years.

In 2014, a new adherence monitoring attached electronic device SmartTurbo, which is designed for a dry powder inhaler named Turbuhaler, is developed [48]. The evaluation stimulated the movements of actual usages. The opposite rotations of the inhaler are detected to identify inhalation. A USB cable connection is used for data transmission from the device to a web-based database.

3. Vision recognition

Medication intakes can be monitoring by a computer vision system with multi-level hierarchical approaches. Instead of tracking the pill itself, the human activities are detected and

tracked remotely for the home-care system. Based on proximity and occlusions [49], the occurred interactions between hands or pill bottles are captured by a camera-based system and recognised by further image processing. Face identification has been introduced to make sure the right user is under monitoring. The face region can be detected and constrained, and the face will be recognised and tracked by template matching with assumptions. Hand localisation [50] based on greyscale sharpness is developed to track the hands, and the orientation of fingers are detected to recognise the movements of opening or closing a medication bottle [51]. The duration, the actual date and time of an event, and the gaps between two intakes are recorded into the electronic system. A high-performance camera may present the information to measure the dosage amount. With a previous advanced training stage, the quality of administering can be estimated to grade normal and abnormal intakes.

Building an entire vision recognising system may lead to a higher cost of time and money. For people with dementia, the research group from Fukuoka University [52] developed a Kinect-based medication adherence monitoring system. The depth-image analysis technology was provided by Kinect, which is an existing motion-sensing line that was produced by Microsoft for video game consoles and PCs.

4. Sound recognition

Acoustic analysis has been introduced in the health care system especially for those diseases with specific sounds, such as breathing and snoring sounds. The functionality of microphones has been improved whereas their size shrinks. Sound detection plays an important role in state-of-art monitoring and other healthcare systems, such as pre-diagnosis of heart disease [53], sleep quality diagnosis [54], environmental sound detection for hearing-impaired people [55], sleep apnea detection [56], coronary artery disease [57], and electronic auscultation of lung sound detection [58].

As the acoustic features of inhalation relate to the expiratory volume [59], sound processing has been introduced into respiratory adherence monitoring. Martin S Holmes indicated clinical implications on measuring whether the patients reach the peak inspiratory flow rate as it presented the feasibility of using acoustics to detect a dry-powder inhalation objectively [20]. Temporal acoustic detection of asthma inhalation was firstly introduced in 2012 [60], presenting empirical evidence to an asthma medication regimen. Raw audio signals which contained inhalation were recorded by a Micro-electro-mechanical (MEMS) microphone and were saved in a Secure Digital (SD) card. The inhalation was detected and demarcated by a developed algorithm after the signals were downloaded to the analysing platform. Faulty usages, including low flow rate and multiple blisters, would be identified by the system. The result of the clinical investigation was published in 2016, showing that the remote monitoring device excluded incorrect usages to acquire objective information. As a clinical decision-support tool, it might also predict exacerbation, aid preventative strategies, detect adverse events, report the efficacy of medication dosing [23], find out the causes of low-adherence behaviours, and leverage findings with targeted interventions [47].

A novel medical wheezing detection technology was developed in 2017 [61]. Instead of a single microphone, an analogue microphone array was fabricated on an insulator to detect asthma attacks. The new algorithm reduced power consumption and computing complexity of acoustic detection and recognition. The Bluetooth Low-Energy (BLE) technology was introduced as the transmitting method.

Another research in 2016 introduced a new system for intake detection [62], with recognising the individual sounds of inhaling, exhaling, inhaler actuation, and background sounds. A wireless microphone was attached to the inhaler and the data were transmitted in a smartphone through Bluetooth for processing. With the functional processor in the smartphone, complex frequency-domain calculation and distinct algorithmic approaches were achieved with 96% as the highest accuracy.

5. Respiratory measurement

Respiratory measurement has been used for health monitoring of pulmonary diseases like CF and asthma. Pressure and air-flow are measured during the research studies [63]. The existing respiratory measurement for CF are usually not in portable size, and the respiratory sensors in use might be costly. An online speaking detection system based on respiratory measurements was developed in 2011 and presented the idea of building a respiratory inductive plethysmograph chest band [64]. The duration data of inhalation and exhalation was detected and collected. The device captured respiration information rapidly whereas the samples would be analysed and recorded by a smartphone. The chest bands provided a more portable respiratory measurement which can be introduced for CF inhalation measurements.

A battery-powered portable electronic I-neb Adaptive Aerosol Delivery system was developed for CF treatment [4], which is capable to release aerosol adapting to the breathing pattern of users with monitoring. A measurement of an inhaled dose has been achieved by placing an absolute viral/bacterial inspiratory filter between the system and breath simulator. As the liquid drug was nebulised during usage, the system is not suitable for dry-powder inhalers. A single piece of this highly-priced system costs hundreds of dollars [65] and being relatively expensive.

2.2.3   The comparison of presented technologies

The electronic components for sensing, signal processing, data storage, and power supply may increase the financial cost in both research period and production stage. The inconvenience of usages may also cause frustration or even abandonment of the device. Some devices demand additional training time for patients or clinical staffs. The developed system should consider the accuracy of detection, the portability, the financial cost, the durability and the power supply, the return and transmission of data, and the presentation and analysis of data.

1. The accuracy of adherence detection

The accuracy of motion detection stays uncertain in terms of whether the patient has completed efficacious medicine administrations. The electromagnetic tablet-dispenser [38] and AdhereTech [47] detect whether the container is opened and record the date and time of the lid openings. There is no accurate information of whether the patients have taken the medicine or not and there is a possibility that the patient forgets to put the lid back on the container. SmartTurbo [48] shows a similar problem, which uses single-wrist motions to simulate inhalations. There is no evidence for ensuring the inhalation has occurred properly and effectively.

The visual detection [49–51] is proven to be reliable as a large amount of information being recorded and considered. With a high-quality static camera, the technologies of face detection, template matching, and hand localisation have been used for recognition of in-take movements, ensuring the right user is taking a correct dose of the right medication within the required time period.

The researches that detect and demarcate inhalation sound from audio recordings show an advantage on accuracy with a sensitivity of 95% and a specificity of 94% comparing the effort from human raters. With an effective sound demarcation and inhalation recognition, the duration of each inhalation can be measured and the average difference of duration between the system and human raters is around or less than 0.1 seconds [60].

2. The portability

The portability of the device affects the convenience of usages and acceptance by patients.

High-performance visual recognition usually requires high-quality cameras and a functional processing system. Therefore, they are usually settled and cannot be moved. The user needs to stand in front of the camera during every consumption of medicine to make sure their faces,

hands, or pill bottles are captured clearly by the camera. In some studies, the sensing and recognition system was bulky with a size that bigger than inhalers and declined the mobility.

The accelerometers and MEMS microphones are in smaller scales and capable to fit into portable devices. The AdhereTech [47] and the acoustic detection system for asthma [60] were both attached to the medical equipment without affecting normal usages. However, when a real-time evaluation system based on a high-performance processor, such as a smartphone application or a PC data-receiving interface, is involved, the device needs to be connected to the terminal through wireless communication or a cable. The real-time data transmission requires extra human actions, which might cause inconvenience and be forgotten or avoided intentionally by the users.

3. The financial cost

The Podhaler device and storage case are disposable and should be replaced when they are not in good condition. The price of each device needs to stay low to fit the short-term usage and to be disposable with the device. The financial cost to develop an electronic device base on the price of components and the technology of manufacture. The essential components for an adherence monitoring system include a microprocessor, sensors, solid-state memory storage, data transmitting system, and a power supply for portable devices. Additional components, such as LEDs and speakers, also increase the total cost.

Building a visual detection system is believed to be expensive as cameras are needed for video capturing. The acceleration sensors and sound sensors are more affordable and been widely used in current healthcare applications based on the development of MEMS technology. A solid-state memory chip is usually priced based on its size of storage, and a high-capacity micro SD card may cost several pounds. According to the given information from existing research studies, however, the price per device is considerable high. The electromagnetic

tablet-dispenser were at a cost 500 Australian dollars in 1999 [38]. It is reported that AdhereTech bottles would cost $1,000 per year per patient according to estimation [47].

4. The durability

The durability of an adherence monitoring device should consider: 1) whether the detection can stay accurate during a long period, 2) whether the functionality can continue on one charge of the battery, 3) whether the data can be successfully stored during the whole treatment, which is crucial for portable devices.

For most of the portable devices, power is only consumed when the device has been woken up by defined triggering signals from sensors and turned to the working mode. For example, the electronic detecting system can be turned on by the removal of the lids from containers or inhalers, which has been widely used and avoid unnecessary power consumption when the system is being idle. Supplied by a normal battery that can be purchased from the current market, most of the devices can theoretically continue to function for months or even years [38, 60, 47].

5. The return and transmission of data

For developed technologies, there are two main methods to return the data of adherence. With higher portability, raw data from activities of medication administration can be detected by the system and stored into memory [60]. The data will later be read from the memory card and received by clinics, hospitals, or pharmacies for further analysis after a treatment period. The second method declines the portability to achieve a fast or even real-time evaluation. The system is connected directly to a PC or a phone during medicine in-takes [62] and the information can be returned to clinics immediately. For the first method, the patient cannot

receive feedback in time, whereas the second method requires more coordination from users as the connection between the device and a terminal is prerequisite.

A terminal that can communicate with a database is presented for adherence evaluation. The data are usually transmitted to a PC [38, 44, 45, 60], a mobile phone [62, 64], or both of them [48]. The transmission can be through a USB cable [48, 44], a built interface box [38], or any wirelesses communication, such as Bluetooth [34, 62] or Wi-Fi.

6. Data presentation and analysis

The information of event occurrence is time-stamped. During the data analysing stage, the adherence is presented in chronological order with the date and time of recordings.

For adherence estimation based on motion detection, as there is no recorded information of accurate time length to consume a daily dose, the test protocol to present adapted patient behaviours by the researchers. In the case of SmartTurbo, each event was tagged as low-use or high-use based on the frequency of usages during the same period [48]. The monitored adherence was an estimated result with limited information about powder delivery.

For the audio recording method, the evaluation was delayed until the data were read from the memory. The audio inhalation identification of asthma inhaler recordings defined the correct using procedure of the Seretide Diskus inhaler, including mouthpiece revealing, medication releasing, an event of inhalation, 10-second breath holding, and turning off the device. The sound clips of different events were recorded previously for a training procedure [60]. Based on the subsequent evaluation, the common technique error during treatment was detected and revealed, such as exhaling into the mouthpiece, no inhalation, low inspiratory flow rate, and multiple inhalations [23]. Adherence to medicine was acquired over 3 months in the later clinical trial (NCT01529697), it was proven that the remotely monitored adherence

holds important clinical information. However, during the 3-month treatment, the users had received no feedback of their technique error until the trial finished [23].

Discussion

With the importance of monitored adherence has been indicated, it was also claimed that the limitation of electronic monitoring exists despite high accuracy and objectivity, and a robust adherence measurement is still in the absence. The comprehensive review of existing technologies lists the strength and weakness of each monitoring method.

It is believed that the visual recognising system is able to collect important information for identification, whereas the construction may lead to high financial cost, high power consumption, bulky equipment, or even low acceptance by users. Comparing to the visual capturing system, the acceleration sensors and sound sensors are manufactured in a smaller size with lower price and lower power consumption. For motion detection, the adherence is estimated from human movements during usages, including lid-removal and rotations of inhalers. The information of whether medicine has been consumed and whether it was delivered properly stays insufficient, which declines its accuracy for inhalation monitoring. Sound detection based on a high-performance processing system is proven to be more accurate in terms of medicine delivery, as the extracted features from acoustic analysis relate to respiratory rate.

A self-evaluation tool is still in need as the adherence and information of inhalation are only recorded and stored without any feedback given in most cases. The users receive limited knowledge about whether their adherence has been recorded correctly.

2.3   Summary

The solution of adherence monitoring has been developing in many med ical fields and the methods based on traditional diaries and electronic system have been proposed in the literature for detecting, recording, and analysing administering of medicine. The electronic devices are recommended as the advanced technologies for medical monitoring thanks to their objectivity, accuracy, and the possibility of employing a fast-reacted evaluation system. In this chapter, the existing popular monitoring solutions have been classified and described briefly.

The technologies of sound detection, motion detection, and vision detection have been discussed. Due to the characteristic of respiratory diseases, the quality of inhalation is the main target for detection, therefore the sound detection is selected with the assistance of a motion detecting system. The motion detection has been used to switch the device from the low-power idle mode to the full-power working mode and thus achieve power saving. The acoustic analysis and detection of inhalation are presented and evaluated in the next chapter, where the methodologies based on frequency-domain features and time-domain features will be estimated separately.

Chapter 3

Acoustic Inhalation Recognition

This chapter briefly introduces different technologies of sound recognition with a further comparison of their performance and computational complexity. It reviews time-domain, frequency-domain, wavelet, envelope, and deep-learning analysis of sound from human activities, especially from the respiratory system, the feasibility of using sound detection to achieve adherence monitoring has been discussed. It presents a comprehensive acoustic analysis of inhalation with a further environment evaluation. A soundproof room has been used for data collection in the early stage.

According to the performance of different features, a combination of three features, including the ZCR, a new feature named short-time Magnitude Difference (STMD), and the pulse length (PL), have been selected to fit the real-time processing. This novel combination has high accuracy in terms of detected rate and duration measurement with low computational complexity and memory requirement, which can fit into a small-scale electronic system with only 1 KiB RAM. The duration difference between the measured duration and reference can be limited to milliseconds based on the fitted threshold value of selected features.

3.1 Introduction

Acoustic recognition has been widely applied in healthcare engineering as a non-intrusive monitoring technology [54, 58]. Especially, for respiratory disease, as the sound of inhalation is related to the quality of medicine administering, acoustic recognition has been successfully established to detect and evaluate medical treatment remotely and is proven to be cost-effective and accurate in its performance [60, 62]. It is proven to be able to identify inhalation, exhalation, and the pulse between the events by analysing extracted acoustic features, thus providing an evaluation of correct inhaler usage. Considered CF as the target disease, acoustic recognition is selected.

The main processing target of the research is to recognise inhaling sounds during medicine administering and measure the duration of inhaling events using the limited 1 KiB random memory to handle data buffers derived from the audio stream. The differences between inhalation and non-inhalation demarcate a sound clip of actual administering. The sound of inhalation is considered to be stationary and continuous during time-limited frames, which is similar to miscellaneous common broadband environmental sounds, such as white noise. With the acoustic complexity of the environment, it is always challenging to discriminate target sound from background noise and locate boundaries of segments on audio files. When the inhalation cannot reach a certain amplitude, especially for users with restricted lung functions or aged patients that cannot inhale strong enough, the sound of inhaling is easier to be affected by surrounding broadband noises.

The duration of inhalation will be measured by detecting the discontinuities of inhalation and locating its endpoints. Endpoint detection is an early stage of sound identification. It contributes to healthcare monitoring and the pre-processing of audio processing technologies like speaker identification and word recognition, which has been widely used to identify

acoustic changes. It plays an essential role in sound recognising and removing mute segments before speech processing with the potential to demarcate inhalation.

For a portable device, low-power computing can contribute to saving energy and extending battery life for a duration-fixed treatment. Alongside with dynamic power management and server virtualisation, improving the efficiency of the algorithm is one of the general approaches to reduce power consumption [66], which can save electrical power by reducing the number of computing cycles and the required hardware storage for calculation. The selection of features needs to consider computational complexity and it may aid to decrease the needed number of coefficients and increase the efficiency of the algorithm.

3.2 Feature extraction and deep-learning methods

3.2.1 Time domain features

The very early version of endpoint detection was proposed for locating endpoints of an utterance processing in 1975 [67], based on two parametrical controlled spectral analysis systems [68, 69], which is inherently capable of performing correctly in any reasonable acoustic environment. The simple efficient processing was considered as one goal of endpoint detection alongside with the reliable locating capability and the adaptability to various environment sounds [67]. The algorithm was based on two features: short-time energy (STE) [70] and zero-crossing rate (ZCR), which are fast to calculate and the most commonly applied feature to form temporal feature sets in endpoint detection [71, 72]. In the early 80s, the improved endpoint detection for isolated word recognition [73] was proposed and composed a concept of the explicit, implicit and a comprehensive combination hybrid endpoint detector, depending on whether the endpoint locating was separated from and prior to the recognition and decision stage. With only the equalised energy was used as the key feature, an energy

pulse detection and the pulse endpoint ordering system was introduced. It also pointed out that in the real-time identification, the start point of the target sound must be detected before it ended in order to use a smaller processing buffer[73].

When the signal is presented as *x(n)*, *n* means the *n*th frame and *N* means the total frame shift number or the length of the signal, the equations of different features are described below:

1. Amplitude-based features

The most common and easy-implemented features in the time domain are the mean value and the peak value. The mean of signal values over the sample length. When μ is the mean value of the signal, it can be calculated as:

$$\mu = \frac{1}{N} \sum_{n=1}^{N} (x_n) \tag{3.1}$$

The peak value represents the maximum amplitude of a signal, it can achieve a rough sound estimation or change detection. The peaks above a certain threshold can also be counted to roughly present the volume of a signal. Peak values are very sensitive to random impulses and become less promising for endpoint detection, which requires an averaging process.

2. Energy-based features

Energy has always been a key factor in endpoint detection. For the audio signals, *E(n)* designates the short-time energy of signal [74]:

$$E(n) = \sum_{n=N-L+1}^{N} (x[n]w[N-n])^2 \tag{3.2}$$

where *L* is window length of the processing window *w[N-n]*.

The root mean square (RMS) value is a simple measure of energy content for signal detection and it is widely used in fundamental signal analysis. It can be defined as:

$$RMS(n) = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (x[n])^2} \tag{3.3}$$

3. Zero-based features

The ZCR is the rate of sign-changes along a set of sample data, calculating how many times the signal changes from negative to positive or conversely. It has been used widely in the existing audio process, including speech recognition, audio information retrieval, and music analysis, being a key and basic factor to identify percussive or continuous sounds in this way [75].

ZCR is defined formally as

$$Z_n = \sum_{n}^{N-1} |sgn(x[n]) - sgn(x[n-1])| w[N-n] \tag{3.4}$$

where

$$sgn[(n)] = \begin{cases} = 1, & x(n) \geq 0 \\ = -1, & x(n) < 0 \end{cases}$$

and $w[n]$ is the windowing function with a window size of L samples

$$w[n] = \begin{cases} = \frac{1}{2L}, & 0 \leq n \leq L-1 \\ = 0, & otherwise \end{cases}$$

As a time-domain feature, the ZCR also presents signal characteristics in the frequency domain and is used commonly to frequency or period measuring, especially for periodic signals [76]. A theorem was introduced to show the relationship between the ZCR and

Fourier Coefficients for signals in a limited range [77, 78]. The ZCR also contributes to pitch detection alongside with Autocorrelation Function (ACF) for fundamental frequency characteristic extraction, whereas the lack of its detecting accurateness for noisy signals and harmonic signals has been mentioned [79].

Based on the zero-crossing detection, other zero-based features, such as the sign-turning points, and the duration between zeros, have presented time-varying signals [80]. A coding concept Time-Encoded Signal Processing and Recognition (TESPAR) was generated in 1999 [81] with the original data being divided into small segments by zero-crossings. It is coded based on the duration and the wave shape between real zeros. It has been widely used in specific simple and steady sound recognition technology. The CARCODE was developed for real-time condition monitoring of a pressure control system on the basis of TESPAR with the peak amplitude and the turning points involved [80].

The peak to peak period detection can also be used to extract fundamental frequency characteristics. The direct time domain fundamental frequency estimation (DFE) [82] provides the classification criterion with the adaptive peak to peak detection after a pre-processed spectral shaping, which needs lower computation complexity and fits especially in a real-time sound identification system.

4. Variation-based features

Variation-based features are common statistical time-domain feature extractions, and the most widely-used variation-based feature is the variance. The variance of a signal is the expected value of the squared deviation from its mean μ. For a random variable $x[n]$, var is presented as its variance.

$$var = \frac{1}{N-1} \sum_{n=1}^{N} (x_n - \mu)^2 \qquad (3.5)$$

For further measurement of the rule how samples change, deviation of the variation is presented. The factor σ represents the standard deviation in the following equation.

$$\sigma(n) = \sqrt{\frac{1}{N-1} \sum_{n=1}^{N} (x[n] - \mu)^2} \tag{3.6}$$

Skewness is asymmetry in a statistical distribution by measuring the third order cumulative. It can be calculated by the equation below:

$$skew = \frac{\frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)^3}{\sigma^3} \tag{3.7}$$

Kurtosis is in a similar way to skewness, which is another descriptor of the shape of the probability distribution of a real-valued variable by measuring its fourth order cumulative. kurtosis characterises the flatness of distribution and has been successfully applied in automated identification of respiratory sounds [83] and lung sounds [84]. Combined to ensemble empirical mode decomposition, it can be used for sound segmentation and extraction in the healthcare system [85]. It can be defined as:

$$kurt = \frac{\frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)^4}{\sigma^4} \tag{3.8}$$

ACF and Cross-Correlation Function (CCF) have been proved to be comparatively robust in pitch detection against noise. Correlation analysis is a common wave-analyse method in time domain. Considering deterministic discrete-time signals *x[n]* and τ as the lag index number, the ACF is defined as:

$$R(\tau) = \sum_{n=-\infty}^{\infty} x(n) \cdot x(n + \tau) \tag{3.9}$$

Average Magnitude Difference Function (AMDF) was introduced in 1974 as a variation of ACF for pitch detection in speech recognition [86] and was usually weighted and combined with ACF for advanced robust pitch detection [87]. In later researches, the falling trend of AMDF is avoided to compose the high-resolution AMDF (HRAMDF) [88], the circular AMDF (CAMDF) [89], the aligned AMDF [90] and the Extended AMDF Extended AMDF (EAMDF) [91], which efficiently reduce detection errors and increase precision. The multiplication of ACF and AMDF in addition to a functional bandpass filter [92] are presented to form a high-resolution pitch detection.

The original AMDF is defined as [86]:

$$D(\tau) = \frac{1}{N - \tau - 1} \sum_{n=0}^{N-\tau-1} |x(n) - x(n + \tau)| \tag{3.10}$$

Inspired by the concept of AMDF and mean absolute difference, which measures statistical dispersion between two independent variables drawn from a probability distribution by calculating the average absolute difference between them, this thesis introduces the short-time Magnitude Difference (STMD). STMD measures the difference between two continuous samples of a variable, it can be described as:

$$STMD(n) = \frac{1}{N - 1} \sum_{n=1}^{N} |x[n] - x[n - 1]| \tag{3.11}$$

3.2.2 Frequency domain features

With the development of frequency-domain feature extraction, more sophisticated processing technologies have been composed in decades. Frequency domain features can demonstrate more detailed characteristics, whereas the calculation complexity may also increase especially when the matrix is extended.

There are various methods to extract frequency domain characteristics. The traditional fast Fourier transform (FFT) uses periodic signals in the individual frequencies to present the target signal, which shows an iconic insight and majorly contributes to transforming the time-domain signals into its spectrum representation. The spectral features were less susceptible to be affected by the changing environment noise.

Despite that the information of signals that varies with time will be lost, short-time Fourier transform (STFT) is defined for improving time-frequency analysis. It usually presents or plots the changing spectra against time and shows how the frequency characteristics change linearly. Instantaneous frequency (IF) [93] obtained from the STFT spectrum with a time-warping method and presents sharper harmonic structures of quasiperiodic signals.

Mel-frequency cepstrum is one of the most popular methods to pick up information from sample data based on the linear cosine transform on a nonlinear Mel-frequency scale [94]. It is a representation of the short-term power signal spectrum. According to existing references and progress, frequency warping using Mel-frequency cepstral coefficients (MFCC) can allow for better representation in audio compression. The MFCC incorporates biologically inspired characteristics and relates to human auditory habits. In 2012, new research combining basic acoustic parameters with MFCC in the segmentation of continuous speech among resonant or obstruent environment [95], with some acoustic parameter could not balance robustness to noise with intrinsic stability. With the compensation of channel distortion, MFCC is proved to be one of those successful features in sound recognition. The original signal is pre-emphasised, framed, added windows to give out information in the time domain, and then leaves through the discrete Fourier transform its discrete spectrum [96]. The Gammatone Frequency Cepstral Coefficients (GFCC) [97] also shows promising performance in sound recognition.

Linear predictive coding (LPC) is an encoding method to represent the spectral envelope from the target signal, which extracts information from the previous samples and determines

a model of weighted coefficients to make a forward linear prediction [98]. To achieve better extraction of acoustic parameters, the prediction error filter needs to be minimised and the target LPC coefficients could be found by using Levinson-Durbin recursion [99]. The LPC is usually used as an autoregressive process in audio processing for sound synthesis, speaker verification, word recognition, and other sound detection technology. Line Spectral Pair (LSP) frequencies are uncorrelated pairs derived from LPC, which have optimal theoretical statistical properties for scalar and vector quantisation [100]. The LSP is proved to be more explicit for sound classification and more robust in a noisy environment [101]. The Perceptual Linear Predictive (PLP) [102] was described in 1989 as a development of linear predictive analysis with engineering approximations of resolution curves and auditory relations to be more consistent with the human hearing habit.

Conventionally delta and double-delta cepstral features are appended in current automatic speech recognition (ASR) systems [103]. Delta-cepstral features were proposed to add dynamic information to existing static frequency features, which can also improve identification accuracy. Delta cepstrum may improve characterisation quality of temporal dependencies to the Hidden Markov Model (HMM) frames, which are assumed statistically independent.

On the basis of the spectral analysis, features with simple logic contribute to the transformed signal. Spectrogram row self-correlation [104] is constructed on spatial frequency based on the knowledge of image processing. The sub-band spectral selection method provides a more robust detection system against various miscellaneous environment noise [105].

$$H(m) = -\sum_{m=1}^{N} p_m \log p_m \tag{3.12}$$

p_m presents the corresponding probability density. N counts the quantity of frequency components from FFT. The spectral entropy-based algorithm is related to the energy variation

instead of the amount. As the spectral entropy is more reliable under non-stationary noises like mechanical sounds than pure energy-based detection, but it shows deficiency under babble noise and melodic noise, the entropy and energy are unusually combined to form a more tolerable feature for detection under different noises [106].

In most cases, the change rate or difference of spectrum energy between sub-bands are measured and analysed. The concept of discrimination information replaces the traditional probability distribution with sub-band energy distribution improves system efficiency [107].

3.2.3   Wavelet transform

The STFT presents both time-domain and frequency-domain characteristics in terms of sine and cosine functions, whereas the time interval is fixed [74]. The wavelet transform can be understood as a complement to the conventional Fourier decomposition and suits better to non-periodic broadband signals even with sinusoidal or impulse transients [108, 109]. Wavelet transform introduces the varied-sized processing window for providing detailed decomposition and reconstruction information [110] with more capability for the extraction of transient features. Wavelets are a widely-used statistical method and contribute in different areas, including signal processing, data compression, DNA analysis, and speech recognition [108].

It can be described as an oscillation which decays by the time [108] and expresses over the entire spectrum of various scales. For the detection and analysis of abrupt changes, the wavelet transform provides simultaneous localisation in both time-domain and frequency-domain. Wavelets have an adaptable window that fits the signal automatically and provides optimal resolution [109]. The wavelet transform is defined as:

$$WT(a,b) = \int_{-\infty}^{+\infty} f(t)\overline{\Psi_{a,b}}(t)dt \qquad (3.13)$$

$$\Psi_{a,b} = |a|^{-\frac{1}{2}} \Psi(\frac{t-b}{a}) \tag{3.14}$$

where a is the scale, b is the time, and $\psi(t)$ is the mother wavelet.

Wavelet theory has been proved to be a promising method for non-stationary signal detection with time-width being adapted to the signal frequency [108]. Recent research studies have described wavelets as an established investigative technique for detection and quantification with high flexibility and relatively efficient. Wavelet packet is an extension of wavelet transform with the approximation coefficients while details in certain levels are decomposed to create the binary tree [74].

3.2.4 Envelope analysis

The technology of envelope extraction has been developed in early radio communication. It can be considered as the amplitude modulation, which is a simple signal presentation in the time domain. Envelope analysis of breathing sounds is considered as a reliable method for feature extraction in classification and recognition of respiratory diseases. It is associated with the pathophysiological relevance of respiratory sound [111] and contributes to healthcare science, such as heart sound analysis. The extraction of amplitude envelope from waveform has also been introduced into cognitive studies of speech to detect and remove breath sounds. With a Hilbert transform and a full-wave rectification [112], the gradient between adjacent samples of the extracted envelope is measured and compared with the gradient thresholds of the trained speeches.

The extraction of the envelope is a perceptual equivalent to the assessment of the amplitude modulation and can be conceived to evaluate inhaling events. The differences between the sound waveform and converted waveform can be used to extract information of the envelope. The coefficient of variation of the envelope is defined as the quotient between the standard

deviation and the mean of the envelope of a signal [113], which describes the information of the envelope profiles and ascertains the differences between them. The envelope extraction can be analytically performed by the non-linear Hilbert-Huang Transform as an improved solution of the original Hilbert Transform, which may be affected by high-order harmonic with high-frequency rough glitches in the extracted result. The result of envelope extraction depends on the carrier frequency band, which remains unknown when the processing initially occurs. The target signal in those frequencies should be the majority. Thus, the prior analysis of the signal before filtering is necessary to improve the envelope extraction.

To extract envelope, the signal will be pre-processed with noisy and mute clips removed first. All the local extrema of the target sample are identified in the next stage. The local maximum and minimum are connected by the parabola parametric spline interpolation line to obtain the upper and lower envelope. The mean line of the sample is defined by the median value between the maxima and the minima. The difference between samples and the mean value is the first component $h(t)$, which should be defined as the demand of an Intrinsic Mode Function (IMF). $h(t)$ should be symmetric and distributes in both positive and negative areas across the x-axis.

As the breathing sounds were measured to be from 100 Hz to 1200 Hz [111], only the IMF with instantaneous frequency in this range will be retained. By extracting the envelope of an instantaneous signal by Hilbert Transform, a result with only positive frequency can be used to express the real part of the original sample.

A simple envelope detector can be constructed by diodes or a rectifier with a fed low-pass filter [114]. After the absolute value of the real and imaginary parts of a complex signal was summed, a first-order low-pass IIR filter can be used to build a signal envelope detection scheme. With the feedback coefficient α, which is valued from 0 to 1, the envelope is proportional to the instantaneous magnitude of the original complex signal [114]. Without squaring calculations, the computational complexity of this envelope detector stays low and

fits into the real-time processing, which has been introduced and developed for environment estimation in this research.

3.2.5 Machine learning models

Distance-based segmentation [115] and model-based detection [116] are two main categories of detection approaches. Compared to the model-decoding-based detection, distance-based segmentation defines a distance measurement initially with an endpoint detection strategy without the basic characteristics of input contents prior to the detection.

Machine learning algorithms calculate and define the prior sets at the data training stage [117]. The knowledge of probabilistic mathematical models is introduced to approximate the optimum of feature combinations. At the discrimination stage, the feature sequences are extracted from target signals and compared with the defined threshold value. Machine learning algorithms contribute to both audio processing and healthcare monitoring technology [56].

With the simple algorithmic structure, clear performance and high recognition accuracy, the Gaussian Mixture Model (GMM) [118] and the HMM [119] have been widely used in sound recognition system in decades. Feature behaviour can be described by the relative weights and the covariance parameters of the models [120].

The Support Vector Machine (SVM) builds mostly non-probabilistic binary linear classifiers and separates different categories with large margins. To achieve a better data implementation, the samples in higher dimensional feature spaces can be implicitly introduced only by using an efficient kernel mapping [121]. The learning process for constructing an SVM is more efficient than previous learning methods, which can provide a unique optimised solution [122]. It has been a promising learning technology for classification and widely applied in the developed learning systems with the ability of generalisation.

Fuzzy modelling describes the signal characteristics based on fuzzy inference rules and is applied to represent complex nonlinear relations of the system [123]. A fuzzy neural network (FNN) can handle the uncertainty from data and has been successfully used in signal processing and pattern classification [124].

The boosting technology combines different classifiers to construct a more accurate solution [121] and solves the two-class classification problem efficiently. AdaBoost, which adapts to weak hypotheses, has been the most widely used boosting algorithm [125] and showed good performance in inhaler medication adherence detection of a pMDI for asthma [62].

More state-of-the-art deep learning models, including the Deep Neural Network (DNN), Recurrent Neural Network (RNN) and i-vector methods [126], have contributed to sound labelling. The clustering and identification technologies like the k-nearest neighbours' algorithm and the Bayesian information criterion (BIC) agglomerative clustering [127], have been highly mentioned and widely used in audio classification.

3.3   Feature selection

Both frequency domain features and time domain features are widely used in sound detection and identification. Simple time domain features are more popular in event detection or segmentation, and frequency domain features show an advance performance in speaker characterisation or further audio extraction.

To select an optimal feature-set to detect and measure inhalation, a novel detecting method aimed to employ fewer features to represent target inhalation with high discriminative capability and reduced computational cost [128] was built in this research. Samples of inhalation and exhalation were collected in the soundproof room and extracted on the PC-based sound detecting system. The relative merits of feature candidates were analysed

in terms of their performance on signal presentation, computational load, and memory requirements.

3.3.1   The complexity of features

Computational complexity is one of the most prominent aspects of a processing algorithm of mathematical functions, which can be very high and unbounded for a Fourier Transform algorithm. A high-performance signal processor allows pipelining which dedicated for certain time-critical functions, such as multiplication and the FFT. The complexity of computing the DFT is $O(n^2)$. An FFT factorising the DFT matrix into a product of sparse and thus reduces the complexity to $O(n \log n)$ of the fast implementation [129]. A wavelet transform is based on FFT with similar or even higher computation complexity. For example, the 2D Haar discrete wavelet transform matrix of calculations is similar to the 2D Fast Fourier Transform with a complexity of $O(4N^2 log2N)$ [25], which is relatively high and fits more in a high-performance processor, such as a PC or a mobile phone. The MFCC is also more common to be developed on a high-performance processing system.

In the execution of feature extraction or other stages in signal processing, the computing speed of the general processor may decelerate with depleting the battery, especially when the exhaustive features are involved. Time domain features refer to the variation of amplitude and show the fundamental acoustic characteristic. They are temporal directly measured or estimated from the original signal with a simple calculation, such as the basic addition or multiplication. The time-domain features, including the ZCR, the STMD, sound volume, deviation, the STE, and the ACF, are simple for calculation with no exponentiation or FFT operation.

An optimal feature selection approach for wearable monitoring systems was proposed and embodied information regarding feature selection, relevance, and redundancy [130]. To opti-

mise the power consumption of feature extraction, the research considered the computation complexity of the selected features, including signal amplitude, deviation, median value, and root mean square power among time-domain features, which clarified that the start to end value was the most low-energy feature and the median value cost most power. It also discussed how the energy was constrained to prevent feature-set from being accounted for during motion recognition. The practical energy consumption experiment of the feature extraction was estimated on a Texas Instruments (TI) MSP430 microcontroller.

To measure the deviation of a signal, the mean value of the processing window needs to be calculated as the difference between samples and the mean value will be stored in the system, which means the data in one window needs to be processed twice with extended calculation time. To measure the ZCR, the STMD, and the STE, each sample only needs to be processed once. The complexity to measure STMD is lower than STE, as the STE is based on the multiplication of two neighbouring samples and the STMD is based on the difference between two neighbouring samples. The STMD can be fulfilled by basic subtraction of bytes.

As only two neighbouring samples are taken into account for calculating the ZCR, the STMD, and the STE, only two integers are needed, which takes only 2 bytes. Besides, counters for summing up the integers during a full window are needed, which will be one byte for each feature. For deviation, one more byte will be used to record the mean value. Overall, the simple non-multiplicative features based on neighbouring samples can be calculated using 3 to 5 bytes, which saves the memory space massively.

For algorithm development, the memory requirement is also important for processing performance and account for the computation time. For evaluating FFT algorithms and wavelet extractions, a large memory space is needed for matrix calculation, especially for higher order functions [26].

3.3.2   Acoustic analysis of inhalation and environment

To select the optimal combination of features, the frequency and time domain characteristics of inhaling and exhaling were extracted in the analysing stage of the research. The samples clips of usage of inhalers were collected and recorded in a soundproof room.

With similar concepts like anechoic chambers, a soundproof room was insulated from exterior sources of noise with thickened doors and designed to absorb reflections of either sound or electromagnetic waves. To decline unwanted noise, such as the low-frequency noise from electronic devices or fluorescent lamps, multiple methods like acoustic quieting, physical barriers, noise mitigation, and noise control were introduced into the soundproof systems. The participated soundproof room for this research was a Clinical Audiometric Equipment Calibration Room with extreme ambient performance. The ambient sound pressure level was lower than a leaf rustling and approximately no more than 10 dBA.

During the recording phase, the inhaler was empty without inhaling any substance. Only the sound of air-flow had been recorded. No drug delivery, powder or propellant were involved following the consideration of health and safety. The inhaling sound with powder delivered from the recorded video has been considered, which has similar broadband acoustic characteristics as the inhalation without powder. The sound volume of inhalation with powder is higher as the powder may introduce more vibration and friction, which also indicates higher energy and higher robustness for detection.

Fig. 3.1 demonstrates the waveform and spectrum of inhalations and exhalations on a cross-platform audio editor called Audacity [131], showing the diagrammatic characteristic of the signal in both the time domain and frequency domain. From the spectrum, the energy of inhalation is spread evenly in the frequency domain, which shows that the inhalation is a broadband sound with constant power spectral density from 450 Hz, sharing similarities with the white noise. The associated condition to generate inhalation with a dry powder inhaler
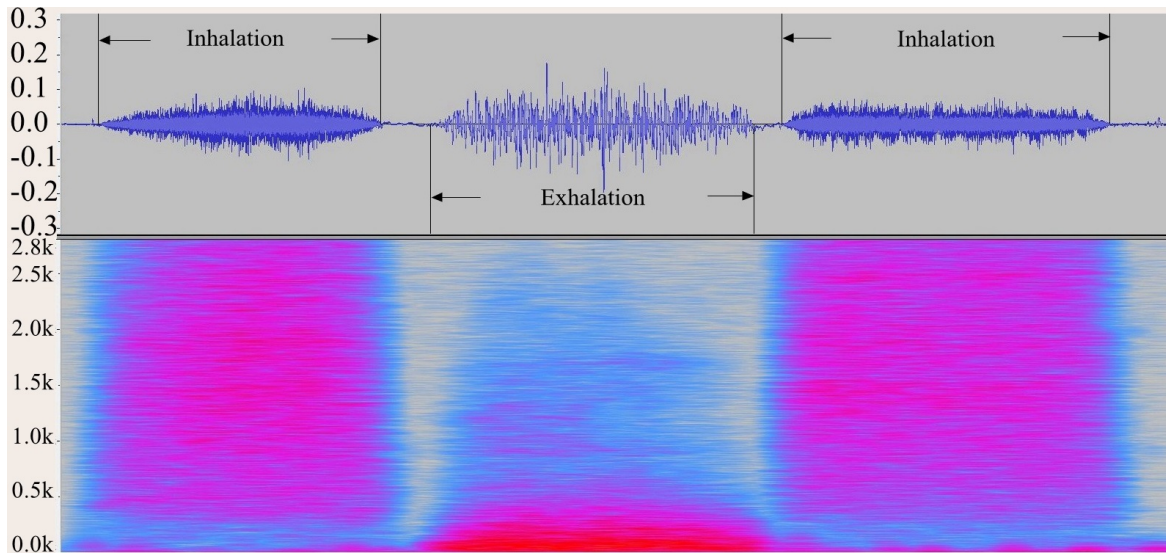
Fig. 3.1 The waveform and spectrum of inhalations and exhalations in a medicine administering

includes the airway friction or vibration of human throat, mouth, drug powder or inhaler itself, which also excites resonances as the signal is broadband with multiple formants. Receiving from multiple sound sources, the volume, and energy of inhalation is distributed over a wide frequency range at a low energy level. In contrast, the exhaling events accompanied by airflow noises are concentrated in lower frequency around 55Hz from the spectrum analysis.

Fig. 3.2 demonstrates the zoomed-in details of inhalation in the time domain, which shows that the signal crosses the zero line frequency with a high ZCR.

The wide distribution of energy impedes accurate detection through conventional spectrum analysis, which may fail to locate the peak or centre frequency of the signal. And the periodicity of a broadband signal is hard to be measured accurately. Particularly, the detection is more difficult in the presence of any sound source that shares a similarity with white noise, such as the mechanical sound or other vibration noises. The friction between the inhaler and any solid surface also incurs vibration noises. These other broadband sound sources are easy to be confused with inhalation and be tagged incorrectly.
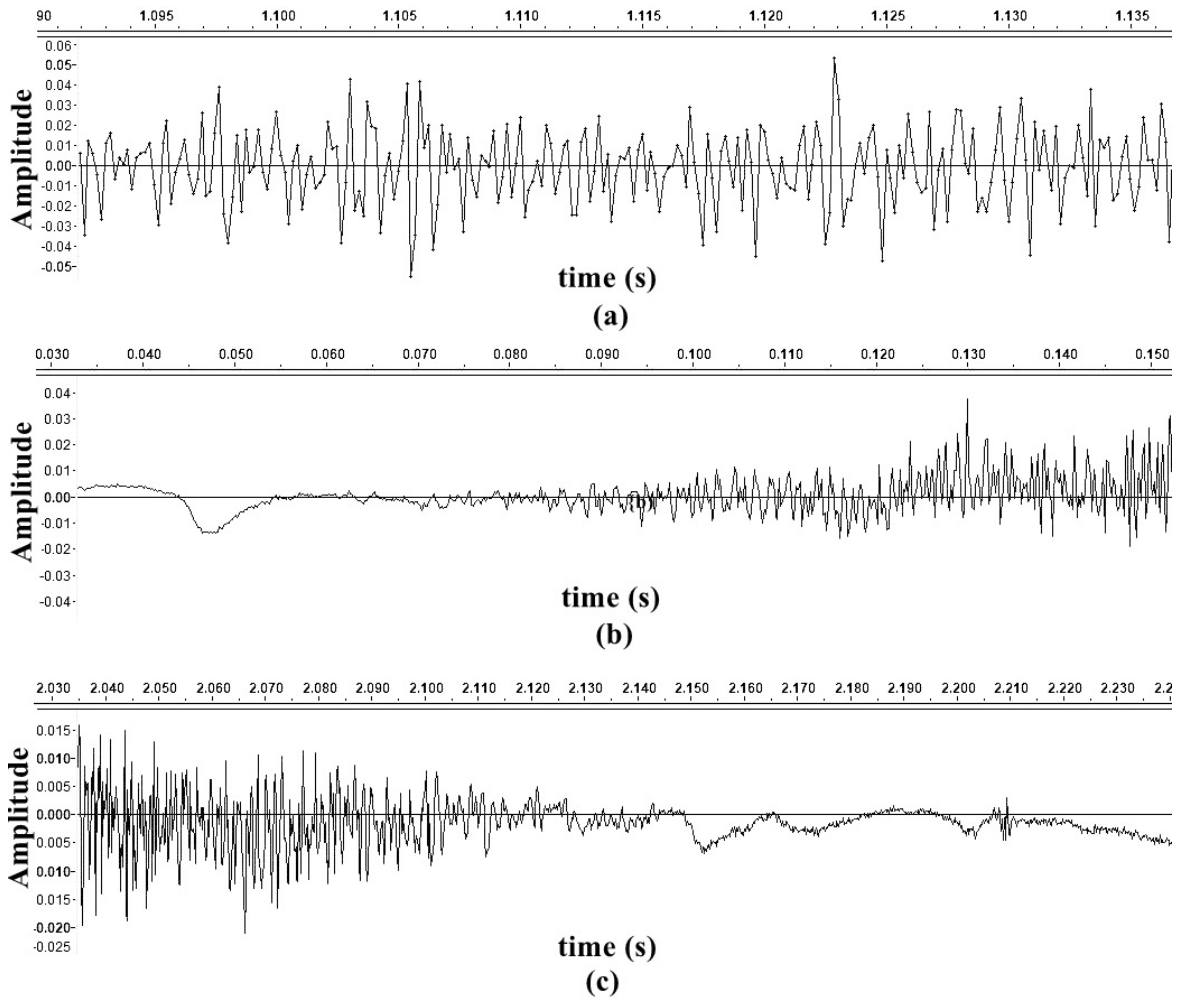
Fig. 3.2 (a) The zoomed-in time-domain details of inhalation; (b) The start of an inhalation in time domain; (c) The end of an inhalation in time-domain.
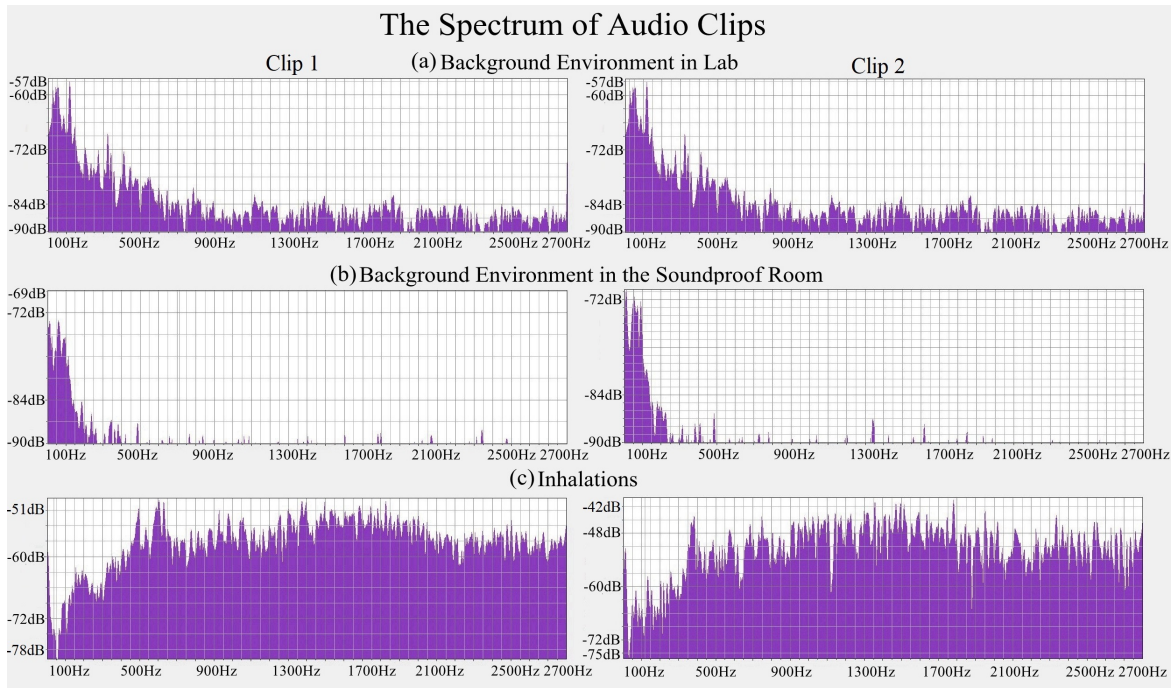
Fig. 3.3 The spectrum of the background sound in the soundproof room

The background noises have been assessed during the recording stage. The spectrum of the environmental sound is presented in Fig. 3.3. Due to the lack of volume of inhalation, the best samples that recorded inside the normal environment were still in inadequate quality, with low SNR from 12 dB to 18 dB, increasing the difficulty of detection. The noise source might be any electronic equipment in the experimental surrounding, which operated a low-volume continuous sound. The sound samples recorded inside the soundproof room provided better performance with lower noise, down to -66 decibels relative to full scale (dBFS) of sound pressure level (SPL) as the RMS value. The measured SNR of the samples from the soundproof room was 40 dB, which was also negligible for adding an exterior 10 to 20 dB artificial noise.

The frequency response of the selected microphone INMP441 is from 60 Hz to 15 kHz [132] and ideally, low-frequency noise will be not be recorded. Notwithstanding, from the spectrum in Fig. 3.4 of a recorded inhalation clip from the soundproof room, it still exhibited a significant low-frequency noise from vibration. The frequency of the artefact noise was
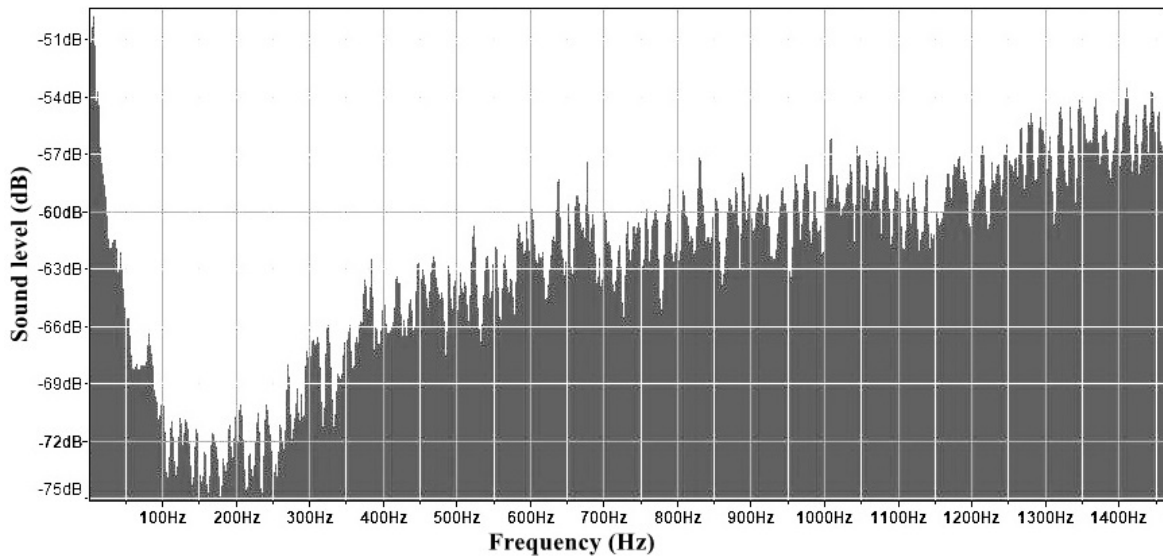
Fig. 3.4 The spectrum of inhalation and the low-frequency vibration artefact

low from 1 Hz to 13 Hz, but it stays high in energy. The low-frequency error accounted for existing single cycle deviations seen in sound traces under the possibility that the MEMS microphone device was controlled to keep the sensing element floating, which is due to an artefact from the control circuitry of the MEMS microphone. The low-frequency noises could theoretically be filtered out by a high pass filter.

3.3.3   The feature extraction

In the feature extraction stage, 200 inhalations have been recorded by a MEMS microphone and transmitted to a PC. The performance of feature candidates, including the conventional low-complexity time-domain features, the STE and ZCR, was extracted using a developed MATLAB system. The features without an obvious change of value around the start points and endpoints of inhalation were considered irrelevant.

Fig. 3.5 illustrates the STE, the ZCR, and the STMD with the original waveform of the audio clip in Fig. 3.1, which contains 2 inhalations and an exhalation. It is clarified that inhalation has high ZCR, low energy, and high STMD, whereas exhalation has high STMD,
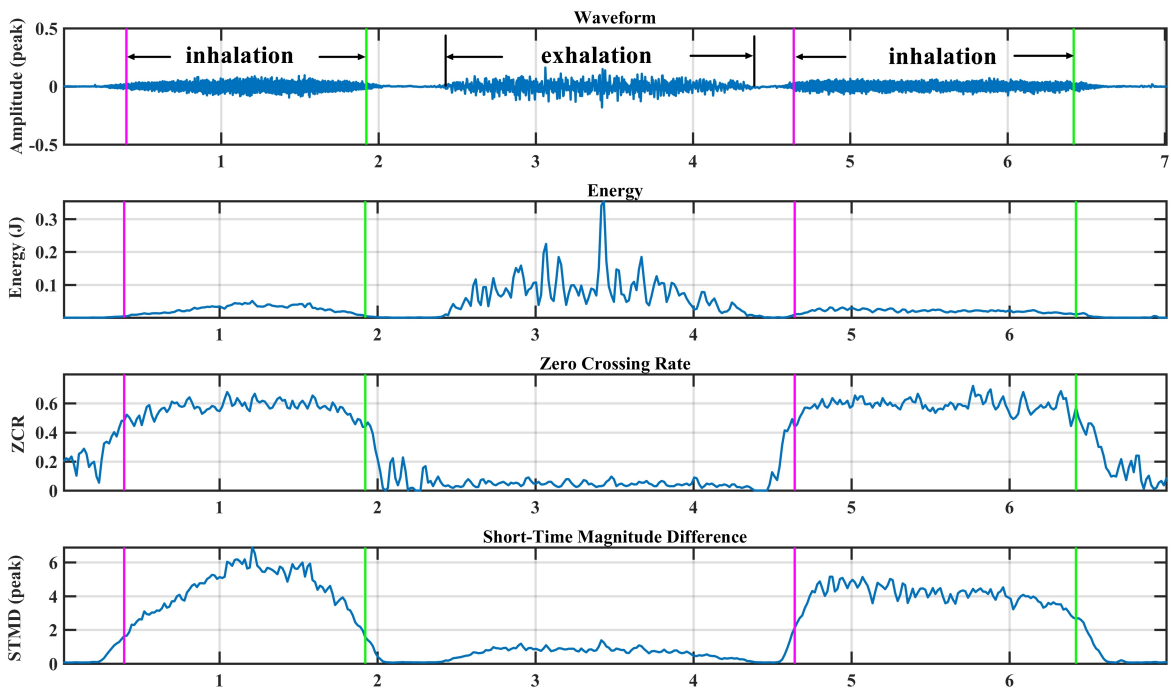
Fig. 3.5 The STE, ZCR, and STMD of inhalation

high energy, and low ZCR. According to the result, the STE will be affected by high-volume low-frequency noise, such as the exhalation sound. The STMD shows higher sensitivity to inhalation than the STE against low-frequency noises.

Another example is presented in Fig. 3.6. The audio clip was recorded in the soundproof room with 25-30 dB SNR. The 15-second audio file, which contained 4 clips of inhalation, 2 clips of exhalation, a whistling sound and a clip of vibration noise, was analysed to instantiate the identification system based on individual features.

The measured value of these features is demonstrated in Table 3.1. It is presented that the ZCR of exhalation is much lower than the inhalation while its energy is higher. The high-frequency whistling has high ZCR and low energy. The ZCR of vibration noise is low and its STE is similar to the inhalation, whereas its STMD value is lower and can be easily distinguished from the target sound. With introducing the STMD, the effect of friction or
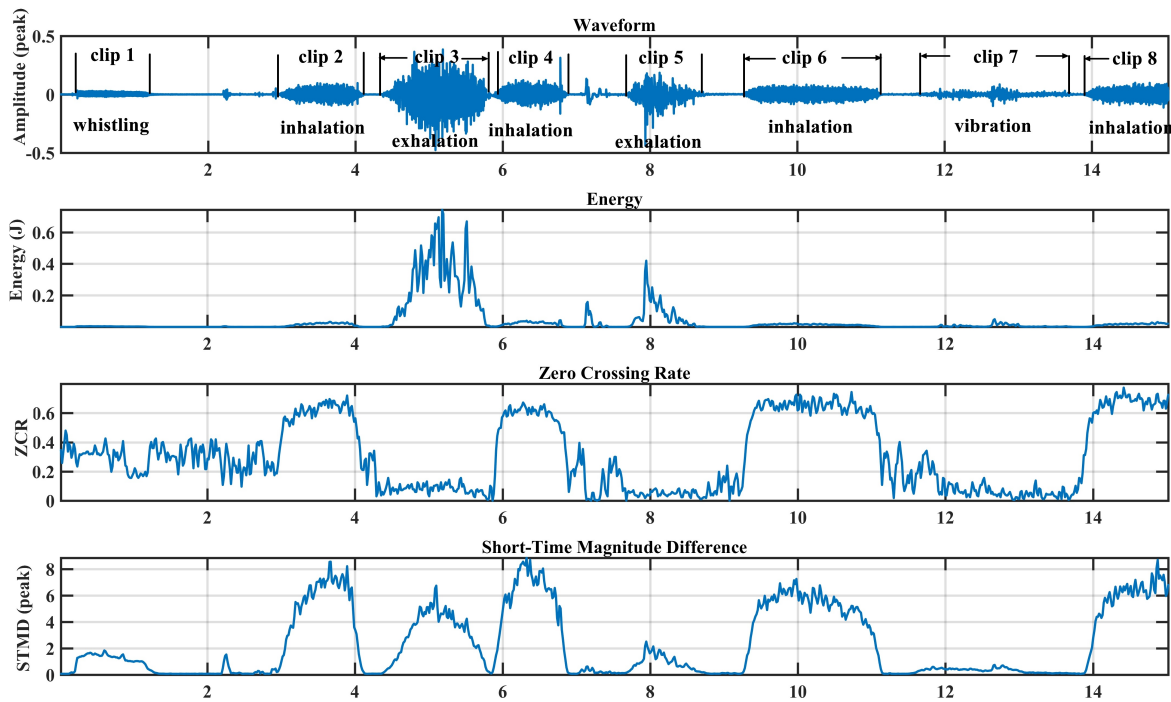
Fig. 3.6 An audio clip with inhalation and exhalation

Table 3.1 The measured feature value of sound clips

Clips	Category	Energy (J)	ZCR	STMD (peak)	Pulse Length (s)
Clip 1	whistling	0.002 - 0.005	0.16 - 0.42	1.2 - 1.8	0.83
Clip 2	inhalation	0.016 - 0.032	0.52 - 0.72	3.6 - 8.6	1.02
Clip 3	exhalation	0.06 - 0.74	0.028 - 0.16	1.5 - 6.7	1.19
Clip 4	inhalation	0.021 - 0.040	0.58 - 0.67	5.1 - 8.9	0.91
Clip 5	exhalation	0.047 - 0.42	0.024 - 0.084	0.93 - 2.5	1.05
Clip 6	inhalation	0.015 - 0.023	0.58 - 0.75	3.3 - 7.3	1.86
Clip 7	vibration	0.008 - 0.048	0.012 - 0.14	0.32 - 0.71	1.62
Clip 8	inhalation	0.01 - 0.032	0.59 - 0.77	2.8 - 8.7	1.05

vibration can be reduced, which commonly occurs during the medicine administering with impacts or scratches between a hard surface and the sound sensor.

The performance of various feature combinations is depicted with the concepts of True Positive (TP), False Negative (FN), and False Positive (FP) [60] for further evaluation. The TP presents the sensitivity of the indication system, clarifying whether inhalation has been correctly detected or neglected. The NP means the number of neglected true inhalation. The FP gauges the probability of misdetection, showing how many non-inhalations have been falsely tagged as inhalation. In other words, the FP presents the specificity of indication.

The definitions of the sensitivity and specificity are shown below:

$$Sensitivity = \frac{DetectedInhalation}{TotalInhalation} = \frac{TP}{TP+FN} \tag{3.15}$$

$$Specificity = \frac{DetectedInhalation}{TotalRecordings} = \frac{TP}{TP+FP} \tag{3.16}$$

Then, the accuracy of the system can be defined as:

$$Accuracy = Sensitivity \times Specificity \tag{3.17}$$

In Table 3.2, the system was built on different feature sets, the size of the feature set is from 1 to a combination of 5. 200 recorded inhalations from soundproof room have been demarcated manually by the experimenters. The features of the cut segments have been measured, and the value has been analysed using a normal Simulated Annealing (SA) function, which was developed in MATLAB. The threshold value of each feature was initially defined as the minimum of the feature array. Then the minimum value of each feature of each inhalation has been compared and analysed. The average minimum of the feature of the inhalations

has been defined as the threshold. For example, the average minimum of the ZCR of 200 inhalations was defined as 0.48.

During the feature selection stage, 4 simple calculated features (the ZCR, the STMD, the STE, and the ACF) and 1 measured feature (the PL) had been tested and their performance had been compared initially. The performance and accuracy were calculated based on the same 200 inhalations. Starting from single feature solutions, the ZCR showed 64.8% accuracy and selected as the optimal feature. The result of segmentation contains very short clips from several milliseconds to hundreds of milliseconds, which might be a very short sudden noise. The measured feature PL was introduced into the system. As the PL can only be presented when a feature has been calculated and a segment has been demarcated with a measured duration length, it can only be added with another existing feature. The threshold value of PL was defined as 0.6 seconds. After the PL was considered, the short noises or inhalation could be eliminated. The sensitivity dropped to 97.5% and the specificity increased rapidly to 84.6%. The highest accuracy of the 2-feature solution was the ZCR with the PL, with 82.5% as the accuracy.

With 3 candidates, there were 6 different combinations of features. The combination of the ZCR, the STMD, and the PL presented the best performance with high sensitivity (97.5%), high specificity (98.5%), and accuracy of 96%. The combination of 4 features presented 88.7% accuracy as the best performance. And with all 5 feature, the accuracy was 70.9%.

To estimate the stability of different combination of features, the 200 samples were tested under diverse noise levels from quiet to noisy, which was quantised as SNR. The natural noise that came from unwanted sources with sundry acoustic characteristics might exist on every occasion of daily life. Loud music, speeches, mechanical noises, traffic noises or weather sounds might add redundant information into inhalation and influence the identification. The result of the optimal feature set is presented in Figure 3.7.

Table 3.2 The accuracy against number of features

Number	Features	Sensitivity	Specificity	Accuracy
1	ZCR	99.0%	65.5%	64.8%
	STMD	98.5%	63.8%	62.8%
	STE	92.5%	54.3%	50.2%
	ACF	82%	56.6%	46.4%
2	ZCR + PL	97.5%	84.6%	82.5%
	STMD + PL	98%	73.2%	71.3%
	STE + PL	92%	71.3%	65.6%
	ACF + PL	78%	75.6%	59%
3	ZCR + STMD + PL	97.5%	98.5%	96%
	ZCR + STE + PL	91.5%	88.7%	80.4%
	STMD + STE + PL	91%	75.8%	69%
	ZCR + ACF + PL	75%	86.2%	66.6%
	STMD + ACF + PL	74%	84.2%	62.3%
	STE + ACF + PL	73%	82.5%	60.2%
4	ZCR + STMD + STE + PL	90%	98.5%	88.7%
	ZCR + STMD + ACF + PL	70.5%	98.5%	69.4%
	ZCR + STE + ACF + PL	68%	90.1%	61.3%
	STE + STMD + ACF + PL	69%	87.8%	60.6%
5	ZCR + STMD + STE + ACF + PL	72%	99%	71.3%
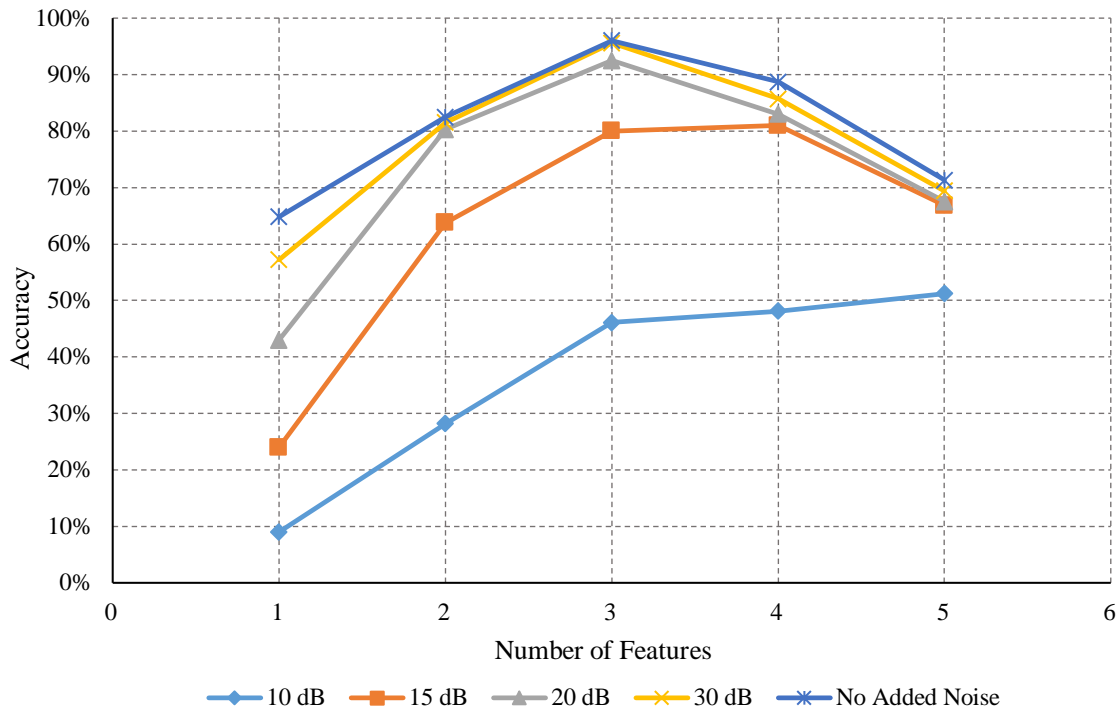
Fig. 3.7 The accuracy against feature number in different noise levels

Based on the result, the performance of feature-set is not linearly proportional to the growth of involved feature number. With a single feature, which was selected from ZCR, the STE, or other conventional features, the boundary of inhalation was demarcated roughly by detecting feature variation, especially in a quiet environment. However, the performance deteriorated when the noise level increased. The sensitivity of the single-feature solution was high as any sound with extracted value in the threshold range was detected, including inhalation and non-inhalation, and thus lead to low specificity. The ZCR was the optimal feature with a high sensitivity to inhalation and the accuracy of 64.8%.

When 2 features were involved in the detecting algorithm, the sensitivity dropped as there were two threshold ranges from different features, whereas the specificity was significantly improved. The accuracy of a combination of ZCR and STMD was 88.7% and it reached 96% when the pulse length was introduced to avoid interfere from short broadband noises like

the sound of friction between the inhaler and a hard surface. The sensitivity of the optimal 3-feature solution was 97.5% while the specificity was improved to 98.5%.

The sensitivity of the best 4-feature solution of the ZCR, STE, pulse length, and STMD continued to decline to 90% with the accuracy of 98.5%. With more limitations from added features, the sensitivity of the system decreased as inhalation might be too soft or too short to fulfil the threshold requirements. Only 72% of the inhalations were correctly detected by the 5-feature solution, whereas the detected segments were 99% correct inhalation. It was also clarified that a system based on more features was more robust under noisy environment. In Section 5.7, a LED feedback system which can recognise the noisy environment has been introduced.

From the result, it was presented that the accuracy of the system would not increase continuously when more features were involved. A multi-feature solution showed an improved performance from the single-feature solution, especially when the features were orthogonal and described the original signal from different aspects. The accuracy stayed stable or even dropped when the number of the feature was above 3 as the added features might share similarity with old features and the orthogonality between features declines. According to the accuracy, it is indicated that the STE and the STMD show high correlation with similar performance, which will be discussed in the next section. An optimal feature set may consist of 3 or 4 features to achieve high accuracy and low computational complexity. The combination of the ZCR, pulse length, and the STMD was selected for inhalation recognition.

3.3.4 The relevance and redundancy of features

The relevance and redundancy of features were analysed at this stage. Relevance analysis is a pre-process of feature selection, which can eliminate irrelevant features of sound recognition and avoid the unnecessary calculation cycles [130]. With a high correlation between these

two features, repeated or similar characteristics will be defined as redundant information. Finding the optimal feature set can reduce the measure-complexity, storage requirements, and training or utilisation times [133]. The feature reduction can also avoid over-fitting, resist noise, defy the curse of dimensionality, and strengthen prediction performance [133].

In statistics, correlation, dependence or association is used to measure the relationship between two random variables or bivariate data. The Spearman correlation and Kendall's tau, which belong to rank correlation coefficients, suits ordinal variables. Spearman correlation assesses monotonic relationships while Pearson correlation assesses linear relationships. Cramér's V is based on the Pearson chi-squared statistic and available for nominal variables [134].

The symmetric uncertainty is a measure of correlation between two discrete random variables [130], which is presented as normalised information gain and suitable for non-linear features. When $H(X)$ and $H(Y)$ represent the entropy of X and Y, their symmetric uncertainty is given by:

$$U(X,Y) = \frac{2I(X,Y)}{H(X) + H(Y)}$$ (3.18)

where $I(X, Y)$ denotes the information gain between X and Y, which is defined as:

$$I(X,Y) = H(X) - H(X|Y)$$ (3.19)

Another common simple measure of dependence between two quantities is the Pearson correlation coefficient, which is also called the product moment correlation coefficient. It simply referred to the correlation coefficient and can be obtained by dividing the covariance of two variables by the product of their standard deviations [135]. Pearson Correlation can be used to detect linear dependencies between two features, which may also be improved to

fit a non-linear function [136]. Pearson correlations are only suitable for metric variables, including dichotomous variables, and the correlation matrix is symmetric.

Pearson correlation coefficients are easy to be calculated, and the features are metric as they are measured based on fix-sized processing window. When σ_X is the standard deviation of X and σ_Y is the standard deviation of Y, the Pearson Correlation of variables X and Y is defined as:

$$\rho_{XY} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{3.20}$$

where $cov(X,Y)$ is the covariance. The covariance between two random variables X and Y is defined as the expected product of deviations from individual expected values:

$$\rho_{XY} = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]. \tag{3.21}$$

where E[X] is the expected value of X and represents the average value of X, and E (Y) represents the average value of Y. E[XY] represents the average value of the product of the corresponding elements from X and Y.

To analyse the orthogonality between the ZCR, the STMD, and the widely-used STE, 25 multi-environment sets are estimated in Fig 3.8. Each set contained 5 audio clips under the same recording environment. The audio clips of various background environments contained natural noises like random music (with the periodic percussion/bass sound and the singing voice tone), mechanical sound (broadband, loud and continuous), speeches (aperiodic, the voice spectrum of different speakers varies massively), whistling (high-frequency, continuous), and other common noises. The sound clips were recorded by the target microphone through the built prototype system. The Pearson Correlation estimation was analysed using MATLAB.

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ZCR and STE	0	32%	20%	32%	8%	4%	0	4%	0	0	0
STMD and ZCR	0	0	4%	20%	36%	12%	12%	8%	4%	4%	0
STE and STMD	0	0	0	0	0	0	0	0	0	64%	36%

Pearson Correlation

■ ZCR and STE   ■ STMD and ZCR   ■ STE and STMD
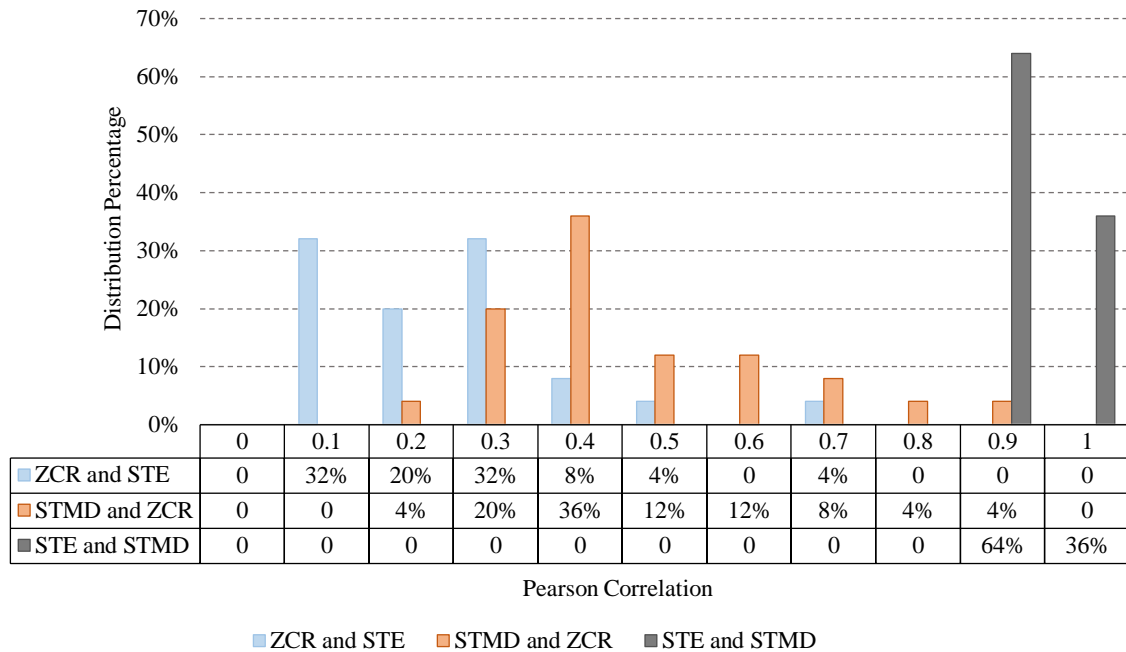
Fig. 3.8 The Pearson Correlation between feature pairs of multi-environment clips

The STE and STMD both partly presented sound volume and they were highly related to each other as the Pearson Correlation coefficients were from 0.78 to 0.96. The separation between the ZCR and STE was high and the Pearson Correlation coefficients were 0.02. The coefficients between the ZCR and the STMD widely distributed from 0.2 to 0.8, as they both described how two neighbouring sample differentiated whereas the STMD was more sensitive to the peaks of the waveform.

25 inhalations were estimated in the subsequent test and presented in Fig 3.9. Each set contained 3 inhalations, 3 exhalations and silent clips in between. The correlation coefficients between the STMD and the STE were high and stayed above 0.8. The coefficients of STMD and the ZCR were more concentrated and mostly from 0.2 to 0.4, which was relatively low.

The difference between two continuous samples was calculated and summed in a fixed-sized window. From the correlation analysis and the performance estimation, the STE and the ZCR are the most orthogonal pair whereas the ZCR and the STMD have the highest capability

	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
ZCR and STE	0	32%	44%	24%	0	0	0	0	0	0	0
STMD and ZCR	0	0	8%	56%	28%	8%	0	0	0	0	0
STE and STMD	0	0	0	0	0	0	0	0	0	68%	32%

Pearson Correlation

■ ZCR and STE   ■ STMD and ZCR   ■ STE and STMD
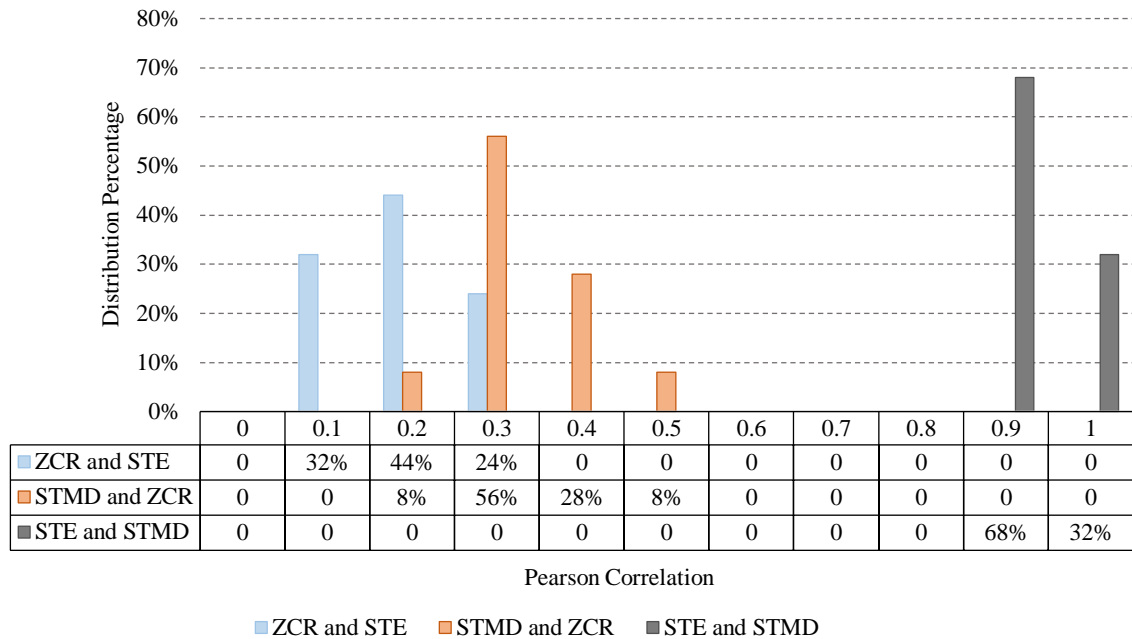
Fig. 3.9 The Pearson Correlation between feature pairs of 25 inhalations

for inhalation identification. The conventional STE shows inferior performance in boundary detection of inhalation in a noisy environment. Besides, the duration length of a single inhalation is considered as the third element to get rid of the effects of short continuous broadband signals.

As a conclusion of analysing performance and relevance, the combination of STMD and ZCR lead to sufficient specificity and sensitivity of inhalation detection theoretically on the basis of an effective noise reduction alongside with high discrepancy between them. As the STMD shows a high correlation with the STE with better performance, it is selected with the ZCR. Comparing to the STE, the STMD has higher accuracy against frequent noises in inhalation, such as the air-flow sound. Without data multiplication, its computational complexity is lower than the STE and fits more for a limited-sized processor.

3.3.5   The performance of selected features

To evaluate the performance of the selected features, a system for inhalation identification and segmentation was developed using MATLAB.

The processing window was sized as the quotient of the sampling frequency divided by 31.25, which was 179 in the research. The samples with the STMD that above threshold ratio, which was defined in the previous training session, were initially considered as a segment of inhalation. The ZCR of it was measured afterwards to locate the start points and endpoints of inhalations more precisely.

In this system, the STMD has been used to define the start points and endpoints of inhalation initially. When a segment is detected, the ZCR of it will be measured. Only the segments with ZCR that higher than the threshold value will be saved. The logic of this methodology is been presented in Table 3.3.
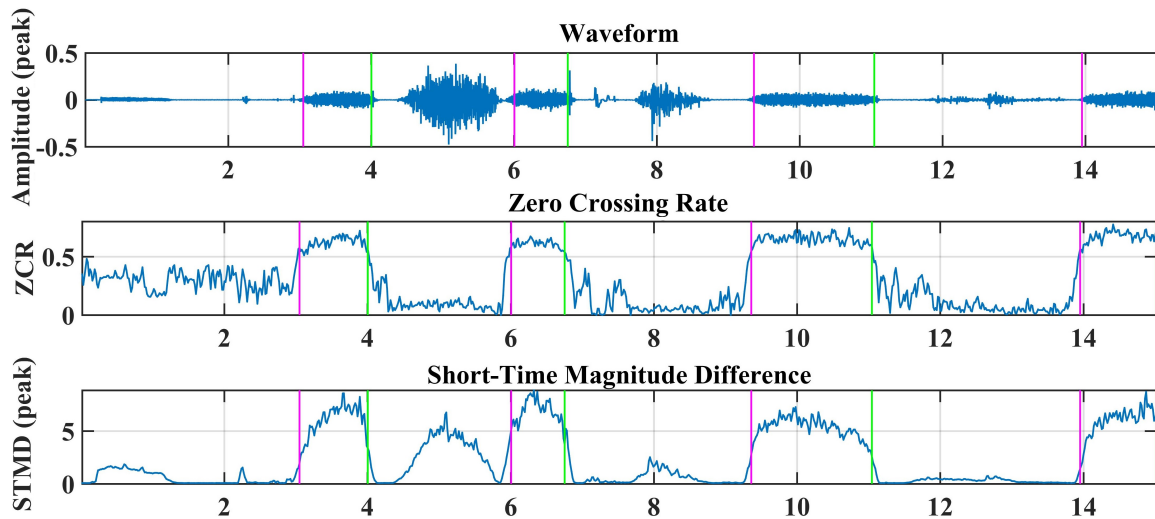
Fig. 3.10 The inhalation detection using MATLAB based on the STMD and the ZCR

Both the STMD and the ZCR contributed to the detection successively. Fig. 3.10 shows the result of the inhalation detection MATLAB. The boundaries of detected inhalation are demarcated by coloured vertical lines, where the magenta lines represent the start points and

Table 3.3 The methodology of inhalation detection using Matlab

```
//ZCR and STMD has been calculated and stored
//Initial segmentation based on STMD
//STMDTh is the defined threshold, which is 0.1 of the max of STMD
STMDIndex = find(STMD>= STMDTh)
inhalation(k).begin = amdfIndex(1);
for i=2:length(STMDIndex)-1,
if STMDIndex(i+1)-STMDIndex(i)>1,
inhalation(k).end = STMDIndex(i);
inhalation(k+1).begin = STMDIndex(i+1);
k = k+1;
end
end
inhalation(k).end = STMDIndex(end);

//Further segmentation based on the ZCR
//ZCRTh is the defined threshold, which is 0.7 of the max of ZCR
for i=1:length(inhalation),
head = inhalation(i).begin;
while (head-1)>=1 && (head-1)<=length(ZCR) && ZCR(head-1)<ZCRTh,
head=head+1;
end
inhalation(i).begin = head;
tail = inhalation(i).end;
while (tail+1)<=length(ZCR) &&(tail+1)>=1 && ZCR(tail+1)<ZCRTh,
tail=tail-1;
end
inhalation(i).end = tail;
end
```

the green lines represent the endpoints of segments. According to the result, clip 2, 4, 6, 8 were correctly marked as inhalation and the other sound clips were neglected.

To detect exhalation, the threshold value of the ZCR is set to be 0.15 of the maximum of the ZCR array, and only the segments with ZCR that lower than the threshold will be defined as exhalation and stored. The exhalation detecting system is based on the same logic as the inhalation detection.
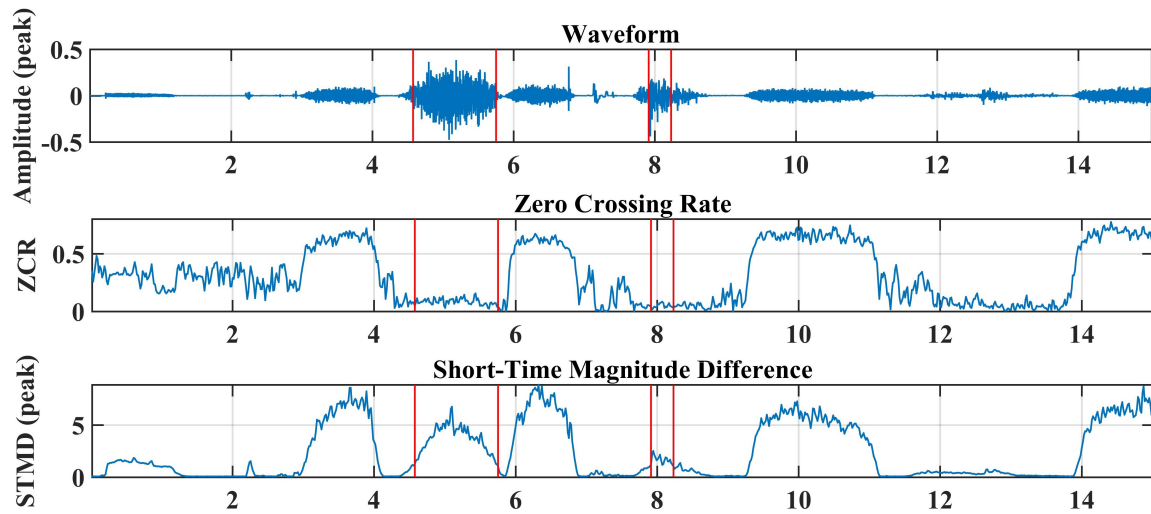
Fig. 3.11 The exhalation detection using MATLAB based on the STMD and the ZCR

Fig. 3.11 shows the simulated detection of exhalation on MATLAB. The segments of exhalation have been correctly detected and demarcated by the red vertical lines. According to the result, clip 3 and 5 were correctly marked as exhalation and the other sound clips were neglected.

The duration difference analysis of detection based on the weighted ZCR and STMD against different noise levels presents its performance and focuses on the length difference between the measured duration and the reference duration. The measured duration is the result of the developed MATLAB system. The reference duration was recorded by another microphone, analysed in Audacity and demarcated by a trained experimenter based on human-ear ability and the waveform. The duration difference was calculated by reference duration minus measured duration from on-chip storage. It can be defined as:

$$DurationDifference = ReferenceDuration - MeasuredDuration \qquad (3.22)$$
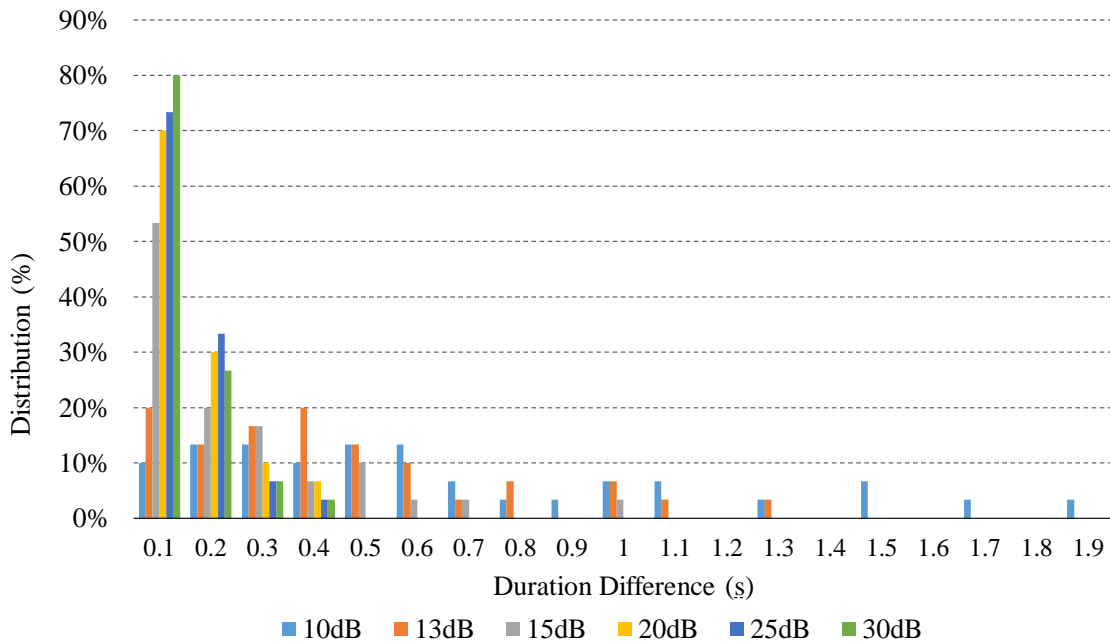
Fig. 3.12 Duration difference against noise levels

From existing labelling of usages and acquired soundproof room data, the start and end positions of one inhalation were converted to suitable formats to plot the duration error versus SNR.

Fig. 3.12 shows how duration difference increases with noise level growing. The added noise was white noise, which had high ZCR and shared similarities with inhaling sound thus affected the recognising system. A sample set of 30 inhalations has been analysed to show the distribution of duration errors. The white noises were added manually through MATLAB in addition to the original noise from the recording system.

When SNR was higher than 30 dB, the maximal length of duration difference was 103 milliseconds and the average value was 46 milliseconds. When SNR dropped to 25 dB, the duration difference was up to 163 ms while the average difference kept low at 58.5 ms. The performance stayed acceptable when the SNR was lower than 15 dB, at which level the average duration difference increased to 400 milliseconds.

The duration difference analysis focused on the detected inhalation and the length difference between measured value and the true one, whereas the redundant detected rate calculated how many sound clips were labelled wrongly as the target sound when they were actually noise clips or other similar types of sound. When the SNR was higher than 20 dB, there was no incorrect detection among more than 200 wave files. However, 23% of detected inhalations were incorrect when SNR dropped to 15 dB. The FP rate increased to 140% with 13 dB SNR and reached 250% under 10 dB SNR.

With further analysis, the redundant detected sounds were mainly short clips which were not in normal inhalation length. With ruling out small clips which were shorter than a defined duration threshold value, the rate was declined and stayed stable with no incorrect detection in 15 dB SNR, 24% in 13 dB and 67% in 10 dB, respectively. The threshold ratio was later trained and adjusted based on the SA system. The duration difference was declined by approximately 45% from the previous version, and the redundant rate was also cut down to half.

Frequency-domain features were not chosen for being energy-hungry and slow to be calculated for a small-size processing RAM. Considering the relatively limited information available to present, frequency domain features are not showing advantages for recognising the inherently broadband signals. The selection approach is aimed to pick out most orthogonal pairs which letting simple features present more information when they matched together. The insignificant characteristic in the frequency domain may lead to the futility of studying spectrum or any spectral features of inhalation.

3.4   Summary

The technologies and methodologies of sound recognition in Healthcare device in an attempt to improve adherence monitoring has been discussed and evaluated. Acoustic features have

been compared in terms of performance and computational complexity. Compared to the frequency-domain features and wavelets, the significant advances of time-domain features include faster calculation and lower requirement of calculating RAM space. Due to the broadband characteristic of respiratory sound in the frequency domain with low-STE in time-domain, and the requirements to build a fast-react real-time system, it seems more sensible to select the applicable simple time-domain features. After the analysis of various feature-set, the combination of three features, including the ZCR, the STMD, and the PL, has been selected for the real-time processing with a small-scale electronic system. Thus method gives speedier calculation and processing than the conventional acoustic analysis. The storing system and transmission system are needed after signal processing, which will be discussed in the next chapter.

Chapter 4

Data Storage and Transmission

In this chapter, an advanced approach based on NFC technology has been considered and developed for data storing and transmission. The duration of detection is saved as a highly-compressed format with real-time stamping. With a new presented coding method, the adherence of a full long-term treatment can be stored as small size data to fit the limited storage.

4.1   The data transmission

Once the data of inhalation has been written into the system, the stored data can be extracted and transmitted from the memory of the constructed prototype to a terminal at any time. For the fingertip-sized device, the conventional communication modules, such as the global system for mobile (GSM) communication, ZigBee, WiFi, and ANT+ are relatively cumbersome with the considerable higher financial cost [137].

4.1.1 The BLE and NFC technology

The BLE is a new version from basic Bluetooth for short burst data transmission [138] with reduced power consumption, which firstly appeared on the market as Bluetooth v4.0. Comparing to Bluetooth, the wireless BLE has shorter setup time [139] from around 100 ms of the basic Bluetooth to a few milliseconds. It also shows an advantage of lower financial cost, ultra-low peak, average and idle/sleep mode power consumption [140]. The device stays in the sleep mode before getting activated and is enabled to run for months or even years on limited coin-cell batteries. It has been widely used in mobile phones, wearables, healthcare, sports, and other areas. Although it has limited ability to handle a large amount of data, it is enough for this research as the inhalation only takes bytes for transmission.

NFC is a short-range contactless communication technology, which detects and enables communication in close proximity [141]. A limited amount of data can be transferred when there are a few centimetres between the devices. It operates within the unlicensed radio frequency industrial, scientific and medical (ISM) band of 13.56 MHz [142] and employs electromagnetic induction between loop antennas. The supported data rates of NFC technology are 106, 212 and 424 kbit/s [142]. Evolved from established supported radio-frequency identification (RFID) standards, including ISO/IEC 14443 and FeliCa [143], the NFC protocols cover data exchange formats. It allows enabled devices, which are usually smartphones, communicate with other devices containing an NFC tag [144]. NFC has been established in electronic devices and been widely used for contactless payment systems [144]. Most Android devices offer full access to their NFC chips, whereas it is still limited to be used for an Apple mobile device as Apple does not offer access to NFC chips for data and information security [141].

The NFC and Bluetooth share similarities as they are both been widely used in mobile phones and provide short-range communication. Comparing to other conventional technologies, the

Table 4.1 Table comparison between BLE and NFC

Category	BLE	NFC
Price	Around £5 (from £2 to £20) Example: CC2540F128RHAT £1.968 (for 1000+) [145]	Around £0.2 (from ££0.1 to £2) Example: M24SR02-YDW6T/2 £0.2556 (for 1000+) [146]
Size	Around $6 \times 6 \times 1$ (mm) Example: CC2540F128RHAT	Around $2 \times 2 \times 0.5$ (mm) Example: MN63Y1213
Power Consumption	<15 mA (read and transmit)	<15mA (read) Can be batteryless
RF frequency	2.4 GHz	13.56 MHz
Adaptability	Most mobiles	Most mobiles expect Apple
Bit rate	1 Mbit/s	424 kbit/s
Network type	WPAN (Wireless personal area network)	Point-to-point
Setup Time	< 0.006 s	< 0.1 s
Range	≈50 m	< 20 cm

chips of the NFC and Bluetooth communication are usually built in a smaller package and lighter in weight. Bluetooth and NFC have their own advantages and disadvantages. The differences between the parameters are presented in Table 4.1.

According to Table 4.1, NFC operates at slower speeds with a much shorter range comparing to BLE. On the other hand, it consumes less power and needs no pairing before transmission [147]. The setup time of the standard Bluetooth is a few seconds (usually around 6 seconds), whereas the communication of NFC can be established automatically in less than 0.1 seconds and no manual configuration for identification is needed. However, the BLE can set up within 6 milliseconds, which is one of its well-known advantages.

The maximum working distance of NFC is approximately 20 cm, and its communication is likely to be limited to 4 cm [147]. Comparing to BLE whose working distance can be 50 metres, the range of NFC is shorter but is also capable to prevent unwanted interception

and interference caused by other devices. The close proximity makes NFC more suitable for crowded occasions, avoiding correlated signals from other users. The maximum data transfer rate of NFC (424 kbit/s) is slower than that of Bluetooth v2.1 (2.1 Mbit/s) and BLE (1 Mbit/s). The power consumption of both BLE and NFC is comparatively low.

The BLE or other aforementioned technologies require the device to be powered for data transmission, whereas NFC can be self-powered from the RF signal thus can recover data even when the battery is discharged. There are two working modes of NFC, the passive mode and the active mode. Typically, both the initiator device and the target device need to be powered in the active mode as they generate their own fields. The target device in the passive mode, on the contrary, answers by modulating the existing initiator-provided electromagnetic carrier field and draws operating power from it. The target device works as a transponder and achieves batteryless transmission [148]. The extra power needed by the passive tag of NFC can only be supplied during the reading stage.

An NFC tag is usually small in size and inexpensive to deploy, whereas the configuration and administration system, such as the reader/writer, can be bulky [142]. For a long-period treatment, the reader will be kept in the clinics or pharmacies and used by clinicians or clinical staffs for data transmission.

4.1.2 The NFC chip and the data reader

The MN63Y1213 [149] is a large-scale integration (LSI) for RFID, which weights 76 mg in a 2.00 mm x 2.00 mm x 0.37 mm package and fits in lightweight small-scale devices. The detected data can be recorded into the built-in 4-Kbit FeRAM non-volatile memory speedily with low power consumption. The Radio Frequency (RF) interface of MN63Y1213 is constructed with the 13.56 MHz contactless IC card standards. The communication modes include the conventional RF mode, serial mode, and tunnel mode [149], which allows

transmission between reader/writer and the host central processing unit (CPU). However, only the RF communication mode allows batteryless operation by getting supplied from the antenna whereas the other two modes require power supplied to the voltage pin.

The firmware for data acquisition is hosted on the microcontroller through an NFC/RFID contactless reader and writer SCM SCL011 [150], which is shown in Fig. 4.1. The matched antenna coil is winding around the inhaler. The extended ring-shape antenna has been designed to fit the cylindrical shape of the inhaler.
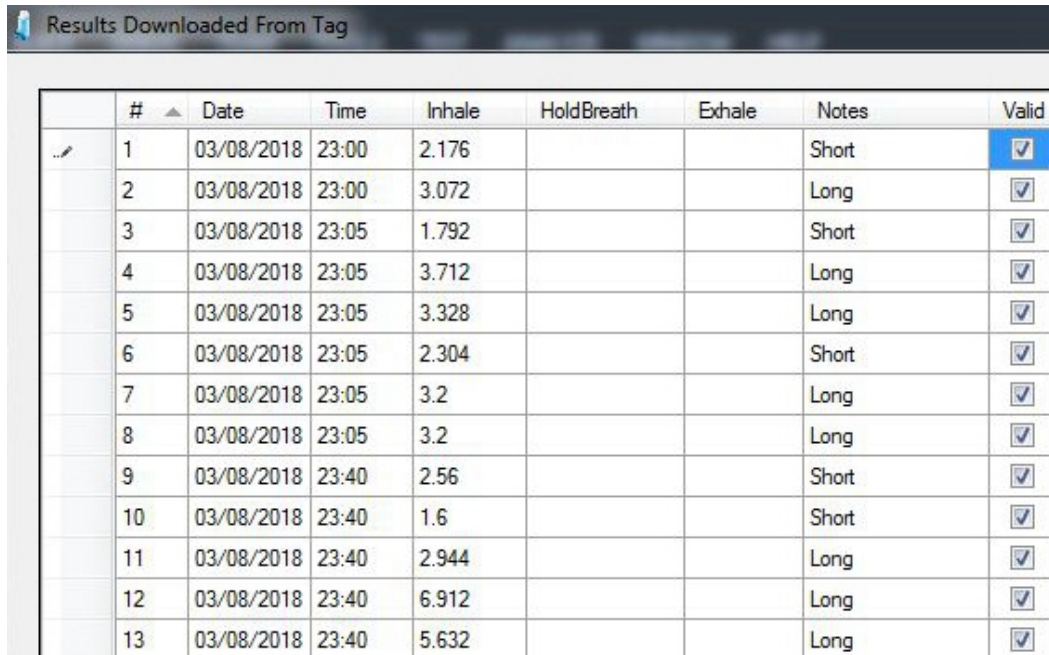
Fig. 4.1 The NFC reader with the extended ring-shaped antenna

Instead of mobile phone applications, a windows C# application was developed on Microsoft Visual Studio 2013 for data-reading. With the same application, the information of patients, such as the patient ID, the real-time clock (RTC) start time, and the storing formats, can be written into the microcontroller and the memory of the NFC tag. The setup stage needs to be completed before treatment and the data of adherence will be collected during the treatment. The inhaler needs to be brought to the clinics or pharmacies when the treatment ends, and the data will be downloaded to a terminal, which is a PC in this research. The interface of the application is designed to be simple and easy to be used.

The interface of the downloaded result is presented in Fig. 4.2. The stored data can be downloaded through the application with a simple evaluation presented. When the duration

	# ▲	Date	Time	Inhale	HoldBreath	Exhale	Notes	Valid
✎	1	03/08/2018	23:00	2.176			Short	☑
	2	03/08/2018	23:00	3.072			Long	☑
	3	03/08/2018	23:05	1.792			Short	☑
	4	03/08/2018	23:05	3.712			Long	☑
	5	03/08/2018	23:05	3.328			Long	☑
	6	03/08/2018	23:05	2.304			Short	☑
	7	03/08/2018	23:05	3.2			Long	☑
	8	03/08/2018	23:05	3.2			Long	☑
	9	03/08/2018	23:40	2.56			Short	☑
	10	03/08/2018	23:40	1.6			Short	☑
	11	03/08/2018	23:40	2.944			Long	☑
	12	03/08/2018	23:40	6.912			Long	☑
	13	03/08/2018	23:40	5.632			Long	☑

Fig. 4.2 Exported result from the NFC chip presented on the PC application

is shorter than 0.8 seconds, it will be demonstrated as 'no inhalation'. With duration from 0.8 seconds to 3 seconds, the inhalation is tagged as 'short inhalation. 'Long inhalation' is considered to be longer than 3 seconds. Any unexpected data will be 'indeterminate'. The stamped date and time of adherence to inhalation are recorded and presented at the same time.

4.2 Data storage

4.2.1 Memory system

To achieve a data analysis, the data of inhalation should be stored safely in the device before data transmission. To avoid information loss due to an unstable power supply or other faulty factors, a non-volatile memory chip is required essentially for the system.

In the existing adherence monitoring studies of CF, the raw audio files were saved directly to memory accessories [60]. The formats of audio coding include the uncompressed audio formats, such as the Waveform audio file format (WAV), and compressed audio formats, such as the lossless compressed Free Lossless Audio Codec (FLAC) and the lossy compressed MP3 [151]. The size of audio files is decided by the compressing method, the sampling rate, and the bit rate. The size of a WAV file of 1 minute is usually several megabytes, and it will also take hundreds of kilobytes for a compressed form like MP3.

The commonly-used non-volatile electronic data storage include electrically erasable programmable read-only memory (EEPROM) [152], FeRAM, resistive random-access memory (ReRAM), erasable programmable read-only memory (EPROM), and other types of storage. The EEPROM has been widely integrated into microcontrollers or other electronic devices allowing individual bytes in EEPROMs to be programmed and erased. The common size of an EEPROM is from 4 Kbit to 1 Mbit, which can store a relatively small amount of data.

As saving raw audio requires larger storage, large-capacity memory cards were introduced in the research studies. SD cards have been commonly used [23] and size of storage can be modified by changing memory cards in this way. Notwithstanding the big amount of raw data that a single SD card can store, it consumes more power. A flash memory chip requires power for reading and writing data and long-period storing, and the power consumption of memory is related to the type and size of storage. Besides, the SD cards are considerable bulky for the target inhalation monitoring system even with a Micro SD card, which is normally sized at 15 mm × 11 mm × 1 mm. The price of a Micro SD is also relatively high.

To achieve non-volatility and avoid potential data loss due to the erratic power supply, a new non-volatile memory, FeRAM, has been used in this research. the FeRAM has a large maximum read/write endurance of static random access memory (SRAM) and high stability and reliability of flash [153]. With higher flexibility, it also requires less power consumption and less time to be written in [154]. To shrink the size of the developed device, with no extra

storing component has been added into the system, the accessory FeRAM of NFC chip is used to store data.

A novel highly-compressed storing method is provided in this work. In contrast of the existing methods which store raw audio files, the important information of adherence for monthly usage, including the duration length of each inhalation and the date and time of occurrence, will be stored using only a few hundred bytes of NVRAM. It is considerable energy and cost efficient over many of the other memory intensive approaches.

4.2.2   The highly-compressed format of inhalation duration

With high writing speed, the 4-Kbit built-in FeRAM for RFID consists of 32 blocks where 5 among them are for system area [149]. The rest 27 blocks are available for data storage. Each block contains 16 bytes. The date and time of setting, the user ID number, system encryption key and authentication key for transmission are written through the PC application earlier than the inhalation detection. To store the data of adherence for a full treatment, several solutions have been developed to fit in the 432-byte (16×27) memory.

The original 16-bit inhalation length

The uncompressed duration data of a single event of inhalation are unsigned 16-bit integers from 0 to 65,535. The pulse length is calculated by counting processed raw samples between the start point and the end point of a detected event.

When N_s is the counted number of detected samples, and the sampling rate SPS, which equals 5578, means the number of processed sample in a second, the duration length DL is

calculated by:

$$DL = \frac{N_s}{SPS} \times 1000(ms) \tag{4.1}$$

In this manner, each number represents 1 ms and the duration of inhalation is ranged from 0 to 65,535 seconds. However, a regular inhalation usually takes from 1 to 10 seconds, depending on individuals. In a later initial user test, 115 inhalations from 15 different volunteers were collected. A temperature sensor and a reference microphone have been used for duration measurement. The histogram of inhalation duration collected from 15 participants is indicated in Fig. 4.3, which presents that the measured duration of inhalation distributes from 1.5 to 8 seconds.

Inhalation Duration (s)	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5	8	8.5
Distribution	0%	0%	2%	11%	6%	22%	14%	12%	16%	5%	5%	2%	0%	1%	3%	1%	0%
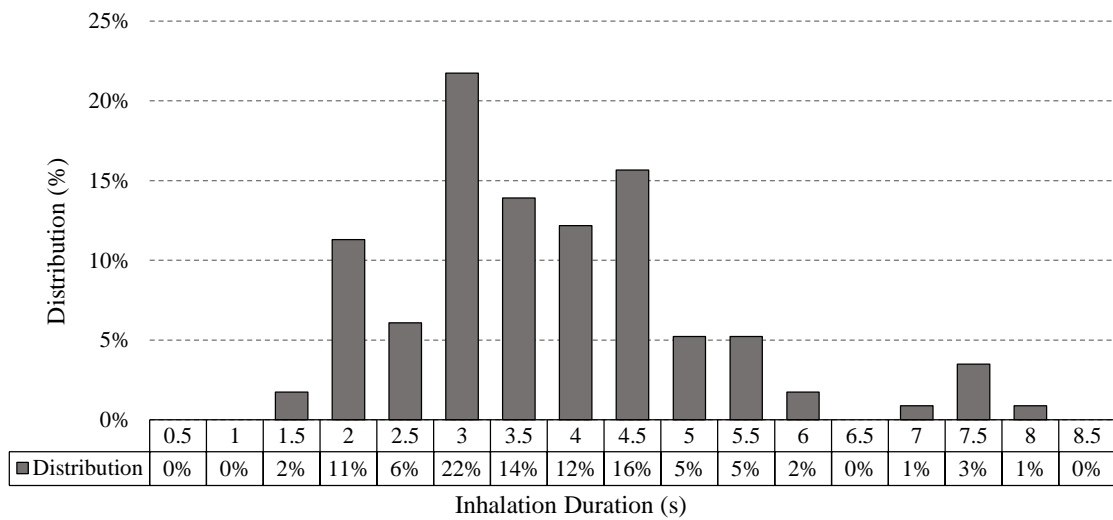
Fig. 4.3 The duration distribution of collected 115 inhalations from 15 participants

From the histogram, it is clarified that inhalation distributes mostly (93%) from 2 to 6 seconds. As the value of the stored 16-bit inhalation data has a wider range than needed, highly-compressed coding methods can be introduced into the storing system to massively shrink the data size. An 8-bit compression and a 4-bit compression of inhalation have been developed with different resolutions.

8-bit coded output with 12-bits range and 5-bit resolution

An unsigned 8-bit integer is coded to present the original 16-bit integer of duration length with a 5-bit resolution. The first 5 bits of the coded number is a right-shifted value of the original data, and the rest 3 bits counts how many bits have been shifted. With this developed coding method, the number of bits required to describe the duration length of each inhaling event has been reduced.

When v presents the samples and n counts the shifted bits, the logic is presented as:

Table 4.2 The code logic of bit shifting

if (v < 128) return v; //small value
while(v>255) n++; v»=1; //right shift
if (n>7) return 0xffff; //overflow maximum value of a 3-bit integer n
return (v & 0xf8)	n; //v is stored in the first 5 bits, n is stored in the last 3 bits

In the decoding stage on the PC application, the duration is calculated as $(v\&0xf8) \times 2^n$ if v is not less than 128. Otherwise, it equals to its original value. A value table is presented in Table 4.3. The error in millisecond presents the possible difference between the original value and the coded number.

When inhalation does not exceed 8 seconds, the resolution is limited to 32-bit following the above coding method, which equals only 0.032 seconds. Comparing to a duration of several seconds, a tolerance of 0.032 seconds is negligible for the following adherence analysis.

The improved 8-bit coded output

The 8-bit coded output with 12-bits range and 5-bit resolution has been presented before and used in the preparation tests and the later user testing experiments with volunteers. This

Table 4.3 The presented coding methods

Inhalation	v	n	Coded Range	Error (Millisecond)
0-127	v	0	0-127	1
128-255	v	0	128-248	1 - 8
256-511	v»1	1	256-496	2 - 16
512-1023	v»2	2	512-992	4 - 32
1024-2047	v»3	3	1024-1984	8 - 63
2048-4095	v»4	4	2048-3968	16 - 127
4096-8191	v»5	5	4096-7936	32 - 255
8192-16383	v»6	6	8192-15872	64 - 511
16384-32767	v»7	7	16384-31744	128 - 1023
32768-65535	0xffff	0	65535	0 - 32767

method transforms every number with the 16-bit range to a corresponding 8-bit number, which may avoid the data loss due to a failed detection without an ending sign. However, according to the measured in Fig. 4.3, the inhalation for a human is normally limited in 10 seconds.

An improved 8-bit coded output cuts 8 bits directly from the original 16-bit inhalation, discarding the LSBs. The methodology, the value table, the coded range and the increment unit are presented in Table 4.4. The error in millisecond presents the fixed difference between the original value and the coded number.

After the data collecting stage from user testing experiments, the duration length is considered not as long as a value that exceeded 10 seconds. Comparing to the first 8-bit method, for a duration less than 10 seconds (10000 as an integer), the coded range is from 64 to 16383. In the range, the error of the new methods is less than 64 milliseconds, whereas the old method is 64 - 511 milliseconds. The revised method will be used in the later experiments, and the 7th number (bit 6) to the 14th (bit 13) will be cut from the original 16-bit inhalation to reduce error.

Table 4.4 The improved LSBs-cutting 8-bit coding methods

Used bits (LSB)	Coded Range (Integer)	Error (Millisecond))
1 to 8	1 - 255	1
2 to 9	2 - 511	2
3 to 10	4 - 1023	4
4 to 11	8 - 2047	8
5 to 12	16 - 4095	16
6 to 13	32 - 8191	32
7 to 14	64 - 16383	64
8 to 15	128 - 32767	128
9 to 16	256 - 56635	256

The highly-compressed 4-bit coded output

For a more highly compressed scale, an exponential function returns a 4-bit power number of 2 from the original data. For example, for number 32, the returned exponent is 5. However, as an exponential value, the resolution of a sample between 4-second (4096) and 8-second (8192) is 4 seconds (4096), which loses the accurateness of detection and measurement.

A new 4-bit coding method for compressing inhalation length has been developed and presented in Table 4.5 based on the measured duration.

Table 4.5 The method of 4-bit coding

Original Range	Calculation	Coded Range	Increment Unit (ms)
v<1	No	0	Short Inhalation
1000≤v<2000	1+(v-1000)/500	1-2	500
2000≤v<5000	3+(v-2000)/300	3-12	300
5000≤v<7000	11+(v-5000)/1000	13-14	1000
v>7000	No	15	Long Inhalation

86% of recorded inhalations were from 2 to 5 seconds, and any measured duration with a pulse length in this range has the highest resolution. From a 2-second inhalation to a 5-second sample, the resolution is 0.3 seconds. With the possible error rate from 6% (0.3/5) to 15% (0.3/2), the 4-bit solution is able to represent important information of duration.

#	Date	Time	Inhale	HoldBreath	Exhale
1	07/09/2018	15:30	2.6		
2	07/09/2018	15:30	2		
3	07/09/2018	15:30	3.5		
4	07/09/2018	15:30	3.5		
5	07/09/2018	15:30	3.8		
6	07/09/2018	15:30	4.7		
7	07/09/2018	15:30	2.9		
8	07/09/2018	15:50	3.5		
9	07/09/2018	15:50	4.7		
10	07/09/2018	15:50	3.2		
11	07/09/2018	15:50	3.8		
12	07/09/2018	15:50	2		
13	07/09/2018	15:50	4.4		
14	07/09/2018	15:50	1.5		
15	07/09/2018	15:50	2		
16	07/09/2018	15:50	4.7		
17	07/09/2018	15:50	3.2		
18	07/09/2018	15:50	4.4		
19	07/09/2018	15:50	3.8		

Fig. 4.4 The presented interface of 4-bit solution

The interface of the 4-bit solution in Fig. 4.4 presents the exported length data of detected inhalation. The samples of duration are abbreviated from the original duration length and categorised by every 0.3 seconds. With a concession of duration accuracy, the length of each inhalation still presents important information about how long effective inhalation continues during each administering. The presented number of each range is a floor value, which means for data between two categories, it will be categorised to the lower range. Each duration has a possible negative increment from 0 to 0.3 seconds.

4.2.3   The compressed method for date/time stamping

The RTC starts counting time once the information of setup date and time have been written into the system through the PC application. The RTC needs to be powered during the entire treatment to ensure its stability. The data of time from RTC are 16-bit unsigned integers, which are accurate to each minute and capable to function continuously within 45 days (64800 minutes).

Instead of recording the actual date and time of each administering event into memory, the time differences (in minutes) between doses are stored. The length of the gap between a detected administering event and the last event is also measured as an unsigned 16-bit integer, which is recorded with inhalation for date/time stamping, clarifying when the adherence was monitored and recorded.

To shrink the data size and save memory space, the integer of time difference can be compressed to an unsigned 6-bit integer in compressed format by the progressive increment with different resolutions. The coded result is presented in minutes. The coding method is depicted in Table 4.6.

Table 4.6 Coded 6-bit time difference in minutes

Original Range	Time Presented	Resolution	Calculation	Coded Range
0-59	First Hour	5-minute	TD/5	1-11
60-1440	Next 23 Hours	30-minute	12+((TD-60)/30)	12-58
1441-3840 (64×60)	hours 24-64	8-hour	12+46+((TD-1440) /(8×60))	58-63
>3840	>64 hours	Limited to 64	No Calculation	64

According to the instructions for use, doses should be inhaled twice daily with 4 capsules for each. It is recommended that patients should take doses every 12 hours [16]. During a normal treatment with no serious forgetfulness issue occurs, it is unlikely that the gaps between doses are longer than 3 days.

4.2.4   Compressed duration with stamped time

The solution with different compression will be selected based on the duration of the treatment and the size of inner storage. As the current constructed prototype contains a 0.5 KiB FeRAM, the discussion will base on the size of 0.5 KiB as the example. With a fixed storage size, the highly-compressed method takes priority when the treatment is longer. As mentioned in Section 1.1.3, 4 capsules will be consumed as a full dose, and two doses should be taken daily. It is indicated from the user information [15] that at least 2 inhalations are needed to empty each capsule. As a result, 400 - 500 inhalation may occur during a 28-day treatment.

The presented 8-bit solutions, including the 8-bit coded output with 12-bits range and 5-bit resolution and the improved LSBs-cutting 8-bit coding methods, both take 8 bits (1 byte). Regardless of the possible error, these two solutions will be discussed together as 8-bit solutions.
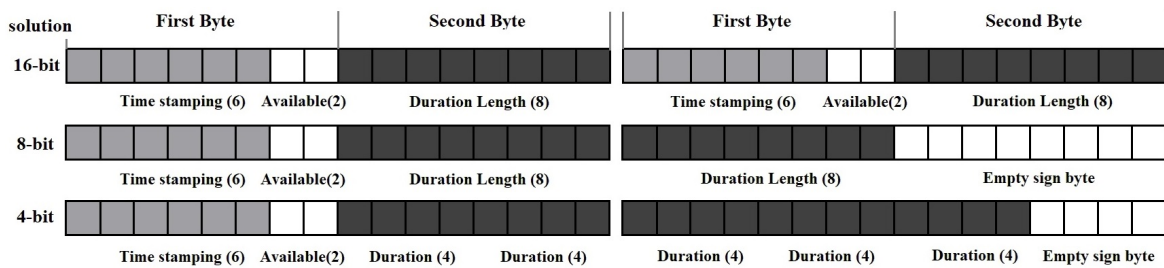
Fig. 4.5 The diagram of data structure using different solutions

The diagram of data structure in Fig. 4.5 presents how the bytes have been occupied in each solution.

a. the 16-bit solution

The 16-bit solution consists of an 8-bit unsigned integer of compressed inhalation length, a 6-bit coded unsigned integer for stamped date and time. The first 6 bits of the first byte is for time stamping. The last 2 bits are flexible for future design. In the future version, it can be

used to record a simple evaluation of how strong and stable the inhalations are depending on the average value of ZCR and STMD. The second byte is used for storing duration. The measured length data are coded into 8-bit integers in 12-bits range and 5-bit resolution.

b. the 8-bit solution

The 16-bit solution uses 6 bits to store time, whereas the 16 intakes of the same doses share the same time stamping as a dose is likely to be taken within 5 minutes. It is worthy of combining similar recorded time samples together to reduce data size. The 8-bit solution abandons time updating if the latest inhalation is taken within one hour since the last inhalation, which means the neighbouring events belong to the same dose. The time will only be updated and recorded when the current inhalation occurs after one hour later than the stored time. It adopts the 8-bit duration length like the 16-bit solution.

An empty byte will be stored into the memory after time-stamped inhalation to label the end of a complete dose. The PC application detects empty bytes to demarcate a full event of medicine administering. The new method cuts the memory size of single doses to roughly half of the original size.

c. the 4-bit solution

The capsule will be removed and checked after two intakes following the instructions. Although it is recommended that 2 intakes should empty a capsule, there is a possibility of an unemptied capsule. Repeating the inhalation helps to inhale the rest of the powder, whereas it requires extra memory to store the added data. To store information of adherence for more than 28 days into the limited memory, a 4-bit solution with the higher resolution has been developed.

Each inhalation will take 4 bits using the 4-bit compressing method in Section 4.2.2. Therefore any inhalation shorter than 1 second will be abandoned, and no 0 value will be adopted to represent inhalation. The coded data will be from 1 to 15 and presents data from 1 second. An empty half byte (0000) will follow the time-stamping data to label the end of a dose. The PC application detects 4-zero binary data to demarcate a full event of medicine administering.

The discussion and context of selecting solutions

The data of original 16-bit inhalation length has been coded to 8-bit and highly-compressed 4-bit coded output with different resolutions. Combining the 6-bit data of time stamping, the essential information of each inhalation event can be stored in the 16-bit, 8-bit, or 4-bit compressed solutions. As the 16-bit and 8-bit solutions are based on the same data resolution with only the time-stamping method varies, both of them have the same duration difference against reference from 0 to 0.26 seconds for an inhaling event that is limited in 8 seconds. The 4-bit solution categorises the duration length with a maximum duration difference against the reference of 0.3 seconds, and the data will be slightly biased with an upward data wrapping.

The storable doses can be calculated based on the available storage size and the selected solution. The quantity of available doses is the quotient of the available 432 bytes and the byte requirement of each dose. Assuming each capsule can be successfully consumed by the minimum 2 inhalations, each dose requires 8 inhalations.

For the 16-bit solution, as each inhalation requires 2 bytes, at least 16 bytes need to be occupied for one dose. The adherence information within around 10 days (13.5 days for a minimum of 16 inhalations daily in the available 432 bytes) can be fitted into the 0.5 KiB FeRAM memory with the aforementioned compressing method. With the limited storage, the 16-bit solution can be selected in short treatment for less than 2 weeks.

For the 8-bit solution, 8 inhalations cost 8 bytes for a dose, with the attached time-stamping byte and an empty sign byte. The minimum requirement is 10 bytes. The adherence information within 21 days for a minimum of 16 intakes daily can be fitted into the available 432 bytes. The 8-bit solution will be suitable for any treatment for around 20 days.

For the 4-bit solution, 8 inhalations cost 4 bytes for a dose, with the attached time-stamping byte and the 0.5 empty sign byte. The minimum storing requirement of each dose is 5.5 bytes. The adherence information within 39 days for a minimum of 16 intakes daily can be fitted into the available 432 bytes. The 4-bit solution will be suitable for a long treatment, such as the 28-day full treatment.

Table 4.7 Storable doses for a 0.5 KiB FeRAM of different solutions

Solution	Bytes per Dose (Minimum)	Storable Doses	Maximum Days
16-bit	16	27	13
8-bit	10	43	21
4-bit	5.5	78	39

In Table 4.7, the number of doses that can be stored in the 0.5 KiB FeRAM using the developed coding solutions are presented. Comparing to the 8-bit solution, the 4-bit solution is more suitable for the small size memory and long-period treatment, whereas the distortion of duration length is considerably higher with a biased increment up to 0.3 seconds.

To keep both the high resolution and accuracy for long-period treatment, an improved NFC replacement is required. A bigger FeRAM is required if a 28-day treatment is monitored using the 8-bit solution. A possible replacement is MN63Y1221 [155]. The built-in FeRAM non-volatile memory of MN63Y1221 is 8-Kbit (approximately 1 KiB), which doubles the size of FeRAM of MN63Y1213. The accuracy of different methods will be further analysed in Chapter 6 after the user testing experiments.

4.3   Conclusion

The aim of the research was to design and develop a monitoring system with advanced transmission and storing approaches. Compared to the BLE technology, NFC consumes less power, and the device can be connected without paring stage. As one of the main considerations is power consumption, and the NFC can be powered by the RF signal and transmit data without battery or any other type of power supply, it shows advantages and is more suitable for a portable long-period electronic system.

The duration of detection is saved as highly-compressed formats with real-time stamping. The size of available FeRAM of the developed prototype, which is the storage of NFC chip to avoid data loss and enhance the readability of data even when the battery goes flat, is limited to 0.5 KiB. With presented 16-bit, 8-bit, and 4-bit coding methods for data compression, the adherence of a full long-term treatment can be stored as small size data to fit the limited storage.

Chapter 5

The Prototype System and Experimental Procedure

The chapter introduces the new inhalation detecting prototype. A low-cost self-printed finger-tip on-chip recognising system has been developed with a TI low-power microcontroller, an omnidirectional microphone, a bicolour LED, a 3-axis accelerometer, and an NFC chip for data transmission. The system has achieved accurate real-time inhalation recognition and duration measurement based on simple time-domain features to reduce computational complexity and memory cost. The energy-effective device was designed with low-power working modes, which extended its durability with the capability to fit a 28-day treatment. The data were stored in FeRAM of the NFC chip to improve the non-volatility and avoid data loss. With a LED direct feedback system, the result of a fast evaluation was presented to the users.

5.1   Introduction

The main target of the research is to design a durable digital signal processing system as the attachment to an inhaler. The design challenges of the electronic device include:

1. The detection and duration measurement of inhalation is reliable with high accuracy against noises.

2. Computational and memory cost of features are low to fit the limited processing RAM.

3. The device is energy-effective with power saving mode and low-power processing.

4. Data recovery with drained battery is available to prevent unexpected information loss.

5. The simplicity of use is concerned to avoid frustration or even abandonment of the device.

To maintain stable usage during a long-period treatment, the power consumption needs to be limited to extend the durability. Wearable or portable electronic devices are more pervasive with increasing requirements of low-power architecture. From existing knowledge, the energy-efficient signal processing can be achieved through software or hardware implementation. Algorithms and software can be optimised for the target architecture, which can also be developed to avoid massive shuffling of data between memory and processor. Improving coding skills with considering strength and weakness in the developed system saves computing clock cycles and power. The energy consumption can also be reduced with the reduction of supply voltage and geometry with the recent improvements in semiconductor technology [66].

5.2   The early design: the RFduino board

In an early prototype, an RFduino board [156] had been used to develop the adherence monitoring system with BLE communication technology. The RFduino includes a high-performance BLE radio transceiver and an bulit-in ARM Cortex M0 microcontroller, which is optimised for small silicon die size and used in the low price chips. The system was programmed efficiently through the simple-to-use Arduino integrated development environment (IDE) using RFduino extensions, which can also be wirelessly communicated with iOS Phones, iPad or Android smartphones or tablets. The size of the surface-mount technology (SMT) module is 15 mm in length, 15 mm in width and 3.5 mm in height, which can fulfil the transmitting and processing function requirements at the same time.

For rapid development and prototyping, the early design was built on the developed RFduino dual in-line package (DIP) shield, which was a modular accessory that can be directly plugged into another RFduino DIP shield. The prototype contained an RFduino module, an RGB LED/Button Shield for direct feedback, a coin Battery shield for power supply, With an external USB shield for programming. The size from a regular RFduino DIP shield was around 22.86 mm $\times$ 23.37 mm $\times$ 18.4 mm, whereas the used coin battery shield was 22.86 mm $\times$ 46.99 mm $\times$ 18.3 mm. The built prototype is 22.86 mm $\times$ 46.99 mm $\times$ 41.7 mm. The early design was too bulky as an attached electronic device and more suitable as an external accessory device.

The accompanied RGB LED/Button board includes three different colours of flashlights, which gives direct feedback to the users. When the green light is flashing, the time counter is working and measuring time pulse. The idea of using LEDs to demonstrate real-time processing has been kept in the future version.

Sound source identification based on the real-time calculation of key factors, such as over-zero rate, variation, and deviation, has been designed with a low-cost system. A MEMS

microphone ADMP401 was used for sound sensing. The Android application receives the data from the RFduino board immediately when the board is active and has detected updated identified data. The device can reject surrounding noises, such as the mechanical sound from the traffic and vehicles, the human voice and music under different occasions.
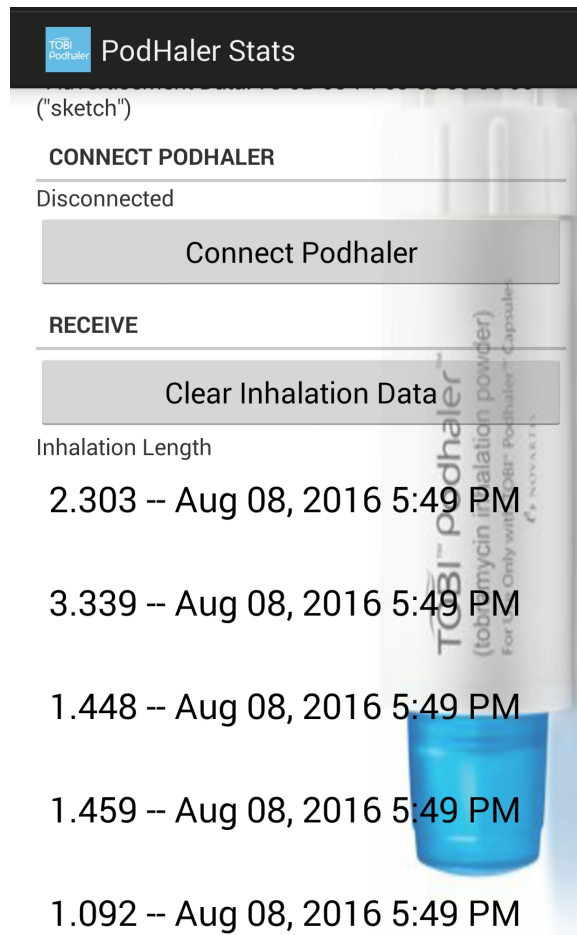
Fig. 5.1 The Android application interface of the data receiving system

The BLE technology was used for data transmission from the sensing system to the terminal. Compare to classic Bluetooth, BLE is more economical and widely supported by the current generation of smartphones and tablets. Both iOS and Android application were built at the early stage of the project. The Android application was developed to receive data information such as the inhalation duration length and the date and time of a recorded event, which is presented in Fig. 5.1. The data can be updated automatically when a new inhalation was

detected with an active BLE module. The data would be analysed on a PC-based platform and the monitored adherence could be analysed and indicated.

More than 800 inhaling events have been tested for detailed threshold adjustment. The original threshold was defined according to the measured value using MATLAB. A yellow LED was used to represent whether the ZCR reached the threshold, and a green LED was used to represent whether the STMD reached the threshold. 200 samples was set as one processing window. The threshold value was adjusted by observing the LED. If the LEDs were flashing during inhalation, the threshold was exact or lower than needed. When the LEDs were not flashing during inhalation, the threshold was higher than needed. The threshold value was adjusted until the LEDs were flashing during inhalation and reach the maximum of allowance.

An early user testing experiments estimated the accuracy of the developed detecting RFduino system with 200 inhaling events. 96.5% of inhalations have been successfully detected and marked as inhalation correctly, and 100% of detected inhaling events have been transmitted to a paired phone successfully.

As an early prototype, the RFduino version with BLE was designed and developed by the writer of the thesis, which provided an initial approach to low-cost low-complexity processing on small-size electronics and short-distance transmission between the device and the terminal, which proved the feasibility of the methodology. The idea of using 1 KiB RAM has been kept in the second version on MSP-EXP430FR573x. The technology of BLE has been compared to the later NFC version based on performance and power consumption. The option of using BLE can still be considered as diversified design if different types of products are in need.

5.3 The PCB based on MSP-EXP430FR573x

5.3.1 The printed prototype

The improved design considers physical size, energy consumption, the ability to store data for a full treatment, and the data safety even with a depleted battery. A low-cost development board with a microcontroller capable of performing digital signal processing operation was needed. Instead of the Arduino platform, the TI MSP430FR573x family of ultra-low-power microcontrollers [153] was selected as the main processor. The system fits both MSP430FR5736 and MSP430FR5738, which was initially designed to fit MSP430FR5736 with MSP430FR5738 introduced in the later printed design. NFC technology has been introduced to replace the BLE for data transmission. The idea of using a limited 1 KiB RAM and a simple processor was decided by the writer of this thesis and The PCB was built and developed by Dr Tim Good. The components were selected and tested together. The core code of audio processing was programmed by the writer.

The revised prototype was a thumbnail-sized thin PCB as shown in Fig. 5.2, which contains:

- A green/red bicolour LED and a crystal

- A TI microcontroller: MSP430FR5736 [153]

- An omnidirectional microphone INMP441 with an I^2S digital output [132].

- A 3-axis accelerometer LIS2DH12 with I2C/SPI digital output interface [157].

- An NFC (RFID Reader/Transponder IC) chip MN63Y1213 [149].

- External components: small battery and NFC antenna coil

The new inhalation-detector is a finger-tip medical monitoring electronic device for inhalation detection, recording, and evaluation, which is designed to be attached to the body of an

Fig. 5.2 The 12.5 mm × 10 mm × 1.5 mm printed circuit board

inhaler. The printed chip is 12.5 mm × 10 mm × 1.5 mm in size. During the testing stage, the chip is sealed underneath a 3D printed plastic protective shell, which is a 22 mm wide ring-shaped cover and fits tightly to the inhaler. The inner diameter of the ring is 26 mm and the outside diameter is 31 mm. The shell is coloured white and not being transparent, whereas the red/green LED can be observed externally through it to provide direct feedback.

The power supply is an ultra-thin curved flat flexible Li-polymer battery 201030. The nominal voltage is 3.7V, and it is rechargeable. It is 2 mm × 10 mm × 30 mm in size, which fits underneath the shell with the processing device. For a final version to be widely produced in future, other suitable batteries may be introduced to replace the current version, considering the price and the stability. The locations of the chip and the battery are presented in Fig. 5.3.

The TI MSP430FR573x microcontrollers feature embedded non-volatile FeRAM and 16-bit CPU with added peripherals, including an ADC, a comparator, 3 serial channels, an internal DMA, a hardware multiplier, an RTC, 5 timers, and digital I/Os [153, 158]. The

(a) The protective shelf with located chip   (b) The location of the bendable battery

Fig. 5.3 The printed protective shelf with the locations of the chip and battery presented

microcontrollers are capable to fit in portable sensing applications with long battery life as being ultra-low-power, lightweight and available in various small packages. In this research, a 4 mm × 4 mm × 1 mm package of MSP430FR5738 has been selected to suit the small-size design of the prototype.

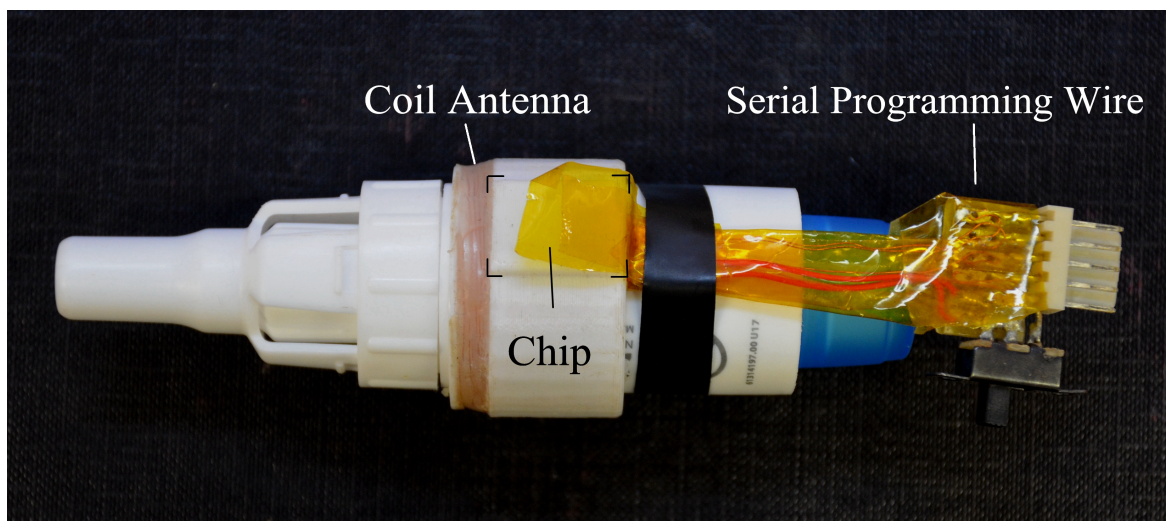Fig. 5.4 The developed prototype with the programming cable and the protective shell

The developed prototype with the location of the chip and the programming cable is presented in Fig. 5.4. The version in actual usage will exclude the programming cable.

The MSP-EXP430FR5739 Experimenter Board was used for programming and debugging MSP430FR573x boards through a serial cable. Pairing with a PC application, it provided an

efficient method to develop, debug, and implement prototypes. It is supported by both IAR
Embedded Workbench and Code Composer Studio (CCS). In this research, CCS 6.1.2 was
used as the programming platform. The used Experimenter Board with the serial cable is
shown in Fig. 5.5.
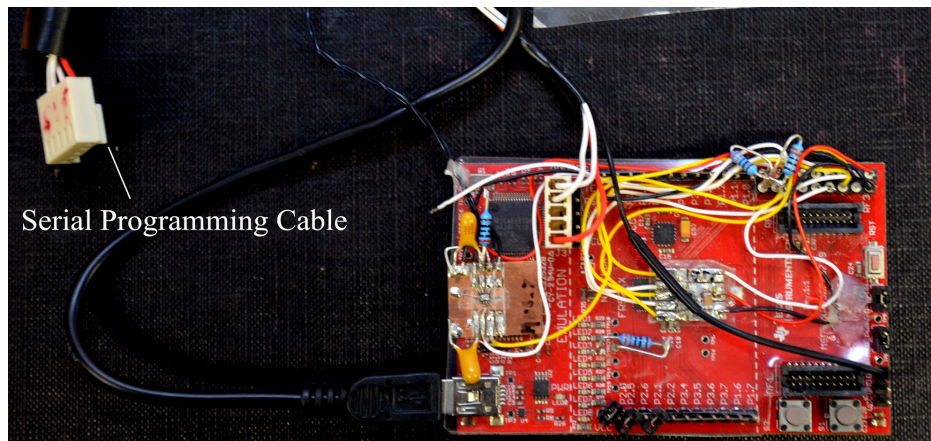
Fig. 5.5 MSP430FR5739 FRAM Experimenter Board with serial programming cable

5.3.2   The low-power inhalation detecting system

The system hardware block diagram for the detection-transmission solution is depicted in
Fig. 5.6.

The accelerometer stays activated when the system is being idle and deep-sleeping, and
the information of movements is collected through a multiplexer (MUX), converted by an
internal ADC, and finally be transferred through SPI. Any sudden movement occurs a signal
of interrupt and is received by the microcontroller. The signals of changing angles are
being received when the system is triggered by the waking-up motion. With an activated
microphone, the accelerometer stops to function. The analogue raw audio is converted by an
ADC and pre-filtered by anti-aliasing filters. The pre-processed signals can be transmitted
through I^2S serial port of the microphone to the microcontroller using the adapted SPI MISO
(Master In Slave Out).
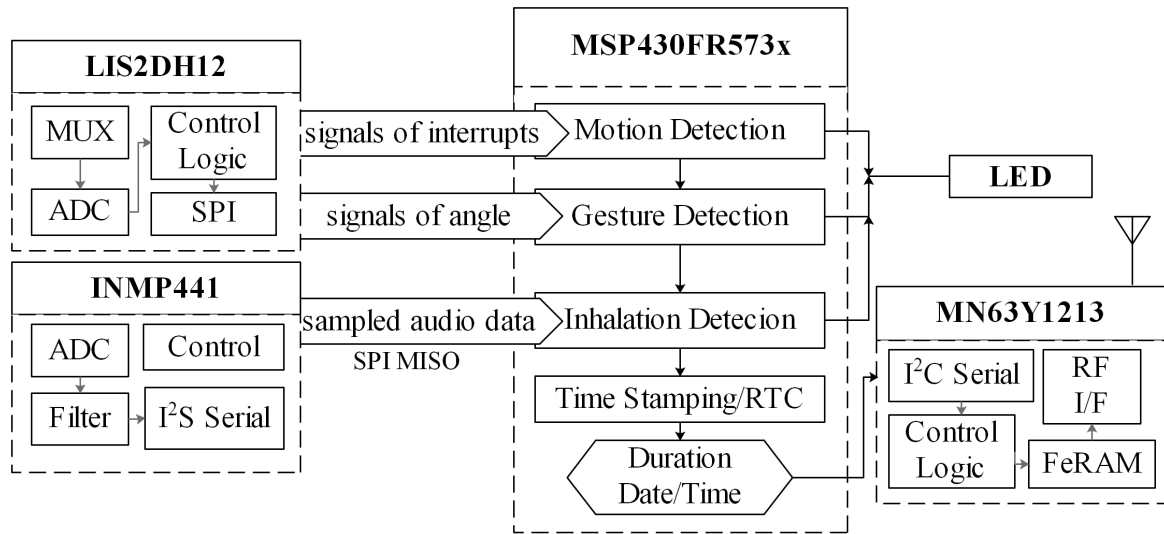
Fig. 5.6 The block diagram of the system

With raw audio signal abandoned, only the duration of each inhalation is extracted, and the system attaches the information time and date (in minutes) to each event. The result of detection is written into the NFC chip directly through I^2C serial without being stored in the main processor.

To achieve a portable system, the size of the applied battery is limited to fit in the inner side of the protective shell and thus restrict its electrical capacity. The selected power supply is a 43-mAh flexible rechargeable Li-Ion battery. The information of the typical working current of each integrated component is presented in Table 5.1, according to their data sheets.

Table 5.1 Power consumption of components

Component	Working Mode	Working Current
MSP430FR5738	Active	81.4 μA/MHz
	Standby	6.3 μA
	Real-Time Clock	1.5 μA
	Shutdown	0.32 μA
LIS2DH12	(Typical)	2 μA
INMP441	(Typical)	1.4 mA
MN63Y1213	(Typical)	Batteryless RF communication

The microphone is shown to be the most power-hungry component in the system. To reduce power consumption, the progress of audio processing will only be activated when the user is inhaling. The working modes of a complete inhalation detection include the sleep mode, the gesture detection mode, the inhalation detection mode, and the data transmission mode. The system flowchart is demonstrated in Fig. 5.7.

1) The sleep mode

To reduce energy consumption and maintain long-period functionality, the system stays in the ultra-low-power sleep mode when it is not activated. The microphone keeps off at this stage. The accelerometer is activated for receiving the triggering signal of preset motion and the RTC is working after the preset of system till the data collection.

2) The gesture detection mode

A shaking motion is defined as the triggering signal from the accelerometer. The attached LED flashes once the system is triggered, highlighting the current working status of the system. The horizontal alignment of the inhaler is detected and measured at the gesture detection stage.

3) The inhalation detection mode

With a horizontal position, the microphone starts to function the real-time inhalation detection. The ZCR and STMD are calculated in parallel. As the calculations of the ZCR and STMD are both based on neighbouring samples, the samples will only be read once, and the parallel calculations can be achieved by using two separated integers as the accumulation of sign differences or value differences.

The pulse length is counted by sample number when the features are within the preset threshold range. The threshold ranges of different features have been evaluated and calculated by a previous training system. To enhance the robustness against the noisy environment, a concept of tolerance threshold has been introduced into the system to avoid triggering
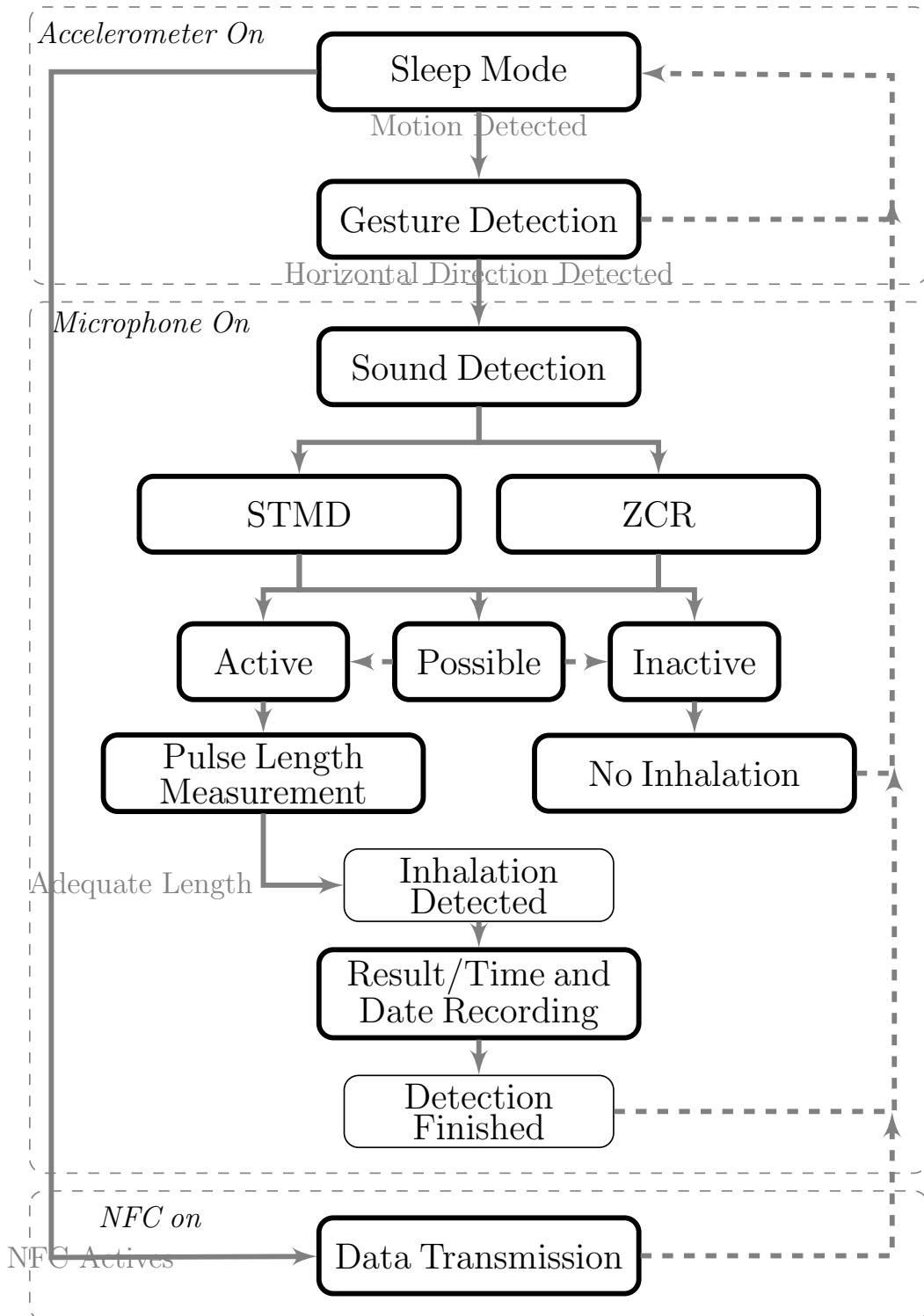
Fig. 5.7 The diagram of the inhalation detecting system

the wrong end of inhalation from the random sudden noise. The tolerance threshold adds 10% - 15% increment to the original threshold, which indicates a 'possible' situation of identification thus reduce instability. When the features of samples in a window are within the inhalation range but in the tolerance range, the counter for duration length measurement will not increase, but the counter for triggering the end of inhalation will not increase and stays in zero.

The detailed methodology of inhalation detection will be presented and discussed in Fig. 5.16. Inhalation will be correctly detected when the measured pulse length is longer than 0.6 second. The duration of a correct inhalation and the event time are stored into the FeRAM of NFC chip.

4) The NFC transmission mode

When an NFC reader activates the system, the stored data can be transmitted from the FeRAM of the NFC chip to a PC through the accessorial application. During the transmission mode, both the accelerometer and the microphone stays off. The power for transmission is supplied by the NFC reader without battery consumption.

During each mode, a protective module is inserted to prevent overflow, bad loops, or other unexpected errors, and also to avoid extra power cost. The system goes back to the sleep mode for unexpected incidents. 15 seconds without horizontal holding in the gesture detection mode or 20 seconds without identified inhalation in the sound detection mode turns the system back to sleep mode.

The measured power consumption of the audio detection mode is around 3 mA, and the daily routine of intakes is shorter than 2 minutes. With the real-time clock and accelerometer functioning, the working current in sleep mode is measured to be from 30 μA to 45 μA. With the 43-mAh battery, the device can operate normally for more than the 30 days expected life of an inhaler and has a pre-usage shelf life in excess of several months.

Fig. 5.8 illustrates the procedure of adherence monitoring of a full treatment from device setting to data analysis and the procedure of data extraction from the electronic medical device through utilised layers to a PC terminal.

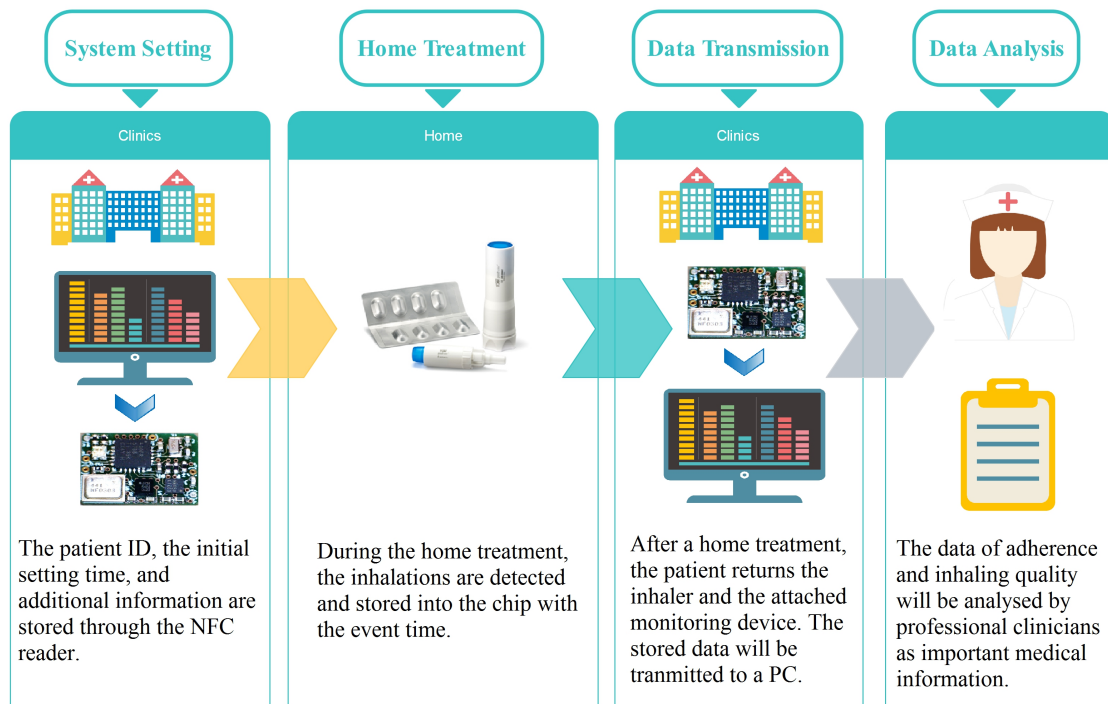| System Setting	Home Treatment	Data Transmission	Data Analysis
Clinics	Home	Clinics	
The patient ID, the initial setting time, and additional information are stored through the NFC reader.	During the home treatment, the inhalations are detected and stored into the chip with the event time.	After a home treatment, the patient returns the inhaler and the attached monitoring device. The stored data will be tranmitted to a PC.	The data of adherence and inhaling quality will be analysed by professional clinicians as important medical information.

Fig. 5.8 The progress of a full treatment

The monitoring device will be preset by professional clinicians using the PC read/write application. The ID of each patient and the date of setting as the reference of time stamping will be stored into the hardware. The battery needs to be applied before system setting and continues supplying the system during the home treatment. After a time-fixed treatment, the device will be returned to the clinics or pharmacies. The wireless and battery-less operation permits the device to be placed in a prepaid postage envelope and posted to a sorting office. The device will be scanned and the data will be transferred into the database without needing to remove from envelopes and the data of adherence will be extracted and analysed by professional clinicians.

This method saves time and cost of clinical visits, it is more economical and convenient to both the users and the clinicians. The patients only need to handle the inhaler with the attached chip without being involved in the troublesome system setting or data management. And they are not responsible for data transmission and thus avoid the risk of data loss due to forgetfulness or improper operations. The adherence is monitored and recorded automatically with regular medicine administering, and the patient needs no extra behaviour for data saving.

For the current version of the developed prototype, the measured power consumption of the system in audio detection mode is around 3 mA. With the real-time clock and accelerometer functioning, the measured current in sleep mode is from 15 μA to 45 μA. The shelf current, which is discharged before the system activates, is low as 0.3 μA. As the audio detection is expected to continue for several minutes each day, and the device stays idle for most of the time. The device can operate normally for more than the expected life of an inhaler and has a pre-usage shelf life in excess of several months or even for a year.

In actual usage, the working current may vary from the measured or calculated value due to temperature or other factors. The built prototype is proven that it has the capability to function continuously for more than 1 thousand of inhalations base on one single charge of battery.

5.4 The motion detecting wake-up system

Motion detection has been widely developed in add-on devices for adherence monitoring. For a previous device, SmartTurbo [67], consecutive opposite rotations of the base of inhaler were interpreted as inhalation. Combining the information of whether the protective cap is removed or not, the overall accuracy of detection reaches 99.2%.

As the developed prototype in this research is not only focusing on inhalation, but also concerning about the duration length and the quality of inhalation, the information of rotations is insufficient for evaluation whereas it is favourable for triggering system from the deep sleep mode as it is proven that the rotations of inhaler can successfully indicate actuation of inhalation. The system is designed to be auto-on/auto-off and will only be woken up by predefined movements of inhalers to avoid unnecessary power consumption.

The selected LIS2DH is an ultra-low-power linear 3-axis accelerometer with I2C/SPI standard output. The advantages of LIS2DH include having two independent programmable interrupt generators and being available in small thin plastic package [157]. In the low-power mode, the power consumption can be down to 2 μA. A two-phase motion detecting system is developed based on LIS2DH, including the detection of interrupts to achieve full functionality of accelerometer and the detection of position to indicate that the device is ready for detecting inhalation.

Three operating modes are provided by LIS2DH, including high-resolution mode, normal mode, and low-power mode, with 12-bit output, 10-bit output, and 8-bit output respectively. The working mode can be selected by changing the value of registers. The output data rate (ODR) may be selected from defined numbers from 1 to 5376, and the current consumption of each operating mode varies with different data rates. For the low-power mode, as the example, when the operating data rate is 1 Hz, the power consumption is the lowest 2 μA, which may increase to 10 μA with 100 Hz as the data rate. For this thesis, the resolution of an 8-bit output is adequate for detection interrupts therefore low power mode is selected for detecting both waking up shaking movements and the angle of alignment. The data rate can be selected between 10 Hz and 25 Hz, with 3 μA and 4 μA as the current consumption, respectively. In the current version, 10 Hz of data rate has been developed.

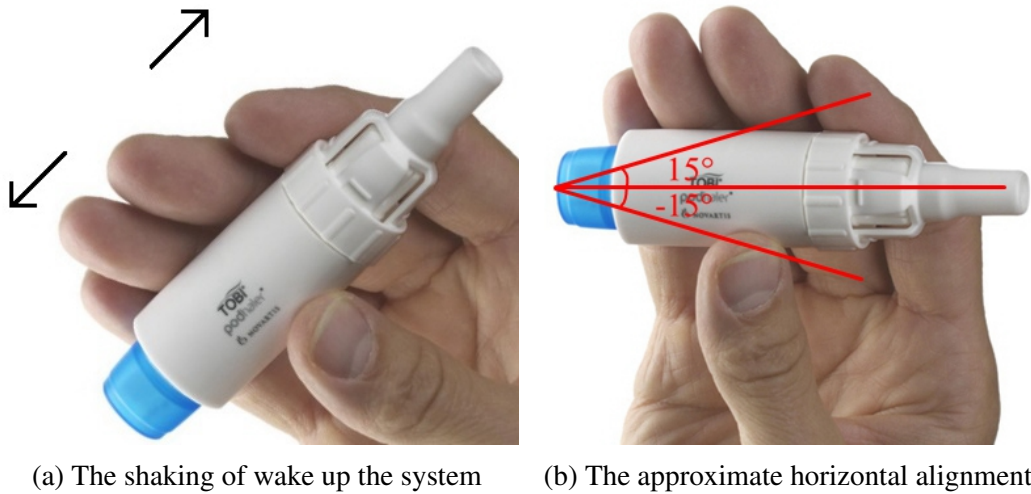The full motion detection is depicted in Fig. 5.9.

(a) The shaking of wake up the system   (b) The approximate horizontal alignment

Fig. 5.9 The two-phase motion detection before detecting inhalation

a. The detection of interrupts

During the sleep mode, the microphone is off and the accelerometer is working in the low power mode.

In the current version, the ODR is 10 Hz with 3 μA current consumption. The threshold value and duration of interrupts are preset to define the wake-up movement. The sensitivity of the system can be adjusted at any time, the currently used version equals the weight of a light shake, which is easy to be operated by users in the early stage. The threshold value of interrupts is defined by the value of the full scale (FS) and the least significant bit (LSB). In the current low power mode, FS is selected to be 2g (g presents the gravitational acceleration), thus 1 LSB equals 16 mg according to the data-sheet [157]. As the value of the quantity of LSB is set to 96, the threshold value of a pre-defined waking-up shake is 1.536g (16 mg $\times$ 96). The threshold duration of interrupts is the minimum duration of two recognised events. It is stored in another register and defined as 1/ODR, which is 0.1 second when ODR is valued 10 Hz.

To avoid unnecessary power wasting due to unwanted triggering signals, especially when inhalers are being carried on travel, the threshold value of waking sign is set high as a light

shake in the developed version. Two continuous light shakes is defined as the waking gesture, which is depicted in Fig. 5.9a. The shakes can be detected regardless of directions and the rotations of the inhaler, and the system then goes automatically into the rotation-detection mode.

b. The detection of approximate horizontal alignment

For a better medicine administering, the inhaler is required to be held horizontally during inhalation. Instead of recognising inhalation from rotations of the inhaler, the horizontal alignment is detected to active audio processing.

When the angle of the inhaler with respect to the ground is $\pm15\circ$ and continues for longer than 0.2 seconds, it is defined as being held horizontally, which is illustrated in Fig. 5.9b. The alignment detecting mode proceeds for 20 seconds, and the system goes back to sleep mode automatically without a recognised horizontal holding when the 20-second period expires. With a 0.2-second horizontal holding, the system turns to the audio processing mode where the accelerometer stops to function and the SPI is used for transmitting audio signals from the microphone.

The sensitivity of the wake-up system

The device has been tested for 500 usages, and 4 intakes occurred for each usage. For each usage, the wake-up motion has been tested once before inhalation. 91% usages were successfully activated by the first attempt of shaking, whereas 7% of usages were activated by the second try and the rest 2% required the third shaking. The probability of unsuccessful activation decreased when more shaking attempts had been done by the same person. During the final 200 usages, there was no failed activation. It was indicated that a training session may help the user to activate the device properly.

During a later user testing session with volunteers, all users had successfully activated the device following the illustrated instructions. 33% of participants did no activated it by the first attempt of shaking, but showed the capability to adapt to the system properly during the second attempt with a 100% successful rate (15/15).

The specificity of the wake-up system

To avoid unwanted activation, the threshold value has been set to be higher than a normal swing from walking, whereas a sudden speed shift from transportation or a movement from running may be recognised as the triggering signal. In a small-scale 30-minute motion detecting experiment, the device was attached to a travel bag to test the specificity of the system. The inhaler was located in the bag vertically. From the record, there was no redundant activation during the 10-minute walking, whereas 3 activation pulses had occurred during the 10-minute travel on the tram. Activation could not be avoided during strenuous exercise. It was also clarified that audio processing had not been triggered as the right angle had not been detected if the inhaler was located vertically, and thus avoided faulty activation and power wasting.

5.5   The sound sensing

5.5.1   The MEMS microphone

It is worth considering the size and power consumption of the voice sensor, whereas conventional microphones are bulky with high energy consumption. Electret condenser microphones have been used in miscellaneous applications, such as being built in telephones and computers to provide a high-quality recording [159]. Piezoelectric microphones are usually used as contact microphones, sensing audio vibrations through contacted solid objects [160]. Nev-

ertheless, the performances of traditional electret condenser and piezoelectric microphones have been improved from the earliest version, the diameters of electret condenser are usually around 9 millimetres with 5 millimetres as the height and piezoelectric microphones are flat and thin with 30 millimetres as the diameter according to the current market of electronic components. With small footprints and low power consumption, MEMS microphones can be integrated into a package with round 4 mm as the longest side.

The first selected microphone was ADMP401 [161], an omnidirectional MEMS microphone with bottom port and analogue output. The internal Analogue-to-Digital converter (ADC) inside microcontroller reads the data one after another from the analogue microphone. The appendant amplifier of the breakout board costs extra energy, and the instability and indirectness of the analogue device may complex the processing. Besides, the size of the ADMP401 breakout board was 1.4 square centimetres, which was over the requiring size. A Pulse-density modulation (PDM) microphone INMP522 [162] was later introduced into the system. The INMP522 is omnidirectional digital output MEMS with low power costing and an impedance converter amplifier followed by a fourth-order Σ-Δ modulator. Besides having a high SNR of 65 dBA and sensitivity of -26 dBFS, it also has an extended wide frequency response from 75 Hz to above 20 kHz, resulting in natural sound with high intelligibility. The usage of PDM microphone was excluded temporarily due to errors to fit the MSP430 microcontroller.

The INMP441 [132] is a high-performance low-power MEMS microphone with a digital (Inter-IC Sound) I^2S interface with high-precision 24-bit data, which is used to collect wave files in the subsequent progress. The I^2S output is a widely used digital interface for audio converters and processors and now being integrated at the edges of systems in components such as new MEMS microphones. The I^2S microphone has similar advantages as a PDM microphone, but its digital data are a decimated base-band audio sample rate output instead of a high sample rate PDM output [163]. It can connect directly to a signal processor or

microcontroller for processing with its standard interface. Its current consumption is around 1.4 mA and can be kept as 2 μA when it is powered down. With high SNR of 61 dBA, high sensitivity of -26 dBFS, and flat frequency response from 60 Hz to 15 kHz, the INMP441 fits the requirements. It is available in a thin 4.72 mm $\times$ 3.76 mm $\times$ 1 mm surface-mount package, slightly bigger than the previous candidates but in the acceptable range.

For an early prototype, the MEMS microphone was attached to the surface of the inhaler body with a small piece of foam covered on the top of it to avoid the influences of airflow. In the improved design, the integrated chip was embed and fitted to the inner side of a 3-D printed protective shell. The sound hole in the microphone package was aligned with a through-hole in the printed circuit boards (PCB) board, which allowed the vibration of air in the chamber and inhaler body to be sensed through. The performance of the microphone and the recognising system would be impeded if the through holes were obstructed partly or completely by additional material, such as glue or tapes.

5.5.2 Audio signal filtering

An initial evaluation of breath sounds detected by the microphone was carried out in a noisy environment with low SNR from 12 to 18 dB. In order to provide a reference, the recording experiments were performed in the soundproof room which yielded a data set under approximately 40 dB SNR. It was degraded by additive noise to create the training and evaluation system to select and assess the performance of appropriate feature-extraction algorithms.

The strong vibrational artefact at around 1 Hz from the control circuitry of the microphone was removed by first-order infinite impulse response (IIR) high-pass filter presented in Figure 5.10, which has been constructed in the processing c code with a simple algorithm. The cut-off frequency of the filter was designed around 15Hz with 60dB for the vibrations.
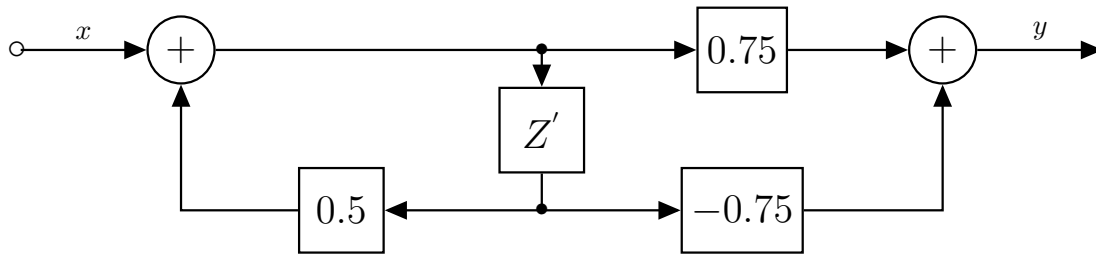
Fig. 5.10 The IIR high-pass filter before real-time feature extraction and sound identification

The coefficients were adjusted to divisible numbers to reduce computational complexity and fit into the real-time one-way processing. For a first-order filter, the transitions between the cut-off region and pass region were gentle and smooth. The -3dB frequency of constructed-in-code IIR filter, which means the ratio between the output an input was 0.707, was calculated to be around 800 Hz with the chosen coefficients. The frequency that signal power was attenuated by 50% was calculated to be around 463 Hz. Although there is no obvious pass-band as the filter is simple-constructed, the performance of filtering low-frequency signal whose frequency is lower than 100Hz has been indicated.

Fig. 5.11 depicted the filtered result of the inhalation and exhalation in Figure 3-1. The inhalations were mostly kept, and the exhalation was restrained as the sound from air vibration was filtered and reduced.

To be more specific, the original signal and the filtered signal are presented in the frequency domain to demonstrate the frequency response of the first-order high-pass filter. In Fig. 5.12, it is shown that the original sound includes the high-energy low-frequency noises from very low frequency to around 300Hz, which is massively reduced in the filtered result. The filtering performance is presented in logarithm in Fig. 5.13 for a detailed demonstration of low-frequency filtering.
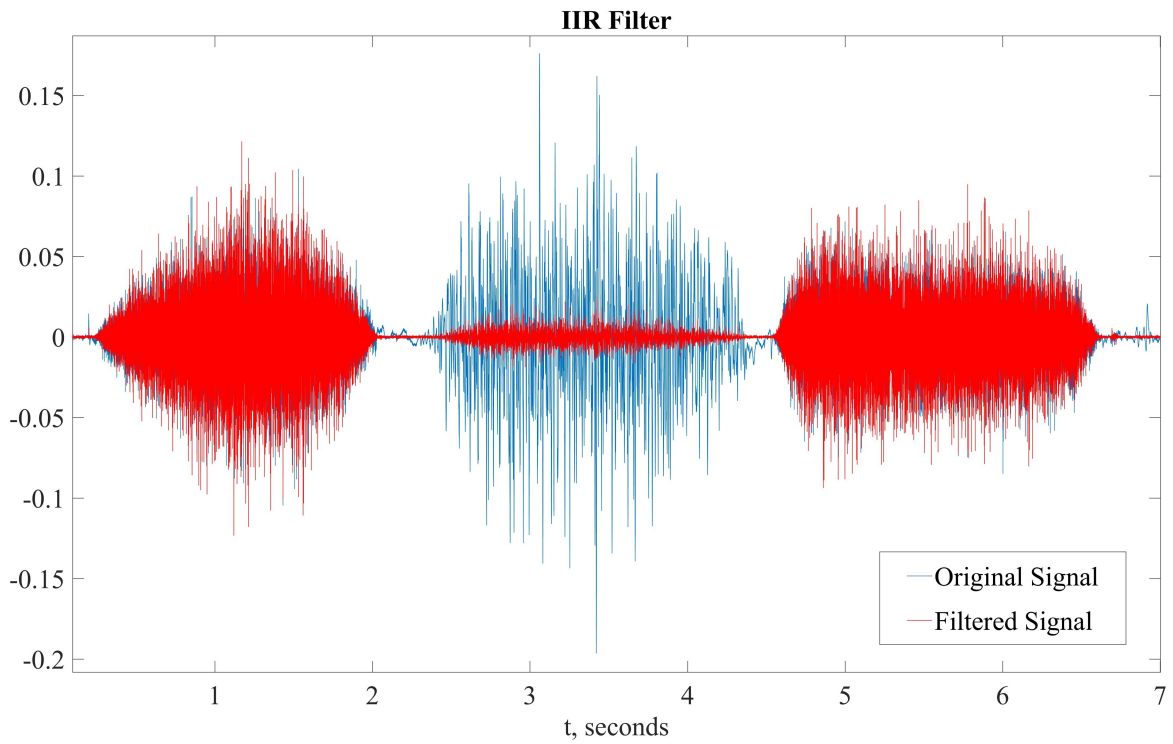
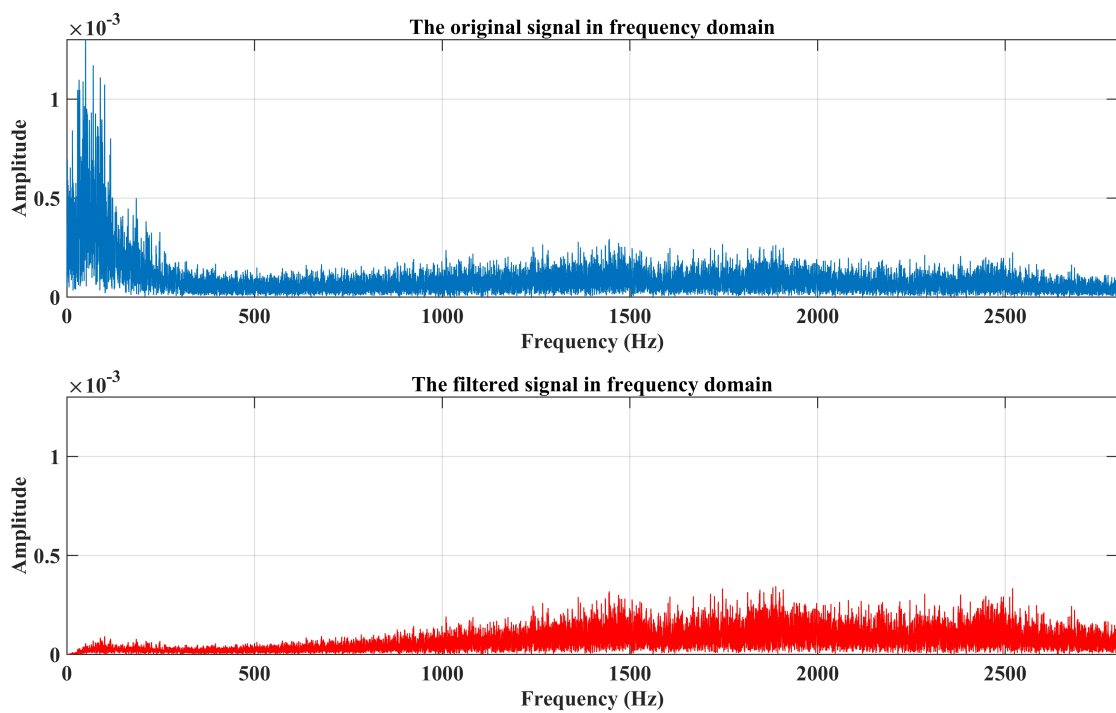Fig. 5.11 The filtered inhalation and exhalation by the constructed IIR high-pass filter

Fig. 5.12 The filtered result in frequency domain
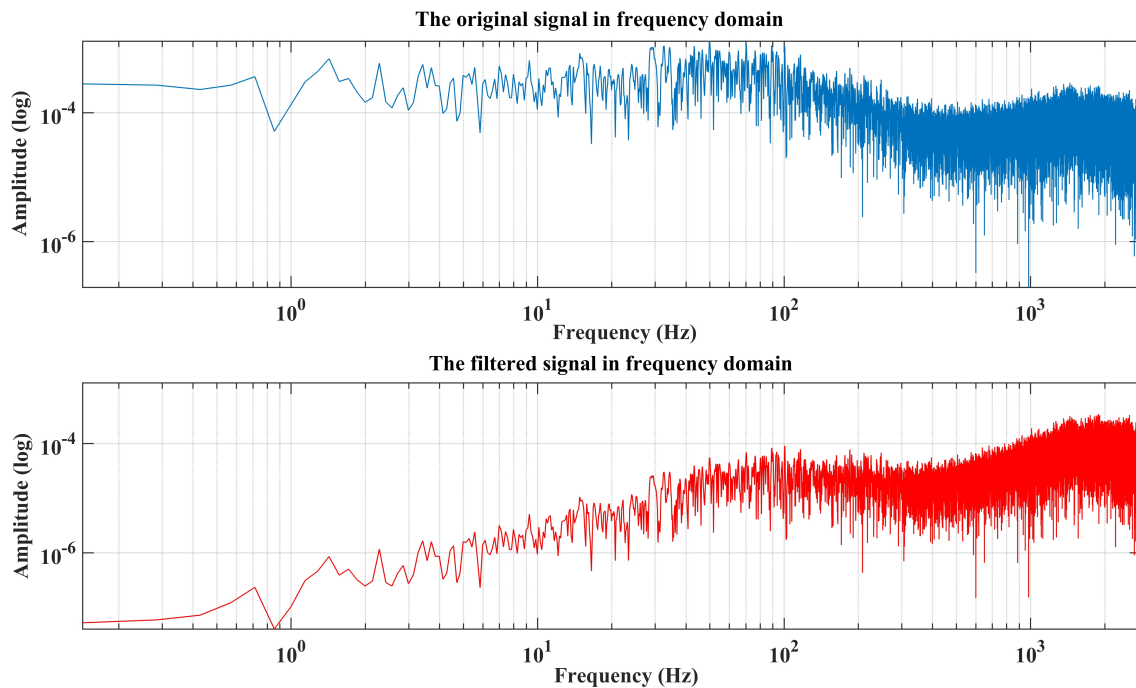
Fig. 5.13 The filtered result in frequency domain (log)

The frequency response of the constructed high-pass filter is depicted in Fig. 5.14. From the change of gain (output/input), it is clarified that the 50% attenuation occurs around 600Hz and the -3dB (0.707) frequency is 875 Hz.

As the low-frequency noises were high in amplitude and energy, the decreased volume of the filtered noise resulted in an evident decline of STMD and STE. The SNR of the filtered result had been improved in the low-frequency region. Thus, the effect from circuitry artefact and airflow vibration would both get mostly eliminated with the simple constructed-in-code filter, improving the performance of the recognition system.

5.5.3   Real-time processing on the prototype

The size of the framing window is limited to fit the finite 1KB RAM and achieve quick real-time processing. The size of processing window can be pre-set as any value from 200 to 600. At the initial stages, a segment of 200 samples is defined as one processing set and
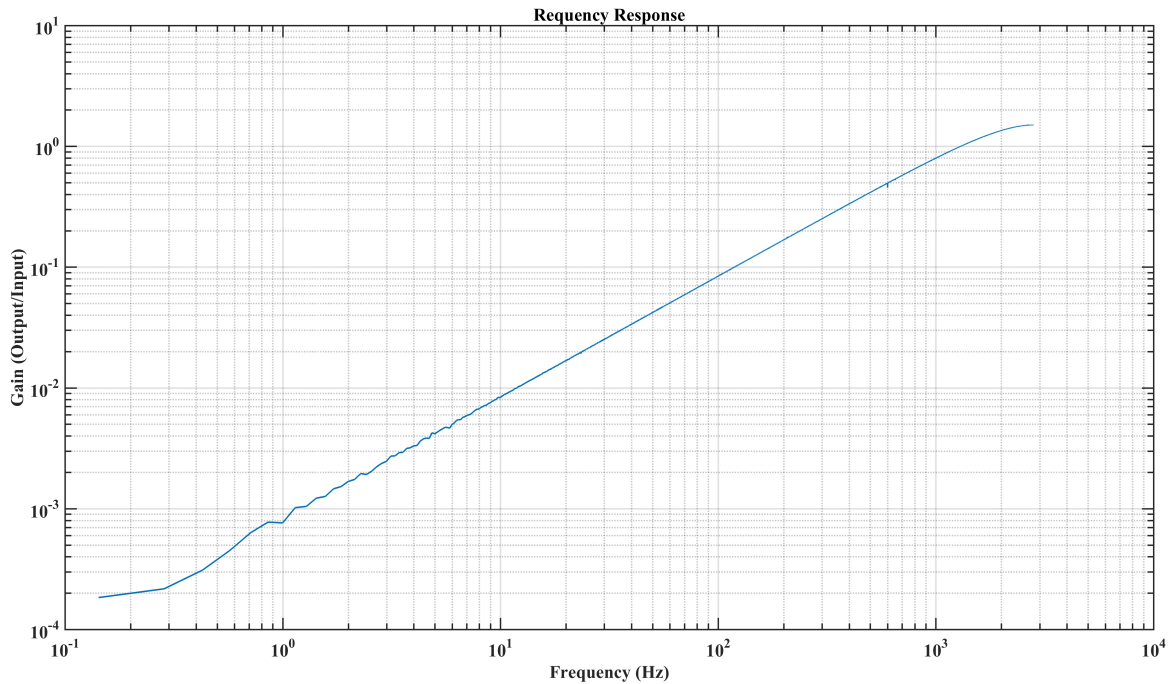
Fig. 5.14 The frequency response of the IIR filter

the size of each set can be adjusted. As the sampling rate is 5578 sample-per-second (SPS) and it is calculated by the SPI clock (357 kHz) divided by 64 bits per conversion transaction, one processing window of 200-samples takes 0.036 seconds. The size of processing window can be pre-set as any value from 200 to 600, which takes 0.1 second. The SPI clock has been defined and limited by the timer source clock, which is 1.5 Mhz. The identification of inhalation begins when 200 samples have been processed and the segments whose features are within pre-set threshold value are considered as a forming part of a clip of inhalation.

In Fig. 5.15, the developed inhalation identification in limited-size windows is depicted.

The chosen features are calculated by samples, stored in buffers, and discerned based on segments. The buffer to store an audio sample of a single transaction is a 64-byte array of unsigned 16-bit integers, and the filtered audio samples will be stored in another 64-byte array. The buffers will be renewed by the latest audio signals. A counter is used for measuring segments, the maximum value of it refers to the size of the processing window.

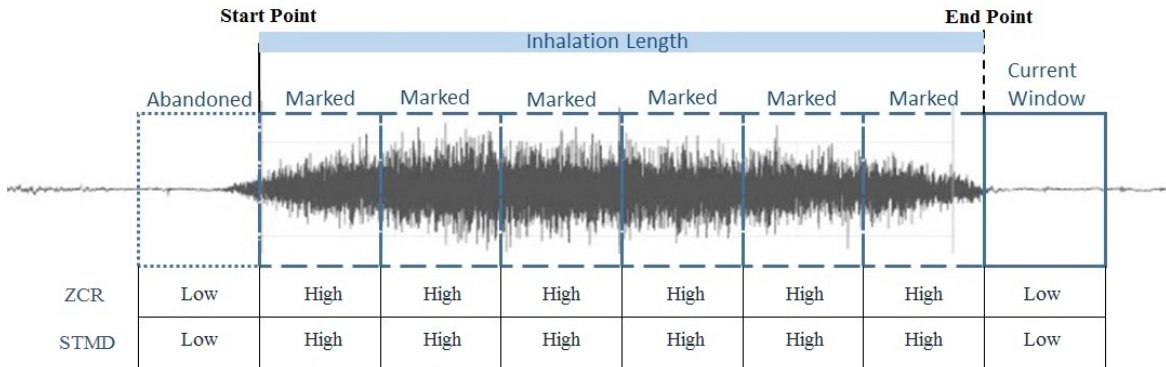Fig. 5.15 The fast inhalation identification in limited-size windows

The data are not saved on segment-based and there is no real buffer to store the 200-sample segment. Respectively, the STMD and ZCR are calculated by comparing the difference and sign-difference between the two latest samples. The absolute difference of the latest samples will be added to a signed 32-bit integer. And if the latest two samples have different signs, an unsigned 8-bit integer to store ZCR will be added by one. The processed audio buffered are discarded and renewed immediately after feature extraction and no raw audio file will be saved without continuously occupying memory.

When a byte is read from the filtered buffer, the counter adds one and it returns to 0 when it reaches 200, which means 200 samples have been read and the feature extraction of the segment is finished. The integers of stored STMD and ZCR will be compared to the pre-set threshold value to decide whether the 200-sample segment belongs to inhalation or not.

In Fig. 5.16, the flowchart of detection is presented. The inhalation_counter is the sign-variable of inhalation segments, and inactive_counter presents non-inhalation segments.

To achieve a duration measurement, the start of inhalation will be detected initially. When a non-inhalation occurs and is detected, the start of the latest segment is marked as the end of the inhalation. The length and the time of occurrence will be recorded into the FeRAM. The duration length itself is considered as the third feature for inhalation determination,
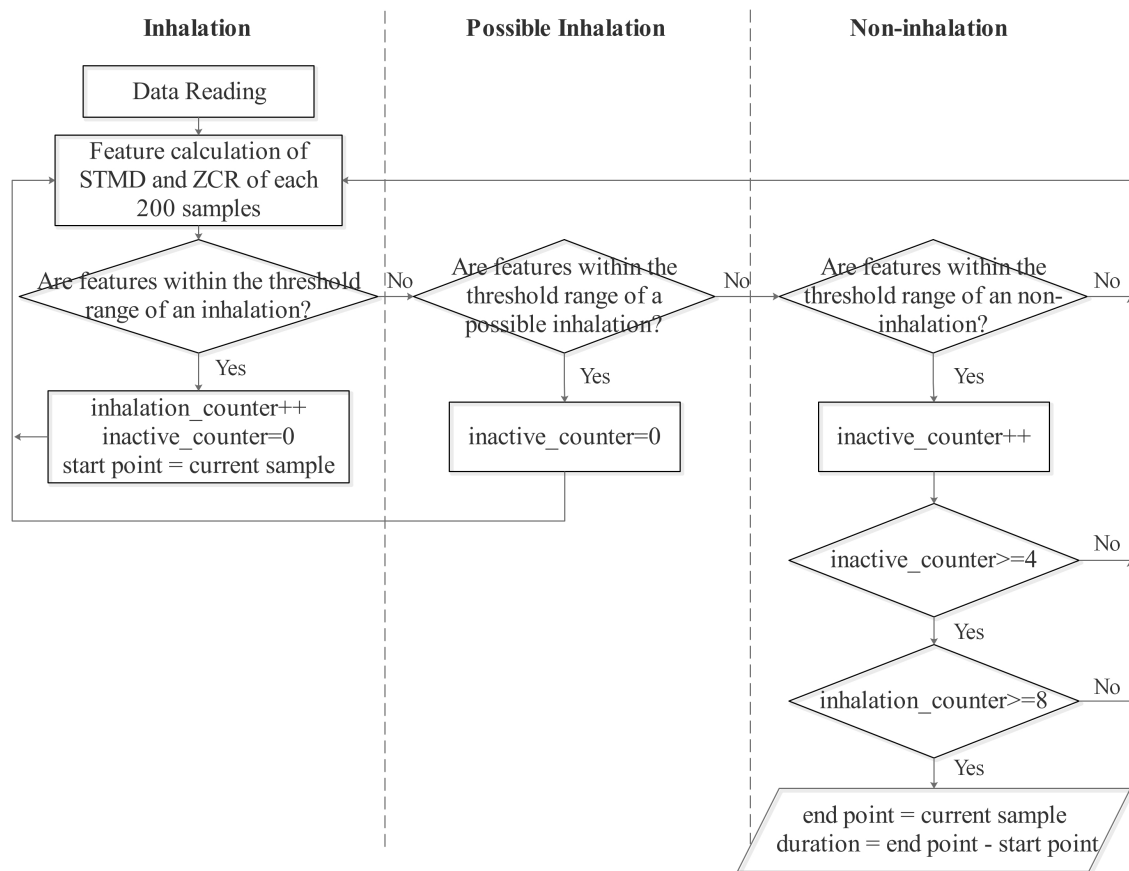
Fig. 5.16 The logic of detecting algorithm

alongside with the STMD and the ZCR. Insufficient inhalation will not be marked as a correct inhalation when the duration length is shorter than 0.6 seconds, which is the pre-set threshold length according to the trained data. The information of an insufficient inhalation will be discarded immediately. As there are algorithmic anomalous errors, such as a start without end, an end before the start, and staggered data in back-to-back events, the restrictions of inhalation detecting have been introduced to the code. The endpoint can be detected when there is one and only activated start point to avoid overlapping start-end regions.

As each 0.036-second occurs fast, and there will be around 27 processing windows during each second, the system is quick-reacted and very sensitive. In an improved inhalation length measurement, 4 continuous non-inhalations after a valid inhalation are considered as the end sign to reduce a premature end of detection. A tolerance threshold has also been introduced into the system to reduce instability from hardware and unstable usages. When the value of one sample set is not within the inhalation range but in the tolerance range, an end will not be triggered. 6 continuous non-inhalations cancel out the stored start sign of inhalation, avoiding overlapping start points.

Without complicated audio signal matrix calculation, the processing provides quick, real-time, effective inhalation identification with limited access memory.

5.6   The LED feedback system

To reduce incorrect usage cases such as short inhalation or non-inhalation and improve the quality of medicine administering, feedback system has been introduced into several designs of inhalers. The Aerochamber [164] is a spacer that helps control delivery of asthma medication, which consists of a plastic tube, a fitted mouthpiece, a valve, and a soft sealed end to hold a metered dose inhaler. An inspiratory flow indicator called Flow-Vu has been

designed to improve the quality of administering. A warning whistle will occur immediately once the patient is inhaling too strong or too quick, and thus reduces improper inhalation.

The Aerochamber is indicating air-flow physically and not being electronic-involved. With achieving electronic real-time discrimination, a feedback system that presents the result of evaluation immediately and directly to users will be beneficial for reducing inadequate inhalation. In the current version, the feedback is presented through the bi-colour LED flashing. The flashing system consists of a green LED and a red LED with various combinations of different lengths and numbers of flashing. The green LED has been used to guide the regular routine of an administering, whereas the red one is designed as the warning signal to indicate that the system is off or not has been used properly.

In Table 5.2, the meaning of each flashing style is presented and it can be customised by the user in the system setting stage. The adopted version indicates the mode changes, a good inhalation, backing to sleep, faulty usage, and 5-second countdown. When the system is activated and switches between working modes, the green LED flashes once for 0.5 seconds. When inhalation is detected and stored in the buffer, the green light flashes once for 1 second. When the system turns back to the sleep mode from any other modes, the red LED shows a long 2-second flashing. The environment can be estimated using the simple envelope which is depicted in Section 3.2.4. For pre-defined faulty usage or noisy environment, multiple flashing occurs to the red LED. To inform the user when the 5-second pause expires, the green LED flashes twice.

5.7   Incorrect usage and presented solutions

Incorrect and inadequate administering has been recorded during previous CF treatment [7]. To use the TOBI Podhaler properly, the medicine should be taken regularly in an exact dosage. Each dose of 4 capsules should be taken around every 12 hours and should not be

Table 5.2 The flashing code of LED

Meaning	Colour	Flashing mode
Mode changes	Green	Flash once, for 0.5 seconds
Good inhalation	Green	Flash once, for 1 second
Back to sleep	Red	Flash once, for 2 seconds
Faulty usage	Red	Flash twice, for 0.1 second each
Noisy environment	Red	Keep flashing, for 0.1 second each
5-second count down	Green	Flash twice, for 0.1 second each

taken less than 6 hours apart [16], which is recommended to be taken during mornings and evenings. An exhalation to the air is needed before inhalation. The inhalation should be deep and smooth with a single breath. At least two intakes empty one capsule. After inhalation, the patient needs to remove the device and hold the breath for around 5 seconds. The possible types of incorrect administering and developed solutions are presented below:

1. Inadequate inhaled medicine or deficiency in duration

A shallow inhalation may lead to an inadequate medicine administering as the powder requires airway propulsion to reach the target organ. An event of over-enthusiastic inhalation with excessive powder may also lead to an administering failure and cause respiratory problems like coughs. The inhalation needs to be steady and deep enough for a successful powder delivery. Inadequate duration length of inhalation may also lead to an unsuccessful administering as the airway may not deliver medicine to the target organ in limited time. The clinicians have the intention to collect information on duration length for further research about how long inhalation should take to achieve the best result of medicine administering.

Existing measurements analyse the recorded data and demarcate inhalation in PC or mobile phones, whereas an on-chip recognition is adopted in this research. For a real-time system, the measurement requires high-speed recognition. To avoid inadequate administering, the threshold value of STE and STMD is pre-set to define the range of a moderate inhalation.

The strength of inhalation shares a positive correlation with energy and variance of received signals, the exceeded or insufficient measured STMD of a single inhalation indicates high or low inspiratory flow rate. In the case of reducing breathing deficiency, the threshold range of duration length has also been set, and inhalation that is shorter than 1 second will be considered too short and neglected.

2. No breath holding after inhalation

5 seconds of breath holding is required to achieve better performance of antibiotic powder, whereas the needed pauses between inhalation and exhalation are possible to be forgotten. Some users found it difficult to estimate 5 seconds accurately. Watching at a watch or a clock is considered inconvenient during adherence. Besides, there are frequent occasions that the user has no time-measuring equipment around.

The green LED will flash twice after a good inhalation with in-range features and sufficient duration length, allowing users to regulate their breathing strength, speed, and duration length if the taken administering has not triggered the LED. The threshold value can be adjusted through the PC application to fit different users. The LED flashing system also contains a 5-second countdown reminder with a quick double flashing. The countdown clock starts after a good inhalation.

3. Missed dose

The treatment requires one dose in every 12 hours and 4 capsules for each dose, whereas it is reported likely that a patient, especially a young adult who stops receiving assistance from parents, may forget to take a routine administering in required period [31].

The RTC is running in an extremely-low-power mode once the system receives power supply. The date and time of patient registration are stored inside as the time reference. The time stamp will be recorded with each inhaling event to recognise any possible missed dose.

4. Noisy environment

Noise level affects the capability of identification massively for the audio processing system. The volume of inhalation is relatively low comparing to natural noises, and the spectrum of inhalation shares similarities with white noise as they are both broadband signals with high ZCR, impeding the segmentation of inhalation under noisy environments.

A pre-processing noise level evaluation system has been introduced and applied before the inhalation detection. When the noise level is high with lower than 5 dB SNR, the surrounding environment will be considered too noisy by the system and a warning sign will be presented from the LED. Normal detection will only be available when the noise level is low enough to detect inhalation accurately.

5.8   Summary

This Chapter has introduced the novel developed prototype of the inhalation detecting device with a fast real-time reacting feedback system, in order to present a low-cost, low-power, convenient adherence monitoring methodology with the potential to aid future medical study. It is an attractive approach as a small-scale energy-effective healthcare monitoring system, which has been designed with extended durability and low-power working modes. It can continue functioning with extended durability for several months with one charge to fit in a long-period treatment. The device has been built on a TI low-power microcontroller with an

omnidirectional microphone, a bicolour LED, a 3-axis accelerometer, and an NFC chip. The data were stored in FeRAM with improved non-volatility to avoid unwanted data loss.

The full process of the system includes motion detection and sound recognition. The two-phase motion detection has been used to wake the device from the pre-set deep sleep mode, which is consist of the detection of interrupts and the detection of approximate horizontal alignment. With this early design, the device can be aroused by a mild double-shake. The system showed an adaptive sensitivity and good specificity in later experiments. The system has achieved accurate real-time inhalation recognition and duration measurement based on simple time-domain features to reduce computational complexity and memory cost. With a LED direct feedback system, the result of the fast evaluation has been presented to the users. The system also has the ability to assist the inhaling process. It can help to solve some existing problems during treatment, for instance, inadequate inhaled medicine, insufficient breath-holding, missed doses, and noisy environments.

The performance of the combined use of optimal feature-set and fast real-time block processing has been introduced, and the performance has been therefore evaluated in the presence of actual users. Subsequently, with more volunteering participants involved, the performance of inhalation detection on the constructed prototype with different sensing methods as the reference will be presented, evaluated, and discussed in chapter 6.

Chapter 6

Result of Preparation Test and Initial User Testing Experiments

The chapter presents the result of inhalation monitoring using the designed prototype system in actual medicine administering. The performance of the detection was tested under different noise levels. A two-phase initial user testing study has been completed. In the first phase, 2000 inhalations have been tested. In the second phase, the inhaling events of 15 volunteering participants have been collected, detected and evaluated to gauge the performance of the monitoring system.

6.1   Preparation Test: Detection against added noise

Before actual user testing on the developed prototype, the experiments had been designed to gauge the performance of the presented methodology and features using MATLAB, which has been described in Section 3.3.4. The aim of this stage is to define and review the threshold value of each feature, including the STMD, the ZCR, and the PL. 500 recorded inhalations have been tested under different levels of SNR. Every sample was noised manually by white

noise in incremental noise level, while white noise was acoustic-characteristically similar to inhalation with high ZCR. The SNR was from 7 dB to 35 dB, adding to the original 40 dB noise from the soundproof room and the recording system.
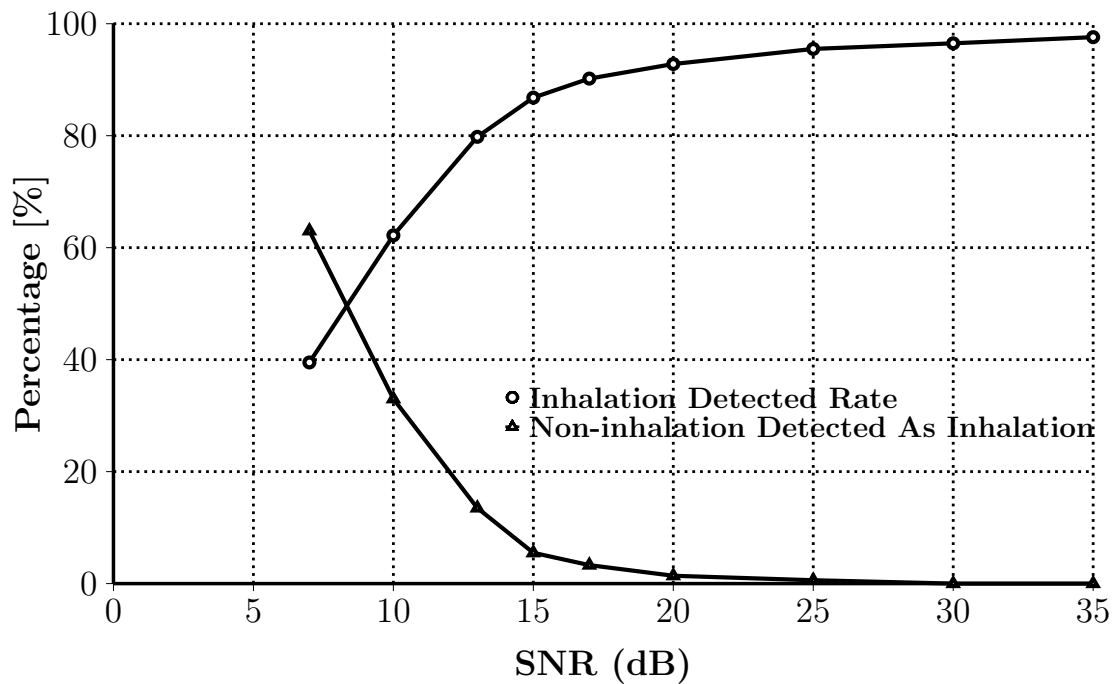
Fig. 6.1 The accuracy of detection against SNR

Fig. 6.1 demonstrates the accuracy of inhalation detection with the concept of TP and FP. More than 80% of inhalations can be correctly detected when the SNR was higher than 13 dB. When the SNR was 20 dB, 92.8% among the inhaling samples had been identified and marked correctly. The sensitivity increased when the noise level dropped, and it reached 97.6% when the SNR was 35dB. Under a noisier environment, when the SNR was around 10 dB, the STMD of noises increased and approached the threshold value of inhalation. The boundaries between inhalations and noises were blurred by the high ZCR and STMD of both signals and thus lost their detectability.

At a high noise level, a clip of continuous white noise or any type of noise with high ZCR has a higher probability to be identified as inhalation as high volume leads to an increase of

STMD. When the SNR was 10 dB, 33% of the tagged inhalations were actually white noise clips. The incorrectly detected rate was declined to 5.5% with the 15dB as the SNR. Only 0.6% of tagged inhalations were FP result when the SNR was 25dB, and no false detection when the SNR was higher than 30dB. To reduce the negative effects of the possibility to pick up similar sounds or noises and mark them as inhalations, the threshold value of inhalation duration length has been adjusted to improve the specificity.

The duration difference is defined as the time difference between the actual inhalation length and the measured duration of inhalation. No extra noise has been added to the original inhalations at this stage. Human perception has been introduced as a comparison to the computational algorithm. The recorded inhalations were imported into Audacity. A trained rater identified and demarcated each inhalation based on the visual and aural inspection by looking at the waveform and listening to audios. Among 500 inhalations, the average duration difference between rated length and detected length was 93 ± 37 (mean $\pm$ standard deviation) ms.

The same set of samples has been tested against a different type of noises. The clips of noises were collected under natural environments by the constructed prototype. The recorded samples, which were WAV files with 5578 as the sampling rate, were transmitted from the prototype to PC by a C# application through a USB cable. The noises included clips from loud Rock music, acoustic music, speeches from female and male speakers, and mechanical noises from working building site. As the natural noises were not non-stationary and the noise level was difficult to gauge, the minimum and maximum of SNR are presented in Table 6.1.

From the result presented in Table 6.1, the broadband noises or high-frequency noises with high ZCR, such as mechanical noises and whistle, showed similar influence as white noise and the level of obstruction from noises depended on their volume, with accuracy as 54.84% and 63.24% respectively. The Rock music involved different instruments and contained both

Table 6.1 The accuracy of detection under different noises

Noise Type	Noise Level	Sensitivity	Specificity	Accuracy
Loud music	7 - 25 dB	78.3%	84.6%	66.2%
Acoustic music	10 - 30 dB	94.2%	96.6%	91%
Speech (male)	15 - 35 dB	98.5%	97.9%	96.4%
Speech (female)	17 - 35 dB	96.5%	96.6%	93.2%
Mechanical Noise	10 - 20 dB	78.5%	69.6%	54.84%
whistle	10 - 30 dB	77.5%	81.6%	63.24%

high-frequency and low-frequency elements, which also declined the accuracy to 66.2%. Regardless of the influence, the used broadband noises were high-volume and rare to be seen during normal medicine delivery. The user can avoid loud music or mechanical site to reduce the loss of accuracy. The soft acoustic music or speeches, which were more common during medication treatment, showed limited influence on the detection.

6.2 Preparation Test: Detection test on the developed prototype

Before more users are taken into account for the user testing experiments, the developed prototype has been repeatedly tested by the same experimenter to review and confirm the functionality of the device.

2000 inhaling events had been taken discontinuously during a 2-month period. The 16-bit solution with the 8-bit compressed inhalation duration and stamped time were used at this stage. The inhalations were monitored by a developed prototype depicted in chapter 5. Only the sound of air-flow was tested during the experiment without drug delivery.

Table 6.2 Performance of initial user testing

Total inhalations	TP	FP	Sensitivity	Specificity	Accuracy
2000	1912	21	95.6%	98.9%	94.6%

The green LED flashing of good inhalation was monitored as the reference for successful detection during the test. A flashing after an inhaling event was considered successful and recorded as TP, and an inhalation without LED flashing was considered as unsuccessful detection and was recorded as TN. When a non-inhalation was detected by the system, and the LED flashed without a true event, a case of FP was recorded.

Base on the equations of sensitivity, specificity, and accuracy, the result of the initial testing experiment is presented in Table 6.2.

95.6% (1912/2000) inhalations had been detected correctly. One external reason for missing detection was the noisy environment. What was more frequent, the undetected inhalations were not stable enough or not with the features in the threshold range. A quick and strong event would shorten the pulse length and its STMD might be higher than the threshold range. The STMD of inhalation with soft and smooth inhaling habit was lower than normal inhalation, which indicated inadequate air-flow for effective drug delivery. There was also potential that the device had not been activated properly.

The TP detected rates of the first three hundred were 97%, 95%, and 96%, and the TP detected rates of the last three hundred were 97%, 96%, and 98%. The accuracy of detection has stayed stable without any deterioration of performance during months of usage with thousands of samples, and the stability of the system was initially guaranteed.

98.9% of 1933 recordings were true inhalations while 21 recordings were redundant noises, such as air-vibration sound from an exhaust fan, high noise-level continuous mechanical sound (with lower than 13dB SNR), and the sound of friction when the inhaler or the ring-shaped protective shelf was scrubbing or rubbing on a hard surface like a wooden table. In

Table 6.3 Categories of the detected redundant noises

Noise Type	Added manually	Detected as inhalation	Detected rate
Whistle	20	9	45%
Air-vibration	-	2	-
Friction	20	4	20%
Mechanical sound	-	6	-

Table 6.3, the categories of the 21 detected clips are presented. The sound clips of whistle and surface friction were caused manually or added intentionally with a certain caused number, and the rest were natural surrounding sound. From the result, the broadband noises were depicted to be the main causes of FP. 45% of the whistles affected the system especially when the duration continued for seconds.

On the basis of event detection, a measurement of inhalation length can present more comprehensive information to the clinicians. The inhalation length error was defined as the difference between the recorded time length and the actual time length of each inhalation. The actual time length was recorded by a reference microphone and an accessorial temperature sensing system for fidelity. The collected data was transmitted to a PC for further analysis.

Table 6.4 Duration errors of detected length

Duration Errors (second)	Sample Amount	Proportion
<0.2	396	19.80%
0.2 - 0.4	1119	55.95%
0.4 - 0.6	372	18.6%
0.6 - 0.8	86	4.30%
0.8 - 1.0	22	1.1%
>1.0	5	0.15%

Table 6.4 depicts the result of duration errors (in absolute value). From the duration distribution, 75.75% of length differences were less than 0.4 seconds. Only 5 among 2,000 samples (0.15%) involved a measuring fault of longer than 1 second time difference. There was still a

possibility of activating early endpoints during inhalation, especially when the inhalation was unstable with low-volume pulses, which was considered as non-inhalation by the system.

6.3 User tests with volunteering participants

After the preparation tests, two groups of volunteering participants have been invited to test the device. The same prototype from Section 6.2 has been used in the entire user testing experiment.

The purpose of this testing study was to evaluate the performance of the prototype monitoring system and to ensure whether it was practical and usable together with assessing its accuracy of measurement, detection, and recognition. It also informed the development and design of technology, gathered feedback from users on the device, and solicited user views on its usability. Instead of a big-group long-term trial, the test was small-scale, low-risk, and short-time with an immediate response.

The participants used an inert inhaler without any actual drug delivery or propellant. The usages were under observation during the entire test. Participants were asked to use the device for 5-8 times during a 20-minute session. The participants were given instructions by reading the user guide, which was a two-page printed manual with step-by-step description figures, with auxiliary verbal directions. The evaluation of self-guide ability was considered as one important part. At the end of the session, the participants were asked about the usages to solicit their views on the technology.

Two types of adherence reinforcement technology were considered in the user testing experiment. The first design measured and recorded the timings relating to the usage of the inhaler. The second version provided immediate feedback through the bi-coloured LED by indicating flashing light indicating the correct usage. The participant tested both versions and

solicited their feedback on the LED system, and the detecting accuracy of both technologies was compared to each other.

The tests were under observation. The behaviour and using habit of participants were monitored with an accessional record. The direct feedback signals from LED were also monitored to manifest whether the light flashed before the actual end of the inhalation.

Participants were contacted through the established research volunteering mailing list of the University of Sheffield after receiving a favourable ethics opinion, reference 016858 (Appendix B). All collected data has been kept strictly confidential. During this user testing, no drug delivery or propellant was involved. The breathing pattern was monitored during the usage without inhaling any substance. Each participant used a brand new one-off disposable 3D-printed mouthpiece to avoid any possible cross-infection risks.

During the user testing sessions with volunteers, the thermistor was used with the designed sound detecting system. The experimental setup is depicted in Fig. 6.2.
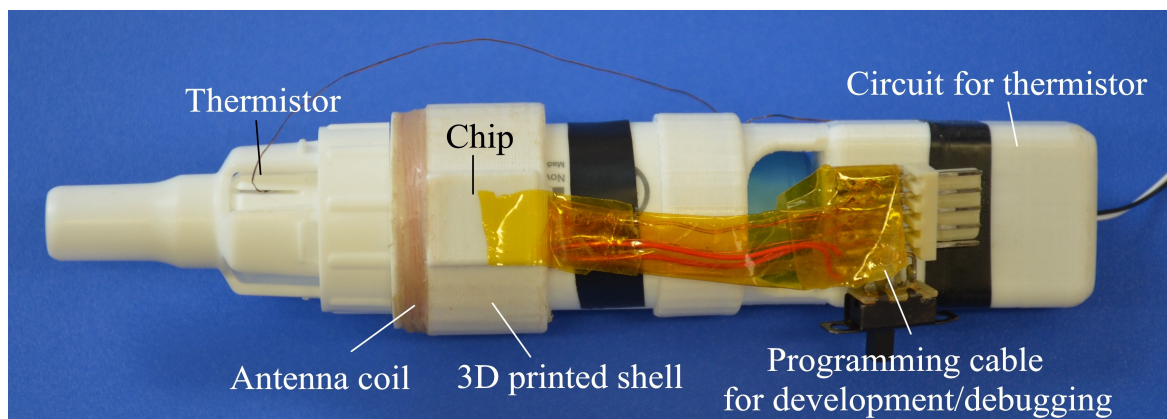

Fig. 6.2 The experimental setup for user testing experiments with volunteers. The thermistor was working along with the sound detecting system.

6.3.1   The first group of volunteers

The first group consisted of 8 volunteers who had some restriction of lung function, such as asthma or CF. The duration of each session was around 20 minutes. As the participants had been using inhalers (pMDI or DPI), each of them had own inhaling habit. The first inhalation was conducted and recorded following only printed paper instruction without further verbal instructions to observe their own inhaling habits and whether their inhalations could be detected by the system. 3 out of 8 participants were able to activate the detection in their first attempts. When an inhaling event was ignored by the system and did not be recorded, it is indicated that the inhalation involved soft or unstable inspiratory flow. Subsequently, the participants were told to inhale stronger or steadier. One participant managed to activate the device until the third attempt.

Table 6.5 Performance of initial user testing in group 1

Total Inhalations	Total Recordings	TP	Sensitivity	Sensitivity (first inhalations excluded)	FP	Specificity
51	42	42	82.4%	90.7%	0	100%

According to the result presented in Table 6.5, 82.4% (42/51) of inhalations were detected in the session. As the first inhalation was not verbally guided and only followed the own inhaling style of each participant, the detected rate of TP was 90.7% (39/43) when the first inhaling events were not included. As all the inhaling events were under observation, it was clarified that the non-detected inhalation could be too short, soft, fast, or unsteady comparing to the pre-set threshold value. No false-positive result has been recorded during the sessions. Comparing to the previous study with pre-recording system [60], which presents the sensitivity of 95%, the specificity of 94% and accuracy of 89%, the sensitivity of the developed real-time system was 4.3% lower with higher specificity. As an early-stage prototype, 9.3% as FN indicates that the system has the ability to detect inhalation.

Subsequent to the sessions, recorded data was analysed against a reference stopwatch and reference temperature sensor for fidelity and the questionnaires studied in terms of usability and practicality to inform further device refinement. The duration difference in the stage was calculated by actual duration minus recorded duration from on-chip storage. It can be defined as:

$$DurationDifference = ReferenceDuration - DetectedDuration \qquad (6.1)$$

The duration difference comparing to the stopwatch

To define the reference true inhalation length, a trained experimenter used a stopwatch to measure the duration of inhalation during the sessions. Due to the time delay of human reflection or other external factors, error existed on this methodology. To estimate the approximate duration difference of using a stopwatch, it was used by the same experimenter to measure 50 inhalation samples for training before the user testing sessions. The result was compared on the measured duration using a PC segmentation MATLAB system with recordings of same inhalations. The duration difference of 50 samples between the PC system and the stopwatch was 0.34 ± 0.22 seconds (mean $\pm$ standard deviation), which should be considered in the later analysis.

Fig. 6.3 presents the duration difference and missed detection of each participant. The actual duration was estimated based on the collected data from a stopwatch and a temperature sensor with an approximate $\pm$ 0.4 seconds duration difference.

The duration difference comparing to the thermistor

Besides the stopwatch, a 223Fu5183-63021 thermistor from ATC Semitec Ltd [165] was used as another reference of duration length. The hardware system was built on the micro-chip
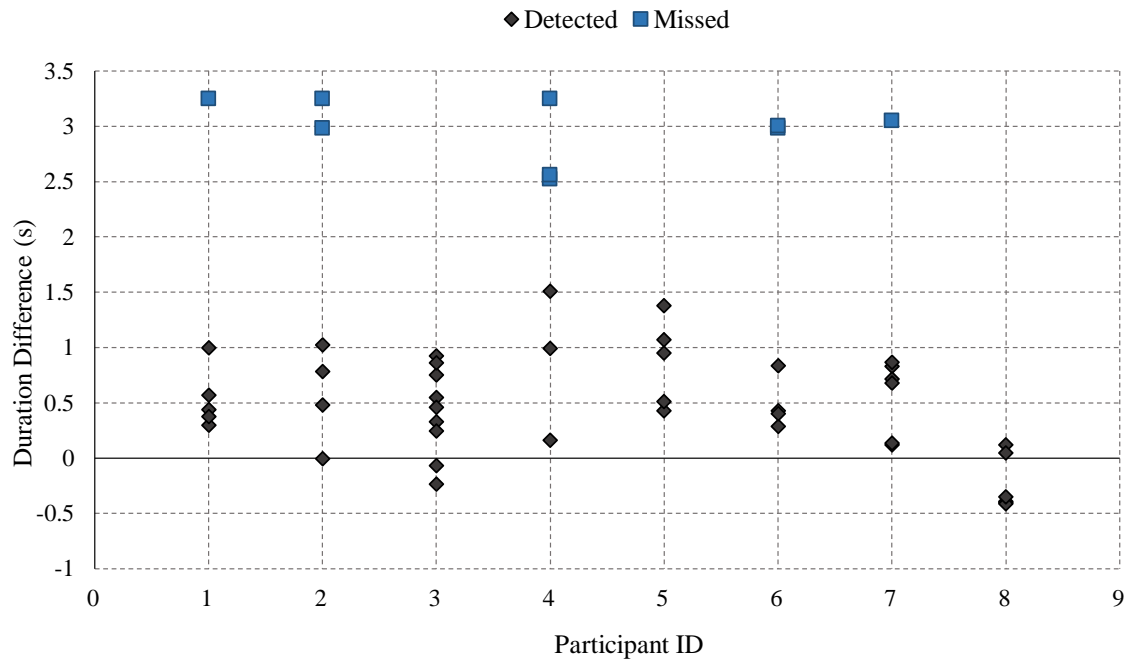
Fig. 6.3 Duration difference and missed detection of each participant in group 1

ATtiny85 [166] by Dr Tim Good. A novel temperature sensing system for demarcating inhalation was developed in this research. The selected thermistor was cylindrical with 0.51 mm as the diameter and 2.226 mm as the length. It was very tiny and able to be inserted into the chamber through the mouthpiece via an air-hole. The inhalation recognising system was developed based on temperature changing due to airflow. It was designed to be very sensitive to any temperature changing inside the chamber. An inhalation would usually cause a drop of temperature due to the air flow in room temperature from outside of the mouthpiece. On the other hand, exhalation was from the human body with a higher human temperature and would increase the temperature inside the chamber.

The changes in temperature were monitored by the sensing system and recorded as a text file. The duration length was then estimated based on the temperature changes. As the air circulation was limited inside the chamber, the internal temperature changing might be restively stable with a possible delay added to endpoints.
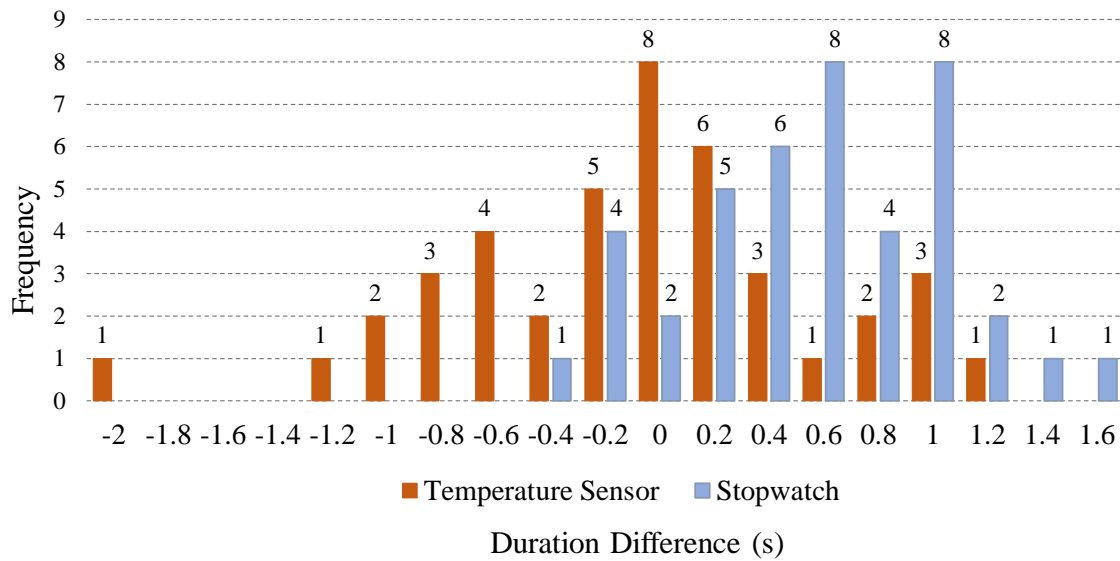
Fig. 6.4 The distribution of duration difference against temperature sensing and stopwatch of Group 1

Fig. 6.4 presents the histogram of duration difference against temperature sensing and stopwatch. The presented range is a ceiling value, which means each number represents a range from the former number to the number itself. For instance, in category 0, the duration difference of samples distributes from -0.2 to 0 seconds. And in category 0.6, samples have duration difference from 0.4 to 0.6 seconds.

According to the histogram, the duration errors against stopwatch of 42 detected samples were from -0.6 seconds to 1.6 seconds. 83.3% of detected inhalations were shorter than the actual duration, which indicated the possibility of early endpoints or delay due to using stopwatch. The accessional record clarified that during 54.05% of the inhalations, the LED flashed before the actual end of the inhalation.

For the temperature sensing system, the duration difference distributed from -2 seconds to 1.2 seconds. Comparing to the estimated result of the stopwatch, the temperature sensing was not significantly biased, which also indicated a possible delay for using the stopwatch.

From the result, the duration difference of 92.8% (39/42) of samples was limited in 1 second, and 52.3% was limited in 0.4 seconds.

6.3.2 The second group of volunteers

The second group consists of 7 volunteers. Before the sessions, the threshold value has been adjusted to avoid early detection. The threshold ranges of the ZCR and the STMD extended for 10%. Each session took around 18 minutes. The result is depicted in Table 6.6.

Table 6.6 Performance of initial user testing in group 2

Total Inhalations	Total Recordings	TP	Sensitivity	FP	Specificity	Accuracy
63	60	58	92.1%	2	96.67%	89.03%

92.1% of inhalations were correctly detected while scant inhalation, which was too mild or too short, were neglected and discarded by the system. The mild short inhalations generally continued for less than 2 seconds with less than 1-second effective medicine delivery that needed strong air-flow to reach the target organ. Undetected inhalations usually occurred to participants in the specific groups, and the odds of insufficient air-flow related to various covariates, including gender, age, and other body conditions of participants [19, 31, 18].

Two participants took the sessions in noisy environments (with averagely lower than 13dB SNR) and the continuous white noise affected the detection. Two FP inhalations have been tagged as inhalations. It was clarified that the continuous broadband noise would be detected redundantly, which may also result in a delayed end of detection and caused extended measured duration data.

Besides air-flow delivery and environment, other covariates have also affected the result of detection in the second group, including the inexperienced skill of using DPIs and misunderstanding of instructions.

Subsequent to the sessions, the result was analysed comparing to a visual and auditory demarcation using a reference microphone (with a Smartphone recording application) and the reference thermistor to test the fidelity of the system.

The duration difference comparing to the human demarcation and temperature sensing

As the average error of stopwatch measurement was still considerably high, another methodology to define the reference duration had been introduced. During sessions in the second group, the original inhalations were recorded by a mobile recording application. Only the sound clips of inhalation had been collected with the consent of each participant. The recordings were marked and time-stamped respectively and manually by the experimenter, which were later transmitted to a PC and analysed on Audacity. The endpoints of segments were recognised visually by looking at the waveform and auditorily by hearing the recordings. The system was still based on human demarcation but showed more reliability than the stopwatch as the segmentation could be repeated to reduce error and it was easier to observe and demarcate the waveform visually.

Fig. 6.5 depicts the duration difference and the faulty-detection of inhalation. Comparing to the result before feature adjustment, the number of missed detection has been declined due to the adjusted threshold value of selected features. It was clearer that missed detection and redundant detection are more likely to occur in specific cases, getting affected by environment and individual differences. For participants with a lower vital capacity, their inhalations were usually shorter and softer, with the higher possibility to be neglected or discarded. On the other hand, the false positive result is often be detected under noisy environment.

Fig. 6.6 shows the histogram of duration difference against temperature sensing and human demarcation. The measurement improves from the first test.
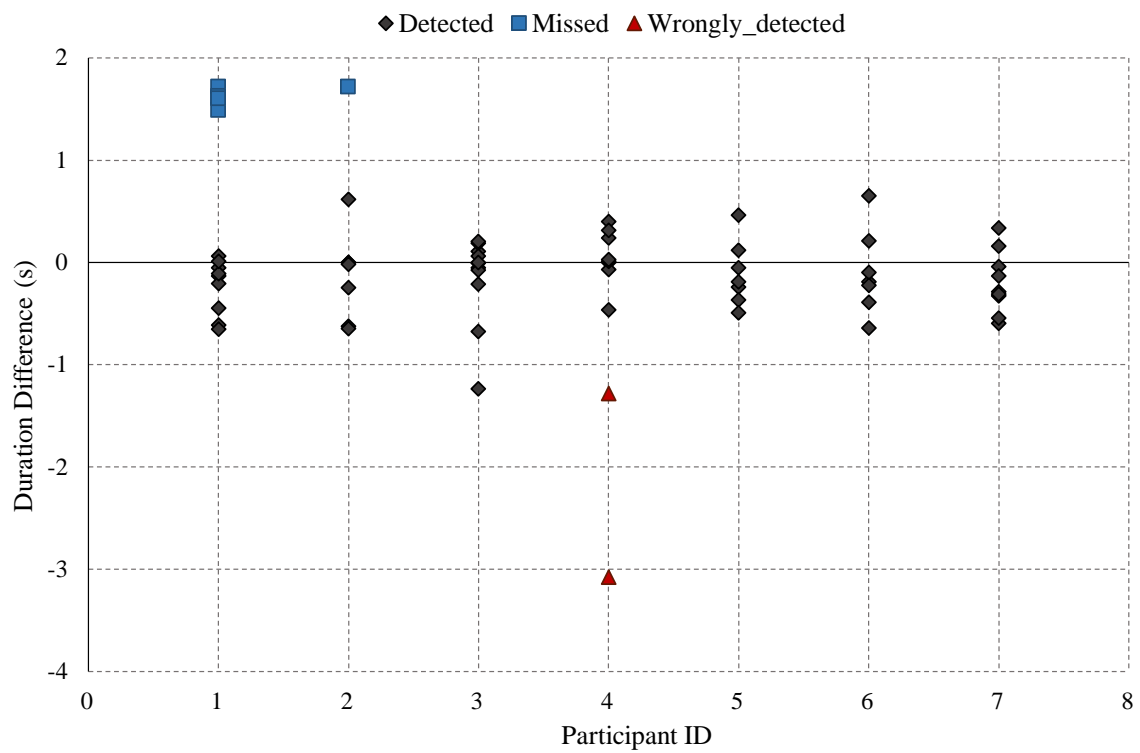
Fig. 6.5 The duration difference and missed detection of each participant in group 2
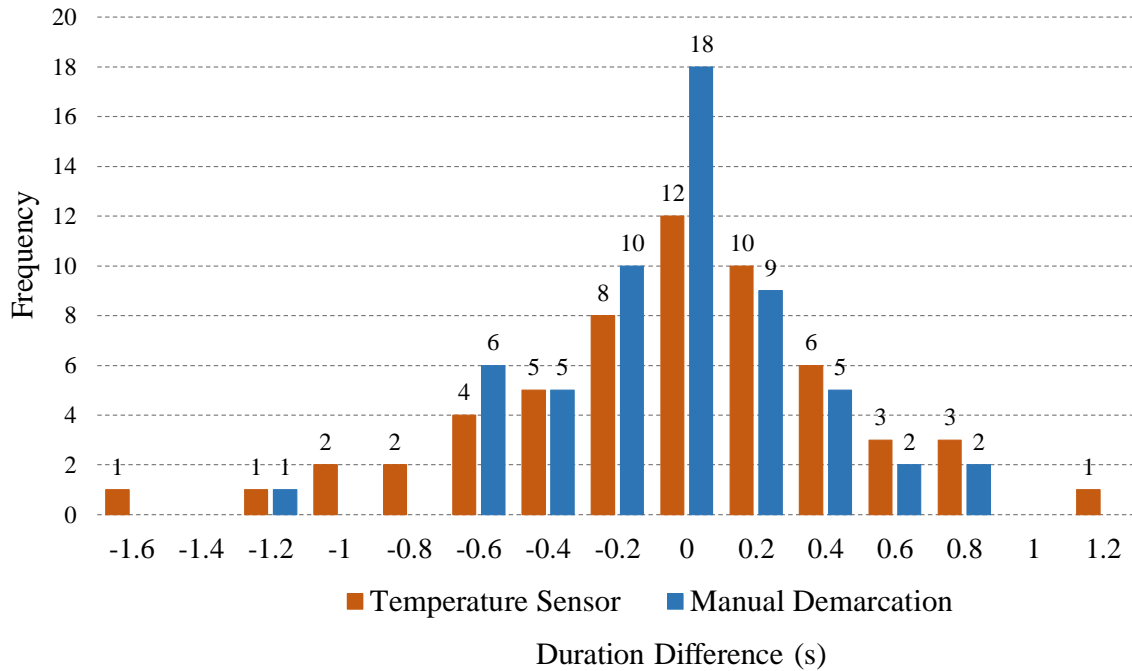
Fig. 6.6 The distribution of duration difference against temperature sensing and human demarcation of Group 2

For human demarcation, the errors were from -1.4 seconds to 0.8 seconds. 98.3% (57/58) of error differences were limited in the 1-second range, and 72.8% (42/58) of measured data had an error of fewer than 0.4 seconds. The duration difference of the temperature sensing system distributed from -1.6 seconds to 1.2 seconds. According to the shown result, the duration difference of 94.8% (55/58) of samples was limited in 1 second, and 62.1% was limited in 0.4 seconds.

The 8-bit compression used the same coding method of duration as the full 16-bit solution, and the coded data value of duration stayed the same. The accuracy of the 4-bit solution dropped slightly with higher compression and the 0.3-second resolution. As the coding method rounds down and truncates the remainder that exceeded the 0.3-second resolution, the errors contained a positive increment up to 0.3 seconds.

The duration difference against human demarcation of group 2 with the 4-bit compressed solution is demonstrated in Fig. 6.7. With the positive increment due to number truncation,
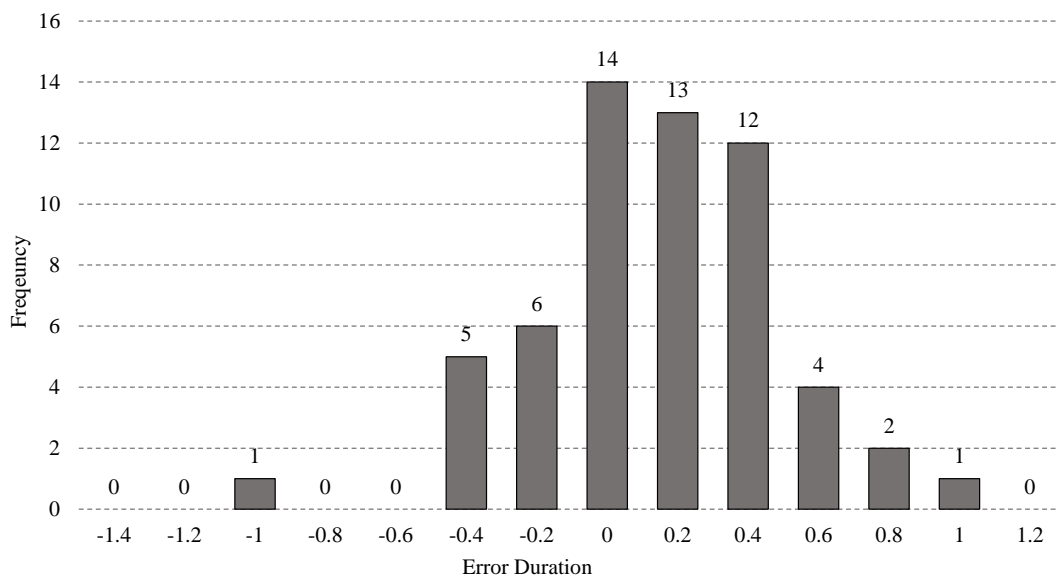
Fig. 6.7 The histogram of duration difference with 4-bit solution against human demarcation

the errors were from -1.2 seconds to 1 second. 98.3% (57/58) of error differences were still limited in the 1-second range, and 77.6% (45/58) of measured data had an error of fewer than 0.4 seconds. The average duration error increased slightly but remained in the acceptable range.

6.3.3 The feedback from participants

The size and portability

The participants agreed that the size of the device was small, lightweight, and convenient to be carried. According to the given information, some existing assistive devices, such as a chamber, might be cumbersome and not easy to be taken along. The electronic monitoring system in this research add slight extra weight and take limited space alongside the inhaler.

They also agreed that the attached electronic chip would not affect medicine delivery or block air-flow during inhalation, and the usage of inhalers was not affected by the additional monitoring system.

The functionality

86% of the participants thought the monitoring system would be useful for both asthma and CF patients while some participants thought it would be useful for CF patients but not necessary for asthma patients as the inhalers were only used when they were uncomfortable for easing the symptom in a short time. The mentioned inhalers are actually reliever inhalers, which delivers medicine directly to the lungs and relaxes the muscles surrounding airways [167]. However, very few people with asthma only need a reliever inhaler, and most of them benefit from preventer inhalers [168]. Preventer medicine helps to build up the asthma protection and should be taken as prescribed every day to prevent symptoms. The preventer inhalers are the main sources of non-adherence to asthma treatment and need to be monitored.

3 participants informed that their inhalations might not reach the standard of a good inhalation during their past usage according to the evaluation from clinicians or physicians. In these cases, the participants were happy with the monitoring system with direct feedback and they could improve the inhalation immediately without getting professional evaluation after a long-time gap.

The current version of the feedback system was designed to flash as a confirmation of a good inhalation. One participant suggested that it would be nice to use it as a sign of enough duration length and the user would know when to stop inhaling. But using the light as a finish sign may cut or shorten the inhalation. All participants in group 2 agreed that the feedback system might aid a good self-evaluation, while some participants pointed out that the users might ignore the light if they have been using it for a long period.

Inhalation without LED feedback might be inadequate when the clips were taken at the beginning of the tests. After looking at the LED, they knew what to do to activate the LED. During a short-time observed test, they could always follow the direction. The main target in the future work is to prove that the functionality will not decline during long-term non-observed tests.

The LED

13% of the participant found the LED was not easy to be observed as it was too small and too close to eyes, especially for people with limited vision. 2 participants preferred an acoustic warning signal, such as a loud whistling sound, as they were easier to be noticed.

The 5-second countdown system

93% of the participants found the 5-second breath-holding was difficult to be estimated without the countdown flashing and they agreed that the reminder was important. Some participants still pointed out that the flashing might be ignored when the patient was holding breath.

6.3.4   The comprehensive discussion of user testing experiments

During the user tests, the functionality and portability of the developed prototype were tested by 15 users (60% female) with feedback solicited. The information of inhalation pulse length (duration), date and time of each event, and identification information of each participant for distinguishing were stored in the memory of the monitoring system. The inhalations were also observed, evaluated, and recorded to an additional document by two members of the experimental group. The participants were giving direct feedback during the test, the related information was recorded for the later refinement.

The tests were conducted under multiple environments with the possibility of being influenced by surrounding noises. According to the previous analysis of detection against SNR, noisy environment did affect detection, and broadband signals might obstruct the system especially. Most user tests (10 of 15) were completed in a quiet room with a noise level of approximate 15-30 dB SNR, with exceptions that were underway under the noisier environment of 0 to 8 dB SNR in group 2. Sometimes the background noises were even louder than inhalation. Noisy environment increases the probability of FP result and extends the duration by redundant detection. The result of detection nevertheless kept its performance with an overall 89.03% of accuracy.

The result might also be influenced by the differences between personal characteristics, such as age, body condition, length of adherence, using skills, and inhaling habits, which can be inherent and not easy to be modified during the short-time experiment. The acoustic features of their inhalations varied due to the differences in air-way delivery, which was closely related to their vital capacity.

The participants were not long-time patients with CF and had not received any professional training with essential instructions for using Podhaler from clinical staff. The instructions were only told verbally with two pages of diagrams during a 2-minute pre-session, while a full professional training may be essentially required for a later trial or an actual long-term treatment. The accuracy of detection is looking for an ascent if necessary training is involved before the session.

During the preparation tests and the user testing experiments, which contains thousands of inhalation events, the system keeps its full functionality based on a single charge of the battery. After the experiments, the voltage from the power supply is still higher than 90% of the value from a full battery. The low-power processing and system design initially show the advantages to retain functionality during a long-period treatment with the ability and stability to handle and process thousands of detection.

The initial small-scale two-phase user tests focused on short-time functionality, portability, and accuracy of detection, and the tests were under direct observation with verbal instructions from trained experimental members. Therefore, practicality and stability during long-term treatment have not been proved. The potential data loss of medication adherence due to forgetfulness and unskilled usage as time goes on has not been checked in each 20-minute session. More information will be gathered in later experiments.

Chapter 7

Conclusions and recommendation of future work

The chapter summarises the performance of the constructed prototype of the novel adherence monitoring system. The developed low-cost, low-power, portable technology shows good performance in inhalation detection, discrimination and duration measurement. It presents the contribution to knowledge alongside with suggestions for possible future work.

7.1   Summaries and conclusions

The summary of the system

A comprehensive analysis of the treatment of CF requires a reliable adherence monitoring system to make sure the patient is receiving sufficient doses of prescribed regimen delivered through effective inhalation, which can be detected and measured during the usage. Therefore an efficient inhalation detecting system may aid evaluating the inspiratory flow rate of medicine delivery and providing crucial clinical information of medication dosing. The need

for a low-cost, low-power, portable technique to monitor adherence remains important. This research has focused on adherence monitoring of CF for long-term home treatment.

One main target of the thesis is analysing the usage of inhaler and demonstrating whether it has been used properly for a further study of medication effect. There are several modalities of incorrect usage such as missing doses, the excessive or scant interval between doses, low inspiratory flow rate, insufficient inspiratory duration, inhaling too fast, and using inhalers without a holding-breath pause or with an unnecessarily long pause.

With the self-printed finger-sized prototype rather than a previous RFduino board, the power consumption and financial cost are declined with a penalty in the reduction of available RAM being reduced to 1 KB. The memory space is limited for a complex calculation like cepstral or multi-dimension measurements. The data need to be kept on a small scale, allowing either the calculation or storage fitting in available resources. With the broadband characteristic of the target signal, frequency-domain features are not providing sufficient useful information to justify their high computational expense, whereas simple time-domain features are offering discernible discriminating information with lower computational effort.

Among candidate features, the ZCR, the STMD, and the pulse length contribute most to identify inhalation separated from complex background interference. The accuracy of the combination of ZCR and STMD reached 88.7%, which is the best 2-feature solution. The accuracy reached 96% when the PL was introduced to avoid interfere from short broadband noises, and short inhalation, which indicated insufficient airflow, will also be eliminated when the detected length is scant. The sensitivity of the optimal 3-feature solution is 97.5%, and the specificity is 98.5%.

The portable ring-shaped device is attached to a TOBI Podhaler, including a lightweight low-power processing prototype of adherence detection, a flat bendable battery, and the antenna coil as an external component. The device is a healthcare monitoring equipment

which can achieve real-time evaluation and record clinical information. The device can function continuously for a long period before the battery runs out and fit a 28-day treatment expectantly from the setting date.

The system contains the gesture detection mode, the inhalation detecting mode, the result storing mode, and the final data transmission mode. It stays in a sleep mode with low power consumption when the system is idle and saves power in this way. It can go back to the sleep mode from the other modes if any unexpected situation occurs to avoid faulty functionality. The data are kept in the small size to fit the limited memory and can be stored inside the device before it is exported to a terminal. The attached bi-colour LED gives out an immediate assessment of inhalation and provides the real-time direct feedback of the self-evaluation to users.

The summary of the result

The noise testing experiments and a two-phase initial user testing experiment evaluated the functionality, practicality, and applicability of the device with high accuracy of detection against different levels of noise.

In the noise testing experiments, noises of different volume were added to the recorded sample. The detected rate was higher than 80% under 13 dB SNR. 92.8% of samples were detected of 20 dB SNR, and it was 97.6% when the SNR was 35 dB. The duration difference against reference among 500 samples is 93 ± 37 (mean $\pm$ standard deviation) ms.

During the preparation tests, the tests were completed under natural environment, with random surrounding noises from 0 dB to 35 dB. More than 80% of inhalations can be correctly detected when the SNR was higher than 13 dB. When the SNR was 20 dB, 92.8% among the inhaling samples had been identified and marked correctly, and it reached 97.6% when the SNR was 35dB. Among 500 inhalations, the average duration difference of 500

inhalations between the reference from human demarcation and the detected length was 93 ± 37 (mean ± standard deviation) ms. The difference is less than 3% when the duration length is longer than 3 seconds.

2000 inhalations were tested by the same user to test the prototype. With the sensitivity rate being 95.6% and the specificity rate being 98.9%, the accuracy of the device was calculated to be 94.6%. Comparing to a previous study with pre-recording system [60], which presents the accuracy of 89%, the performance of the developed system has been initially proven to be reliable and ready to be used by different participants.

In the user testing experiments with volunteering participants, two groups of volunteers have been invited to use the monitoring system during regular inhalation. 82.4% of the inhaling events have been detected correctly under their own inhaling habit and quality of inspiratory. The detected rate of TP was 90.7% when the first non-instructed inhalations were not taken into account. The error duration of 92.8% (39/42) of samples was limited in 1 second, and 52.3% was limited in 0.4 seconds.

The features were slightly adjusted for the second group of participants to adapt to the inhaling habits of more users. With the sensitivity rate being 92.1% and the specificity rate being 96.67%, the accuracy of the device was calculated to be 89.03%, which was similar to the previous study with pre-recording system [60] with 89% as the accuracy. 98.3% of the error differences are less than 1 second, and 72.8% of them are less than 0.4 seconds. For a duration that longer than 4 seconds, the duration difference that shorter than 0.4 seconds presents 90% accuracy, which was adequate enough for the further study about whether the inhaling habit and inhalation length will affect the performance of medicine.

7.2   Suggestions for further work

1. The user testing sessions have initially proved the performance of the detection system. A long-term clinical trial is in need to further prove that the adherence monitoring system can keep its stability and functionality for a long treatment. The result and feedback will be used as part of an iterative user-centred design process, and subsequent to this stage of the research, the technology will be further refined before wider testing with the NHS patients and healthcare professionals next spring. The coding to meet medical device standards will be discussed with medical professionals and the manufacturer of devices. The improved 8-bit coding compression of inhalation length will be used in the medical experiments and trials.

2. The experiment and the processed data are the air-flow inhaling sound without any drug delivery or propellant. In the next stage, with further cooperation with clinics, the sound of actual drug delivery will be collected for feature adjustments. An iterative user-centred design process will analyse the collected data. Subsequent to user testing sessions, the technology needs to be refined before clinical testing with NHS patients and healthcare professionals.

3. After analysing the result and feedback from the volunteering participants, a self-evaluation to identify soft inhalation is considered useful. The current version abandons the incorrect inhalations without clarifying the doses are missed or not being taken properly. Although counting the left capsules may help to understand how many doses have been taken, a soft inhalation detector can recognise improper inhalation and presents the habit of users.

4. A double shake is used as the wake-up activating signal for the current version. In the future design, an automatic wake-up system would be developed to avoid the effect of forgetfulness.

References

[1] Cystic Fibrosis Trust. https://www.cysticfibrosis.org.uk/what-is-cystic-fibrosis, 2002. [Accessed 19 July 2017].

[2] Webb AK Abbott J, Dodd M. Health perceptions and treatment adherence in adults with cystic fibrosis. *Thorax*, 51(12):1233–1238, 1996.

[3] NHS choices. Cystic fibrosis. http://www.nhs.uk/conditions/cystic-fibrosis/Pages/Introduction.aspx, 2013. [Accessed 25 April 2014].

[4] Kesser KC Geller DE. The i-neb adaptive aerosol delivery system enhances delivery of α1-antitrypsin with controlled inhalation. *Journal of Aerosol Medicine and Pulmonary Drug Delivery*, 23(Suppl 1):S–55–S–59, 2010.

[5] T. F. Boat R. E. Wood and C. F. Doershuk. State of the art cystic fibrosis. *Amer. Rev. Resp. Dis.*, 113:833–878, 1976.

[6] Steven D Freedman Brian P O'Sullivan. Cystic fibrosis. *The Lancet*, 373(9678):1891–1904, 2009.

[7] Winefield HR Kettler LJ, Sawyer SM. Determinants of adherence in adults with cystic fibrosis. *Thorax*, 57:459–464, 2002.

[8] Goodacre L. Sutton C. Pollard K. Conway S. Daniels, T. and D Peckham. Accurate assessment of adherence: self-report and clinician report vs electronic monitoring of nebulizers. *Chest*, 140(2):425–432, 2011.

[9] Colin Tidy. Nebulisers in general practice. https://patient.info/doctor/nebulisers-in-general-practice, 2016. [Accessed July 2017].

[10] Usmani OS Bonini M. The importance of inhaler devices in the treatment of copd. *COPD Res Practice*, 31:1–9, 2015.

[11] R. Dal Negro S. Pedersen A. Magnan J. Seidenberg P.J. Barnes J.C. Virchow, G.K. Crompton. Importance of inhaler devices in the management of airway disease. *Respiratory Medicine*, 102(1):10–19, 2008.

[12] Innes J. A. Lenney, L. and G. K. Crompton. Inappropriate inhaler use: assessment of use and patient preference of seven inhalation devices. *Respiratory Medicine*, 94:496–500, 2000.

[13] Crompton GK Paterson IC. Use of pressurised aerosols by asthmatic patients. *British Medical Journal*, 1(6001):76–77, 1976.

[14] Copley Scientific. Spray force and plume temperature testing (cold freon). http://www.copleyscientific.com/home/inhaler-testing/special-applications/spray-force-and-plume-temperature-testing-cold-freon, 2017. [Accessed 20 March 2017].

[15] Novartis Pharmaceuticals Corporation. http://www.tobipodhaler.com, 2017. [Accessed 24th July 2017].

[16] Novartis Pharmaceuticals Corporation. Patient information and instructions for use. http://www.tobipodhaler.com/info/about/how-to-use-tobi-podhaler.jsp, 2018. [Accessed: December 2016].

[17] Drugs.com. Tobi podhaler prices, coupons and patient assistance programs. https://www.drugs.com/price-guide/tobi-podhaler, 2019.

[18] Lask B. Understanding and managing poor adherence in cystic fibrosis. *Pediatr Pulmonol*, Suppl 16:260–1, 1997.

[19] Riekert KA Eakin MN. The impact of medication adherence on lung health outcomes in cystic fibrosis. *Current opinion in pulmonary medicine*, 19(6):687–691, 2013.

[20] et al Holmes, M. S. A method of estimating inspiratory flow rate and volume from an inhaler using acoustic measurements. *Physiological Measurement*, 34(8):903–914, 2013.

[21] Donald Meichenbaum and Dennis C. Turk. *Facilitating treatment adherence: A practitioner's guidebook*. Springer, 01 1987.

[22] Maryna Marynchenko Pooja A. Chopra James Signorovitch Yana Yushkina Kristin A. Riekert Alexandra L. Quittner, Jie Zhang. Pulmonary medication adherence and health-care use in cystic fibrosis. *Chest*, 146(1):142–151, 2014.

[23] MacHale E et al Killane I, Sulaiman I. Predicting asthma exacerbations employing remotely monitored adherence. *Health care Technology Letters*, 3(1):51–55, 2016.

[24] Alicia O'Cathain Martin Wildman. Development and evaluation of an intervention to support adherence to treatment in adults with cystic fibrosis. https://www.sheffield.ac.uk/polopoly_fs/1.652514!/file/ACtifPilotProtocol.pdf, 2016. [Accessed: 1st March 2018].

[25] Liu Chun-Lin. A tutorial of the wavelet transform. http://disp.ee.ntu.edu.tw/tutorial/WaveletTutorial.pdf, February 2010.

[26] Manish Soni and Padma Kunthe. A general comparison of fft algorithms. https://pdfs.semanticscholar.org/0ae7/232f983a6356ebf10ccd44193b659dad6824.pdf, 2011.

[27] Villa KF et al Nasr SZ, Chou W. Adherence to dornase alfa treatment among commercially insured patients with cystic fibrosis. *J Med Econ*, 16:801–808, 2013.

[28] Academic Directorate of Respiratory Medicine. Prestigious nihr programme grant awarded to sheffield cystic fibrosis research team. http://www.lungsheffield.org/news/prestigious-nihr-programme-grant-awarded-to-sheffield-cystic-fibrosis-research-team, 2018. [Accessed: 1st March 2018].

[29] P. R. Byron. Drug delivery devices: issues in drug development. *Proc Am Thorac Soc*, 1:321–328, 2004.

[30] Hamnett T et al Conway SP, Pond MN. Compliance with treatment in adult patients with cystic fibrosis. *Thorax*, 51:29–33, 1996.

[31] Reid AJ Horne R Shields MD McElnay JC Goodfellow NA, Hawwa AF. Adherence to treatment in children and adolescents with cystic fibrosis: a cross-sectional, multi-method study investigating the influence of beliefs about treatment and parental depressive symptoms. *BMC Pulmonary Medicine*, 15:43, 2015.

[32] L. A. Leung J. R. Slagle, S. M. Finkelstein and W. J. Warwick. Monitor: an expert system that validates and interprets time-dependent partial data based on a cystic fibrosis home monitoring program. *IEEE Transactions on Biomedical Engineering*, 36(5):552–558, May 1989.

[33] L. M. Verbrugge. Home diaries. *Med. Care*, 18:73–95, 1980.

[34] W. J. Warwick et al. S. M. Finkelstein, J. R. Budd. Home monitoring in cystic fibrosis: A model program. *Proc. IEEE/Eng Med. Bio. Soc.*, 7:657–659, 1985.

[35] Koleva Y et al Shi L, Liu J. Concordance of adherence measurement using self-reported adherence questionnaires and medication monitoring devices. *Pharmacoeconomics*, 28:1097–1107, 2010.

[36] G. A. Gony. Computer-assisted clinical decision making. *Meth. Inform. Med.*, 12:45–55, 1973.

[37] Watson A Hamnett T Conway SP, Pond MN. Knowledge of adult patients with cystic fibrosis about their illness. *Thorax*, 51(1):34–38, 1996.

[38] Carlin JB et al Starr M, Sawyer SM. A novel approach to monitoring adherence to preventive therapy for tuberculosis in adolescence. *Journal of Paediatrics and Child Health*, 35(4):350–4, Aug 1999.

[39] A. Lay-Ekuakille, G. Griffo, R. Morello, C. De Capua, and F. Spano. Sensing system for cystic fibrosis: Modeling the detection and characterization of sweat. In *2017 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, pages 287–291, May 2017.

[40] R. Fallahzadeh, B. Minor, L. S. Evangelista, D. J. Cook, and H. Ghasemzadeh. Demo abstract: Mobile sensing to improve medication adherence. In *2017 16th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 279–280, April 2017.

[41] Timothy L Robertson Greg D Moon Kit-Yee Au-Yeung Mark J Zdeblick George M Savage Hooman, Hafezi. An ingestible sensor for measuring medication adherence. *IEEE transactions on bio-medical engineering*, 62(1):99–109, 2015.

[42] A. Dua, W. A. Weeks, A. Berstein, R. G. Azevedo, R. Li, and A. Ward. An in-vivo communication system for monitoring medication adherence. In *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, March 2017.

[43] R. K. Rosen P. R. Chai and E. W. Boyer. Ingestible biosensors for real-time medical adherence monitoring: Mytmed. In *2016 49th Hawaii International Conference on System Sciences (HICSS), Koloa, HI*, pages 3416–3423, 2016.

[44] X. Zhang and C. Xu. Real-time endpoint detection of upper limb movement based on energy threshold and residual methods. In *2012 5th International Conference on BioMedical Engineering and Informatics*, pages 334–338, Oct 2012.

[45] R. Fallahzadeh N. Hezarjaribi and H. Ghasemzadeh. A machine learning approach for medication adherence monitoring using body-worn sensors. In *2016 Design, Automation & Test in Europe Conference & Exhibition (DATE), Dresden*, pages 842–845, 2016.

[46] D. Berry, J. Bell, and E. Sazonov. Detection of cigarette smoke inhalations from respiratory signals using decision tree ensembles. In *SoutheastCon 2015*, pages 1–4, April 2015.

[47] D. DeMeo and M. Morena. Medication adherence using a smart pill bottle. In *11th International Conference & Expo on Emerging Technologies for a Smarter World (CEWIT)*, pages 1–4, 2014.

[48] L. Gradinarsky and T. Lööf. Inhalation adherence monitoring using smart electonic add-on device: Accuracy evalutaion using simulated real-life test program. In *2014 4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH)*, pages 145–147, Nov 2014.

[49] S. Ammouri G. A. Bilodeau. Monitoring of medication intake using a camera system. *J. Med.Syst.*, 35(3):377–389, 2011.

[50] D. Batz, M. Batz, N. da Vitoria Lobo, and M. Shah. A computer vision system for monitoring medication intake. In *The 2nd Canadian Conference on Computer and Robot Vision (CRV'05)*, pages 362–369, May 2005.

[51] D.M.Meunier H.H.Huynh, J.Sequeira. Real time detection, tracking and recognition of medication intake. *J. of Computer, Electrical, Automation, Control and Information Engineering*, 3(12):2801–2808, 2009.

[52] V. Moshnyaga, M. Koyanagi, F. Hirayama, A. Takahama, and K. Hashimoto. A medication adherence monitoring system for people with dementia. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 000194–000199, Oct 2016.

[53] W. Zhang, J. Han, and S. Deng. Abnormal heart sounds detection based on the scaled time-frequency representation and feature selection. In *2016 Computing in Cardiology Conference (CinC)*, pages 1177–1180, Sept 2016.

[54] T. Rosenwein, E. Dafna, A. Tarasiuk, and Y. Zigel. Detection of breathing sounds during sleep using non-contact audio recordings. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1489–1492, Aug 2014.

[55] A. I. S. M. Ayu and K. K. Karyono. Audio detection (audition): Android based sound detection application for hearing-impaired using adaboostm1 classifier with reptree weaklearner. In *2014 Asia-Pacific Conference on Computer Aided System Engineering (APCASE)*, pages 136–140, Feb 2014.

[56] Y. Castillo, D. Blanco-Almazán, J. Whitney, B. Mersky, and R. Jané. Characterization of a tooth microphone coupled to an oral appliance device: A new system for monitoring osa patients. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1543–1546, July 2017.

[57] I. Frederix, S. Sankaran, K. Coninx, and P. Dendale. Mobileheart, a mobile smartphone-based application that supports and monitors coronary artery disease patients during rehabilitation. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 513–516, Aug 2016.

[58] S. Aras, M. Öztürk, and A. Gangal. Endpoint detection of lung sounds for electronic auscultation. In *2016 39th International Conference on Telecommunications and Signal Processing (TSP)*, pages 405–408, June 2016.

[59] T. A. McCartan, T. E. Taylor, I. Sulaiman, R. W. Costello, and R. B. Reilly. Changes in inhaler inhalation acoustic features during induced bronchoconstriction: A pilot study. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3749–3752, Aug 2016.

[60] M. S. Holmes, M. Le Menn, S. D'Arcy, V. Rapcan, E. MacHale, R. W. Costello, and R. B. Reilly. Automatic identification and accurate temporal detection of inhalations in asthma inhaler recordings. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2595–2598, Aug 2012.

[61] A. A. Shkel and E. S. Kim. Wearable low-power wireless lung sound detection enhanced by resonant transducer array for pre-filtered signal acquisition. In *2017 19th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*, pages 842–845, June 2017.

[62] S. Nousias, J. Lakoumentas, A. Lalos, D. Kikidis, K. Moustakas, K. Votis, and D. Tzovaras. Monitoring asthma medication adherence through content based audio classification. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–5, Dec 2016.

[63] J. A. Tenreiro Machado C. M. Ionescu and R. De authorser. Analysis of the respiratory dynamics during normal breathing by means of pseudophase plots and pressure–volume loops. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 43(1):53–62, Jan 2013.

[64] M. M. Rahman, A. A. Ali, A. Raij, M. Al'absi, E. Ertin, and S. Kumar. Demo abstract: Online detection of speaking from respiratory measurements collected in the natural environment. In *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks*, pages 137–138, April 2011.

[65] GoSouthernMD. I-neb aad system. https://www.gosouthernmd.com/store/store/38219-I-neb-AAD-System, 2018.

[66] R. Mian S. Alsalman, R. ur Rasool. Low-power computing. *Asia Pacific Journal of Contemporary Education and Communication Technology*, 2(3):39–50, 2016.

[67] L. R. Rabiner and M. R. Sambur. An algorithm for determining the endpoints of isolated utterances. *American Telephone and Telegraph Company*, 54(2), 1975.

[68] H. F. Silverman and N. R. Dixon. A parametrically controlled spectral analysis system for speech ieee trans. *On Acoustics, Speech and Signal Processing*, ASSP-22(5):362–381, 1975.

[69] R. W. Schafer and L. R. Rabiner. Parametric representations of speech. *Proc. IEEE Specch Recognition Symposium*, 1974.

[70] J. C. Junqua, B. Mak, and B. Reaves. A robust algorithm for word boundary detection in the presence of noise. *IEEE Transactions on Speech and Audio Processing*, 2(3):406–412, July 1994.

[71] PS Vikhe NN Lokhande, DNS Nehe. Voice activity detection algorithm for speech recognition applications. In *International Conference in Computational Intelligence (ICCIA) 2011 Proceedings published in International Journal of Computer Applications (IJCA)*, 2011.

[72] Y. Sun and J. Zhou. Research on algorithm of segment and classification of audio in broadcast. In *2017 10th International Symposium on Computational Intelligence and Design (ISCID)*, volume 1, pages 316–319, Dec 2017.

[73] Rosenberg A. E et al Lamel L, Rabiner L. An improved endpoint detector for isolated word recognition. *IEEE Transactions on Acoustics Speech & Signal Processing*, 29(4):777–785, 1981.

[74] T. Zhang, H. Huang, L. He, and M. Lech. A robust speech endpoint detection algorithm based on wavelet packet and energy entropy. In *Proceedings of 2013 3rd International Conference on Computer Science and Network Technology*, pages 1050–1054, Oct 2013.

[75] S. E. Bou-Ghazale and K. Assaleh. A robust endpoint detection of speech for noisy environments with application to automatic speech recognition. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 4, pages IV–3808–IV–3811, May 2002.

[76] R. W. Wall. Simple methods for detecting zero crossing. In *IECON'03. 29th Annual Conference of the IEEE Industrial Electronics Society (IEEE Cat. No.03CH37468)*, volume 3, pages 2477–2481 Vol.3, Nov 2003.

[77] F. Bond, C. Cahn, and J. Hancock. A relation between zero-crossings and fourier coefficients for bandwidth-limited functions (corresp.). *IRE Transactions on Information Theory*, 6(1):51–52, March 1960.

[78] E. C. Titchmarsh. The zeros of certain integral functions. *Proceedings of the London Mathematical Society*, s2-25:283–302, 1926.

[79] R. G. Amado and J. V. Filho. Pitch detection algorithms based on zero-cross rate and autocorrelation function for musical notes. In *2008 International Conference on Audio, Language and Image Processing*, pages 449–454, July 2008.

[80] N Freebody and J Watton. A time-encoded signal-processing approach to fault classification of an electrohydraulic pressure control system and its application to a hot steel strip rolling mill. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 213(5):407–426, 1999.

[81] TESPAR Shannon and Approximation Strategies. The zeros of certain integral functions. *ICSPAT 98*, 18:445–453, 1999.

[82] H. Bořil and P. Pollák. Direct time domain fundamental frequency estimation of speech in noisy conditions. In *2004 12th European Signal Processing Conference*, pages 1003–1006, Sept 2004.

[83] F. Jin and F. Sattar. A new automated approach for identification of respiratory sounds. In *2007 IEEE International Conference on Multimedia and Expo*, pages 356–359, July 2007.

[84] A. Mondal and H. Tang. Respiratory sounds classification using statistical biomarker. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2952–2955, July 2017.

[85] C. D. Papadaniil and L. J. Hadjileontiadis. Efficient heart sound segmentation and extraction using ensemble empirical mode decomposition and kurtosis features. *IEEE Journal of Biomedical and Health Informatics*, 18(4):1138–1152, July 2014.

[86] M. Ross, H. Shaffer, A. Cohen, R. Freudberg, and H. Manley. Average magnitude difference function pitch extractor. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 22(5):353–362, October 1974.

[87] T. Shimamura and H. Kobayashi. Weighted autocorrelation for pitch extraction of noisy speech. *IEEE Transactions on Speech and Audio Processing*, 9(7):727–730, Oct 2001.

[88] T. E. Tremain. The government standard linear predictive coding algorithm: Lpc-10. *Speech Technology Magazine*, page 40–49, April 1982.

[89] W. Zhang, G. Xu, and Y. Wang. Pitch estimation based on circular amdf. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages I–341–I–344, May 2002.

[90] H. Tanaka M. S. Rahman and T. Shimamura. Pitch determination using aligned amdf. In *Proc. INTERSPEECH 2006*, pages 1714–1717, 2006.

[91] G. Muhammad. Noise robust pitch detection based on extended amdf. In *2008 IEEE International Symposium on Signal Processing and Information Technology*, pages 133–138, Dec 2008.

[92] Muhammad G Abdullah-Al-Mamun K, Sarker F. A high resolution pitch detection algorithm based on amdf and acf. *Journal of Scientific Research*, 1(3):508–515, 2009.

[93] Kobayashi T. Abe, T. and S. Imai. The if spectrogram: A new spectral representation. *Proc. ASVA 97*, page 423–430, 1997.

[94] Hugo L. Rufiner Leandro D. Vignolo, Diego H. Milone. Geneticwavelet packets for speech recognition. *Expert Systems with Applications*, 40:2350–2359, 2013.

[95] Z. Yessenbayev. Robust segmentation of speech signal using mfcc and acoustic parameters. In *2012 Sixth Asia Modelling Symposium*, pages 103–108, May 2012.

[96] Y. Zhang, K. Wang, and B. Yan. Speech endpoint detection algorithm with low signal-to-noise based on improved conventional spectral entropy. In *2016 12th World Congress on Intelligent Control and Automation (WCICA)*, pages 3307–3311, June 2016.

[97] B. Ayoub, K. Jamal, and Z. Arsalane. Gammatone frequency cepstral coefficients for speaker identification over voip networks. In *2016 International Conference on Information Technology for Organizations Development (IT4OD)*, pages 1–5, March 2016.

[98] A. Kyriakides, C. Pitris, A. Fink, and A. Spanias. Isolated word endpoint detection using time-frequency variance kernels. In *2011 Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, pages 1041–1045, Nov 2011.

[99] Kamil Aida-zade, Cemal Ardil, and Samir Rustamov. Investigation of combined use of mfcc and lpc features in speech recognition systems. *Signal Processing*, 01 2007.

[100] J. S. Erkelens and P. M. T. Broersen. On the statistical properties of line spectrum pairs. In *1995 International Conference on Acoustics, Speech, and Signal Processing*, volume 1, pages 768–771 vol.1, May 1995.

[101] Lie Lu, Hong-Jiang Zhang, and Hao Jiang. Content analysis for audio classification and segmentation. *IEEE Transactions on Speech and Audio Processing*, 10(7):504–516, Oct 2002.

[102] Hermansky H. Perceptual linear predictive (plp) analysis of speech. *Acoustic, Soc, Am.*, 87(4):1738–1752, 1990.

[103] K. Kumar, C. Kim, and R. M. Stern. Delta-spectral cepstral coefficients for robust speech recognition. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4784–4787, May 2011.

[104] J. Fu, S. W. Wang, X. L. Cao, M. L. Jiang, S. H. Zhang, and X. Zhao. The research on speech endpoint detection algorithm based on spectrogram row self-correlation. In *Proceedings of 2012 2nd International Conference on Computer Science and Network Technology*, pages 212–216, Dec 2012.

[105] C. Zhang and M. Dong. An improved speech endpoint detection based on adaptive sub-band selection spectral variance. In *2016 35th Chinese Control Conference (CCC)*, pages 5033–5037, July 2016.

[106] Wang K C Wu B F. Robust endpoint detection algorithm based on the adaptive band-partitioning spectral entropy in adverse environments. *IEEE Transactions on Speech & Audio Processing*, 13(5):762 – 775, 2005.

[107] H. Zhiyan and W. Jian. Research on speech endpoint detection under low signal-to-noise ratios. In *The 27th Chinese Control and Decision Conference (2015 CCDC)*, pages 3635–3639, May 2015.

[108] M. Sifuzzaman and M Rafiq Islam. Application of wavelet transform and its advantages compared to fourier transform. *Journal of Physical Sciences*, 13:121–134, 2009.

[109] C. Yali, L. Dongsheng, J. Shuo, and N. Xuefen. A speech endpoint detection algorithm based on wavelet transforms. In *The 26th Chinese Control and Decision Conference (2014 CCDC)*, pages 3010–3012, May 2014.

[110] M. R. K. Mollaei M. Eshaghi. Voice activity detection based on using wavelet packet. *Digital Signal Processing*, 20(4):1102–1115, 2010.

[111] Shaohua Feng Yanyan Jiang Jianwei Ma Bin Zhang Lu Liu, Siming Ma. Envelope extraction of anaesthesia breathing sound signal on hilbert huang transform. *Procedia Engineering*, 29:2693–2698, 2012.

[112] V. Rapcan, S. D'Arcy, and R. B. Reilly. Automatic breath sound detection and removal for cognitive studies of speech and language. In *IET Irish Signals and Systems Conference (ISSC 2009)*, pages 1–6, June 2009.

[113] Bassi A Vivaldi EA Díaz JA, Arancibia JM. Envelope analysis of the airflow signal to improve polysomnographic assessment of sleep disordered breathing. *Sleep*, 37(1):199–208, 2014.

[114] Richard Lyons. Dsp tricks: Approximate envelope detection. https://www.embedded.com/design/configurable-systems/4210780/DSP-Tricks--Approximate-envelope-detection, 2010. [Accessed: 7 June 2017].

[115] P. Delacourt and C.J. Wellekens. Distbic: A speaker-based segmentation for audio data indexing. *Speech Communication*, 32(1):111 – 126, 2000. Accessing Information in Spoken Audio.

[116] R. Bakis et. al. Transcription of broadcast news shows with the ibm large vocabulary speech recognition system. *in DARPA Speech Recognition Workshop*, 1997.

[117] Theodorou T. An overview of automatic audio segmentation. *International Journal of Information Technology & Computer Science*, 6(11), 2014.

[118] H. Beigi and S. Maes. Speaker, channel and environment change detection. *in World Congress on Automation. Citeseer*, 1998.

[119] B. H. Juang and L. R. Rabiner. Hidden markov models for speech recognition. *Technometrics*, 33(3):251–272, 1991.

[120] A. Dufaux, L. Besacier, M. Ansorge, and F. Pellandini. Automatic sound detection and recognition for noisy environment. In *2000 10th European Signal Processing Conference*, pages 1–4, Sept 2000.

[121] H. Lo, J. Wang, and H. Wang. Homogeneous segmentation and classifier ensemble for audio tag annotation and retrieval. In *2010 IEEE International Conference on Multimedia and Expo*, pages 304–309, July 2010.

[122] V. Vapnik. An overview of statistical learning theory. *IEEE Trans. on Neural Networks*, 10(5):988–999, 1999.

[123] T. Takagi and M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-15(1):116–132, Jan 1985.

[124] G. Wu, Z. Zhu, and A. Li. Fuzzy neural networks for speech endpoint detection. In *2012 International conference on Fuzzy Theory and Its Applications (iFUZZY2012)*, pages 354–356, Nov 2012.

[125] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119 – 139, 1997.

[126] J. Li, W. Dai, F. Metze, S. Qu, and S. Das. A comparison of deep learning methods for environmental sound detection. In *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 126–130, March 2017.

[127] Jun Huang, Y. Dong, Jiqing Liu, Chengyu Dong, and H. Wang. Sports audio segmentation and classification. In *2009 IEEE International Conference on Network Infrastructure and Digital Content*, pages 379–383, Nov 2009.

[128] Huawen Liu, Jigui Sun, Lei Liu, and Huijie Zhang. Feature selection with dynamic mutual information. *Pattern Recognition*, 42(7):1330 – 1339, 2009.

[129] Mathias Lohne. The computational complexity of the fast fourier transform. https://folk.uio.no/mathialo/texts/fftcomplexity.pdf, 2017.

[130] Hassan Ghasemzadeh, Navid Amini, and Majid Sarrafzadeh. Energy-efficient signal processing in wearable embedded systems: An optimal feature selection approach. *Proceedings of the International Symposium on Low Power Electronics and Design*, pages 357–362, 07 2012.

[131] Audacity. http://www.audacityteam.org/, 2018. [Accessed 25 June 2015].

[132] Inven Sense. Inmp441. https://www.invensense.com/products/digital/inmp441/, 2014. [Accessed: 7 July 2015].

[133] Isabelle Guyon and André Elisseeff. An introduction to variable and feature selection. *Machine Learning Research*, 3:1157–1182, March 2003.

[134] Sigma Plus Statistiek. Pearson correlations - quick introduction. https://www.spss-tutorials.com/pearson-correlation-coefficient/, 2018. [Accessed: 4 November 2018].

[135] J. L. Rodgers and W. A. Nicewander. Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42(1):59–66, 1988.

[136] X. Zhi, S. Yuexin, M. Jin, Z. Lujie, and D. Zijian. Research on the pearson correlation coefficient evaluation method of analog signal in the process of unit peak load regulation. In *2017 13th IEEE International Conference on Electronic Measurement Instruments (ICEMI)*, pages 522–527, Oct 2017.

[137] Patrick Mannion. Comparing low-power wireless technologies. https://www.digiauthor.co.uk/en/articles/techzone/2017/oct/comparing-low-power-wireless-technologies, 2017.

[138] Bluetooth SIG Inc. Radio versions. https://www.bluetooth.com/bluetooth-technology/radio-versions, 2018. [Accessed: 1 August 2018].

[139] 2018 Bluetooth SIG Inc. Topology options. https://www.bluetooth.com/bluetooth-technology/topology-options, 2018. [Accessed: 1 August 2018].

[140] 2018 Bluetooth SIG Inc. Sig introduces bluetooth low-energy wireless technology, the next generation of bluetooth wireless technology. https://www.bluetooth.com/news/pressreleases/2009/12/17/sig-introduces-bluetooth-low-energy-wireless-technologythe-next-generation-of-bluetooth-wireless-t note = [Accessed: 1 August 2018], 2018.

[141] Cameron Faulkner. What is nfc? everything you need to know. https://www.techradar.com/uk/news/what-is-nfc, May 2017. [Accessed: 23 Dec 2017].

[142] Blue Calypso Inc. Near field communication (nfc). http://bluecalypso.com/what-we-do/activation-methods/nfc, 2018. [Accessed: 20 April 2018].

[143] NFC Forum. Technical specifications. https://nfc-forum.org/our-work/specifications-and-application-documents/specifications/nfc-forum-technical-specifications/, 2018. [Accessed: 15 Sep 2018].

[144] Square Inc. Near field communication. http://nearfieldcommunication.org/, 2017. [Accessed: 2 Jun 2018].

[145] Premier Farnell Limited. Cc2541f128rhat. https://uk.farnell.com/texas-instruments/cc2541f128rhat/ic-soc-bluetooth-2-4ghz-40vqfn/dp/2281643, 2018. [Accessed: 1 Jun 2018].

[146] Premier Farnell Limited. M24sr02-ydw6t-2. https://uk.farnell.com/stmicroelectronics/m24sr02-ydw6t-2/nfc-rfid-tag-2kbit-13-56mhz-tssop/dp/2396467, 2015. [Accessed: 1 Jun 2018].

[147] Square Inc. Near field communication versus bluetooth. http://nearfieldcommunication.org/bluetooth.html, 2017. [Accessed: 15 Dec 2017].

[148] Henning Siitonen Kortvedt and Stig F. Mjølsnes. Eavesdropping near field communication. https://www.researchgate.net/publication/265976861_Eavesdropping_Near_Field_Communication, 02 2019.

[149] Panasonic Industrial Devices. https://industrial.panasonic.com/content/data/SC/ds/ds4/MN63Y1213_E.pdf, 2015. [Accessed: 7 July 2015].

[150] Identiv Inc. Scl010/scl011 contactless smart card reader. https://support.identiv.com/scl010-scl011/, 2018. [Accessed: 10 May 2018].

[151] P. Noll. Mpeg digital audio coding. *IEEE Signal Processing Magazine*, 14(5):59–81, Sept 1997.

[152] Y. Tarui, Y. Hayashi, and K. Nagai. Electrically reprogrammable nonvolatile semiconductor memory. *IEEE Journal of Solid-State Circuits*, 7(5):369–375, Oct 1972.

[153] Texas Instruments Incorporated. Msp430fr573x mixed-signal microcontrollers. http://www.ti.com/product/msp430fr5736, 2011. [Accessed: 12 May 2015].

[154] Fujitsu Semiconductor America Inc. Non-volatile ferroelectric random access memory (fram). https://www.fujitsu.com/us/Images/SPBG_FRAM_Overview_BR.pdf, 2015. [Accessed: 10 Dec 2017].

[155] Panasonic Semiconductor Solutions Co. Ltd. Mn63y1221: Nfc tag lsis. https://industrial.panasonic.com/tw/products/semiconductors/nfctags/lsis/nfc-tag-lsis/MN63Y1221, 2016.

[156] RFduino. Rfduino. http://www.rfduino.com/, 2013. [Accessed: 15 April 2015].

[157] STMicroelectronics. Mems digital output motion sensor: ultra low-power high performance 3-axis "femto" accelerometer. https://www.st.com/resource/en/datasheet/dm00042751.pdf, 2011. [Accessed: 20 August 2016].

[158] Texas Instruments Incorporated. Msp430fr57xx family user's guide. http://www.ti.com/general/docs/lit/getliterature.tsp?baseLiteratureNumber=slau272&fileType=pdf, 2018. [Accessed: 20 Sept 2018].

[159] West J.E. Sessler, G.M. Self-biased condenser microphone with high capacitance. *Journal of the Acoustical Society of America*, 34(11):1787–1788, 1962.

[160] Woon Seob Lee Seung S. Lee. Piezoelectric microphone built on circular diaphragm. *Sensors and Actuators A: Physical*, 144(2):367 – 373, 2008.

[161] Analog Devices Inc. Admp401: Omnidirectional microphone with bottom port and analog output obsolete data sheet. https://www.analog.com/en/products/admp401.html, 2012.

[162] Invensense. Inmp522: Ultra-low noise, low sensitivity tolerance, pdm digital microphone. https://www.invensense.com/products/digital/inmp522/, 2014.

[163] Jerad Lewis. Analog and digital mems microphone design considerations. *Analog Devices Technical Note MS2472*, 2013. [Accessed 18 March 2016].

[164] Allergan. Aerochamber plus flow-vu. http://www.aerochambervhc.com/, 2011. [Accessed: 10 August 2018].

[165] ATC Semitec. Thermistors. https://atcsemitec.co.uk/, 2018.

[166] Atmel. Atmel 8-bit avr microcontroller in-system programmable flash. https://www.microchip.com/wwwproducts/en/ATtiny85, 2013.

[167] Asthma UK. Reliever inhalers. https://www.asthma.org.uk/advice/inhalers-medicines-treatments/inhalers-and-spacers/reliever/, 2019.

[168] Asthma UK. Preventer inhalers. https://www.asthma.org.uk/advice/inhalers-medicines-treatments/inhalers-and-spacers/preventer/, 2019.

Appendix A

Participant Information Sheet

Title of research project: Dry powder inhaler adherence monitoring device

Name of lead researcher: Dr Mohammed Benaissa, University of Sheffield

Thank you for your interest in taking part in this study. Before you decide whether to take part, it is important that you understand why the research is being done and what it will involve for you. Please read the information provided carefully, and discuss it with others if you wish. Please ask us if there is anything that is not clear or if you would like more information.

What is the purpose of this research?

The purpose of this study is to gather feedback on a prototype technology used to assist with adherence in using dry powder inhalers. Adherence is the term for the success of a patient to follow the instructions regarding correct therapeutic usage or treatment. The specific application being focused on is an antibiotic inhaler used in the treatment of cystic fibrosis (CF). This is the first stage of user testing and the aim is to help us to make a system which is

practical and usable together with assessing its measurement performance. CF is a relatively rare condition so initial testing is being carried out with volunteers with any restriction on lung function.

Why have I been chosen?

You have been asked to take part because you have some restriction of lung function (eg asthma) and are a member of staff or student registered at the University of Sheffield.

Do I have to take part?

It is up to you to decide whether or not to take part. If you do not want to take part, it will not affect you in any way. If you agree to take part, we will discuss this information with you again, and ask you to sign a consent form to show you have agreed to take part. If you agree to take part but then decide that you do not want to, you can withdraw at any time. This would not affect any services you may receive.

What will happen to me if I take part?

If you decide to take part, you will be invited to use an inhaler which does not contain any drug, powder or therapeutic component. The 'inert' inhaler has been equipped with a small device which is being tested to record the presence and duration of the breath-in, hold breath for a short time and exhale cycles required to use this type of inhaler. There is a second sensor which is being used as a comparison with the one under development which is wired to a laptop. The enrolment and experiment can be completed in a single visit to the Electronic and Electrical Engineering Department (EEE) which should take less than an hour to complete. The steps involved are:

1. We will invite you to attend a session at the Electronic and Electrical Engineering Department (EEE) at the University of Sheffield in July 2018, when we will tell you more about the study.

2. If you agree to take part, you will be provided with brief instructions on how the inhaler is expected to be used.

3. We will ask you to inhale and exhale using an inert inhaler (there is no drug present) through a disposable mouthpiece a few times according to the instructions. The breathing pattern and period of holding breath will be recorded by the system.

4. You will also be asked to give your feedback as to the via a short questionnaire.

5. This completes your kind involvement with the research project.

What is the technology that you are using in this study?

There are two types of adherence reinforcement technology being considered. The first measures and records the timings relating to usage of the inhaler. The second version provides some immediate feedback via a coloured LED (small indicator light) as to the correct usage. The technology comprises a small package of electronics mounted to an inhaler which uses a microphone and motion detector to record usage information. The data can then be downloaded later to confirm adherence.

What will happen if I do not wish to take part?

Participation is completely voluntary and you do not have to take part. This will not affect the nature or standard of care you receive from any health or social care service in any way.

What are the possible benefits and risks of taking part?

While there are no immediate benefits in taking part, it is hoped that the information you will give us will help to improve the design of inhaler adherence monitoring technology for people with respiratory conditions which require the regular use of dry powder inhalers. We do not anticipate that there will be any particular risks or disadvantages to taking part in this study.

Are there any expenses or payments involved?

You will not receive any payment for taking part in this study.

What if I change my mind during the study?

You are free to withdraw from the study at any time. You will not have to give any reasons for your withdrawal.

Will my taking part in the project be kept confidential?

All information which is collected about you during the course of this study will be kept strictly confidential. The information you give will not be used in any way that could identify you. You will be identified by a code rather than by name and the information will be stored in password-protected encrypted computer files at the University of Sheffield which can only be accessed by the research team. You will not be able to be identified in any reports or publications.

Who will have access to the data and where will it be held?

All the data will be held in confidence at the University of Sheffield. It will be used for the purpose of this research. The research team will have access to the data. With your consent, the anonymised data may also be used by other researchers at the University of Sheffield.

What will happen to the results of the study?

We aim to publish results of the study in academic journals and present our findings at professional conferences. Nobody will be able to identify you in any reports or publications.

What if something goes wrong?

If by participating in this research you have any questions, please contact the project team (details given at the end). This type of research is not considered to be harmful. If you find using this technology upsetting, or if you wish to seek advice or reassurance about your own health, then either contact your GP or the university welfare services.

If you do wish to complain or have any concerns about how you have been approached or treated in the course of this study you can contact Mike Hounslow, Vice-President & Head of Faculty (Engineering).

Who is organising and paying for the research?

The research is self-funded at the University of Sheffield. The overall project area relies on close links with Respiratory Medicine at Sheffield Teaching Hospitals NHS Foundation Trust.

Who has reviewed this study?

This study has been reviewed and approved by the Research Ethics Committee at the University of Sheffield.

Appendix B

Ethics Approval Letter

Approved: 12/02/2018

Fei He

Registration number: 130264118

Electronic and Electrical Engineering

Programme: Dry powder inhaler adherence monitoring device

Dear Fei

PROJECT TITLE: Dry powder inhaler adherence monitoring device: initial user testing study

APPLICATION: Reference Number 016858

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 12/02/2018 the above-named project was approved on ethics grounds, on

the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 016858 (dated 06/11/2017).

- Participant information sheet 1036960 version 2 (05/03/2018).

- Participant consent form 1036961 version 1 (06/11/2017).

University research ethics application form 016858 (dated 06/11/2017).

The following optional amendments were suggested:

I still didn't see revision to clarify what inert powder will be used. If inert powder used, please specify what it is. Need to also assess any issue potentially caused by the powder. for example irritation?

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since written approval will be required.

Yours sincerely

Peter Rockett Ethics Administrator Electronic and Electrical Engineering

Appendix C

Initial Testing Protocol

Dry powder inhaler adherence monitoring device: initial user testing study

1. Purpose

This protocol describes the processes and procedures for a study to gain initial user feedback on the prototype inhaler adherence monitoring device. The participants will be involved in the initial stage of testing on the prototype technologies. The technology and will be developed using the feedback gathered during the initial testing and evaluated with the target patient population (cystic fibrosis) using an observational study.

2. Introduction

One of the key indicators of success in the treatment of cystic fibrosis (CF) using inhaled dry powder antibiotic delivery is the adherence to the treatment regime. The treatment is undertaken regularly by the patient in their home. Correct and regular usage of the inhaler is needed to maintain the required concentration of antibiotic. One solution to aid adherence may be the use of technology to provide monitoring and feedback.

From an engineering and economic perspective, the development of suitably practical, small and disposable monitoring system would be an appealing solution.

This protocol concerns the initial user testing of prototype systems with a view to obtaining feedback from people with respiratory conditions. It also provides an opportunity to assess the relative merits of different ways of presenting instructions and providing automatic feedback.

3. Technology

The systems being developed is a small package of electronics which can measure the usage of a dry power inhaler and store this information for later retrieval using a suitable reader. The device may have a multi-colour LED (indicator light) to provide real-time feedback to the use as to the correct usage. The technology has been developed by members of the research team in Electronic and Electrical Engineering (EEE) at the University of Sheffield.

4. Objectives

The objective of this sub-study is to gather usage data and user feedback on the prototype device to inform the development and refinement of the technology. This is the first stage of user testing, which will continue during later stages towards testing with CF patients in an NHS setting.

5. Study Design

It is anticipated to undertake a number of cycles to develop the prototype technology including its internal testing / validation (to ensure safety and functionality), followed by a period of volunteers to its usability, gather usage data and solicit their views which will then inform

the design process to refine the approach. In this sub-study, feedback from participants will be collated using a short questionnaire completion at the end of the study session.

6. Participants

We will undertake this study with approximately a dozen volunteers.

Subject inclusion criteria:

- A restricted lung function of some description (eg Asthma)

- Aged 18+ years

Subject exclusion criteria:

- Unable to provide informed consent

- Unable to communicate in written & verbal English

7. Recruitment

Participants will be identified through the established research 'volunteers' mailing list at the University of Sheffield. This list is co-ordinated by Corporate Information and Computing Services (CICS) at the University of Sheffield, and therefore all potential participants are either students or members of staff.

An invitation email will be distributed through this mailing list to advertise the study and invite eligible people to participate. Potential participants are required to opt-in to the study by contacting the research team. If they express an interest, they will be sent a participant information sheet by email to tell them more about the study. Potential participants are then invited to contact the research team again if they are interested in taking part at which point

they will be informed of the date, time and location of the sessions at the EEE department. These will be in July 2018. We anticipate that the whole recruitment, consent and testing process will take up to a maximum of one hour.

Written informed consent will be taken from all participants by one of the research staff. All participants will be free to withdraw at any time without providing any explanation if that is their wish.

8. Study procedures

Once they have given written informed consent, participants will be given instructions on how to use the technology and asked to use the inert inhaler a few times whilst usage data is acquired. This will be from the device under test and a baseline measurement system connected to a laptop.

The participants will be asked to trial the technology a few times during the session and then will be encouraged to provide some feedback via a short questionnaire.

8.1 Potential harm to participants

This is a low-risk project. The study requires people to use an inert dry powder inhaler (no drug, powder or propellant is present). A 3d printed mouthpiece is provided for each volunteer and discarded after their use. We are not asking them to take any drug or change their usual care but rather evaluate their experience of using the technology only and the ability of the technology to accurately measure their usage of the inert inhaler. The breathing pattern will be monitored without inhaling any substance. There is therefore no potential for physical harm to any participants.

9. Confidentiality and anonymity

Personal information (e.g. name, address, email and telephone number) will only be recorded for purposes of communication between participants and the research team, and will only be accessed by members of the research team at the University of Sheffield. Their participation will be confidential and handled anonymously. Participants will be given an explanation about how the data will be processed, and an understanding that the data gathered in the study will not be reported, discussed or made available in such a way that will enable them to be identified. The only intended use of personal identifiable information is as part of the recruitment and consent process. Consent forms will be achieved for 5 years then destroyed. The anonymised data will be retained and archived with the project and may be published as part of a thesis or academic paper.

10. Data handling

All research data will be stored electronically within protected storage on servers belonging to the University of Sheffield. This will only be accessible by members of the research team and all information will be kept confidential. Paper copies of questionnaires and consent forms will be stored in the study master file. All electronic data will be accessed on university computers that are password protected, and maintained within the university server in an encrypted database with encrypted off/cross-site backups.

11. Research team

Lead Researcher: Dr Mohammed Benaissa, Senior Lecturer, Department of Electronic and Electrical Engineering, University of Sheffield

Research Engineer: Dr Tim Good, Research Associate, Department of Electronic and Electrical Engineering, University of Sheffield

PhD Student:Miss Fei He, Department of Electronic and Electrical

Appendix D

Participant consent form

Title of research project: Dry powder inhaler adherence monitoring device

Name of lead researcher: Dr Mohammed Benaissa, University of Sheffield

1. I confirm that I have read and understand the Participant Information Sheet explaining the above research project and I have had the opportunity to ask questions about the project.

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without there being any negative consequences. In addition, should I not wish to answer any particular question or questions, I am free to decline.

3. I understand that my responses will be kept strictly confidential.

I give permission for members of the research team to have access to my anonymised responses. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in the report or reports that result from the research.

4. I agree for the data collected from me to be used in future research.

5. I agree to take part in the above research project.

6. I agree to be contacted about future research.

Name of participant Date Signature

Name of researcher taking consent Date Signature

Appendix E

Invitation email

Invitation email for Dry powder inhaler adherence monitoring device

Subject Header: Inhaler correct usage monitoring: Volunteers wanted to help developing monitoring technologies: single session of less than an hour

Dear volunteers,

Do you have any form of restricted lung function? For example, that required the use of an inhaler. Do you want to take part in research to help us develop better technologies for self-management?

We are looking for adults with any form of restricted lung function (eg Asthma) to help us to test our prototype technologies for measuring correct usage of a particular class of inhaler called a dry powder inhaler.

This study does not involve the delivery of any drug an inert inhaler is used which has been instrumented to allow its usage to be monitored. Such inhalers are typically used to deliver

antibiotic therapy for the treatment of cystic fibrosis. This is a relatively rare disorder so volunteers with any restricted respiratory function are requested.

The research involves a single session which will include the consent process and provide an opportunity for the volunteer to use the prototype instrumented inhaler a few times following a brief set of instructions. At the end of the session, the volunteer will be asked to provide any feedback they wish via a short questionnaire.

This is the initial user testing of the prototype so a second measurement technique is attached to the inhaler connected to a laptop.

Individual sessions will be organised in the Mappin Street area or other mutually convenient campus location. Time of day is flexible and some time before the end of June.

The technologies have been developed by researchers in the Electronic and Electrical Engineering Department. The research has been granted ethical approval by the University of Sheffield.

Please get in touch to express your interest in taking part: fhe3@sheffield.ac.uk

With best wishes,

Fei HE

Department of Electronic & Electrical Engineering, University of Sheffield

Appendix F

User Guide of Initial User Testing

Waking Up

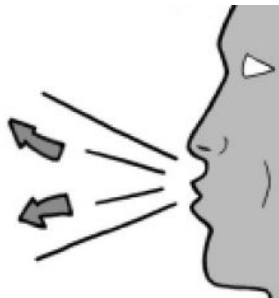1. Shake the inhaler twice, until the green light flashes for 1 second.

2. Hold the inhaler horizontally until the green light flashes again.

The Detection of Inhalation

1. To start **breathe out** (exhale) fully to empty your lungs of air.
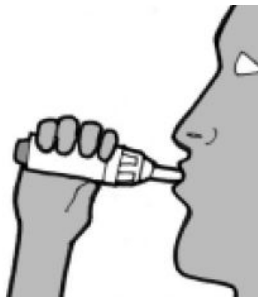

2. Place your mouth over the mouthpiece and close your lips tightly around it.

3. **Inhale** deeply with a single breath. Keep the inhalation at a **steady** pace. The green light will flash once during a good inhalation.

4. **Hold your breath** for 5 seconds after completing the inhalation. The green light will flash twice indicating the time has elapsed.

5. **Exhale** and take a few normal breaths away from the Podhaler device.
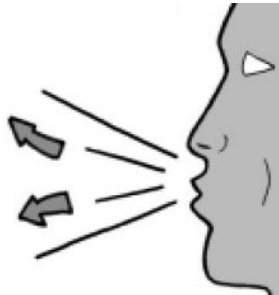
6. Place the inhaler back on the table and rest with normal breathing for around 1 - 2 minutes.

Repeat the set of steps for around 5 - 8 times.