Time-Frequency Analysis based Methods for Classification of Newborn Cry Signals

by

Saraswathy a/p Jeyaraman
(1141310639)

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

School of Mechatronic Engineering
UNIVERSITI MALAYSIA PERLIS

2016

UNIVERSITI MALAYSIA PERLIS

DECLARATION OF THESIS

Author's full name	:	Saraswathy a/p Jeyaraman
Date of birth	:	27 February 1987
Title	:	Time-Frequency Analysis Based Methods for Classification of Newborn Cry Signals
Academic Session	:	2011/2012

I hereby declare that the thesis becomes the property of Universiti Malaysia Perlis (UniMAP) and to be placed at the library of UniMAP. This thesis is classified as:

☐	**CONFIDENTIAL**	(Contains confidential information under the Official Secret Act 1972)*
☐	**RESTRICTED**	(Contains restricted information as specified by the organization where research was done)*
☐	**OPEN ACCESS**	I agree that my thesis is to be made immediately available as hard copy or on-line open access (full text)

I, the author, give permission to the UniMAP to reproduce this thesis in whole or in part for the purpose of research or academic exchange only (except during a period of _____ years, if so requested above).

Certified by:

SIGNATURE | **SIGNATURE OF SUPERVISOR**

(NEW IC NO. / PASSPORT NO.) | **NAME OF SUPERVISOR**

Date: _________________ | Date: _________________

NOTES: * If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentially or restriction.

ACKNOWLEDGEMENT

Praise to the Lord…

First and foremost, I would like to show my deepest gratitude to my parents, '*Amma* & *Appa*' for trying to provide me with all the best in life through it would mean a lot of sacrifice on their behalf. Thank you for your unconditional love and encouragement, support and understanding all this while. Thank you also for giving me the absolute freedom to choose the path of life to walk. This achievement would serve to be one of the first great accomplishments of many more to come.

I am grateful to my supervisors Dr. Hariharan and Prof. Dr. Sazali Yaacob for their never-ending support, patience, guidance, perspective advices and comment. Dr. Hariharan in particular has been an exceptional mentor and we often got engaged in stimulating discussions throughout the project. His help and assistance gave me the strength and driving force to reach for greater heights. The knowledge departed from him had also enlightens me in many ways.

I like to acknowledge Dr. Carlos Alberto Reyes-Garcia, Dr. Emilio Arch-Tirado and his INR-Mexico group, and Dr. Edgar M. Garcia-Tamayo from the Instituto Nacional de Astrofisica Optica y Electronica –CONACYT, Mexico for providing the 'Baby Chillanto Data Base'. I owe a debt of gratitude to Dr. Carlos Alberto Reyes-Garcia and Dr Pilar Gomez-Gil for providing the valuable information regarding the database when I requested.

Thanks are also extended to Dr. Gyorgy Varallyay Jr from the Department of Control Engineering and Information Technology, Budapest University of Technology and Economics, Budapest, Hungary and to her co-workers for sharing the deaf cry database with us.

I owe a debt of gratitude to Dr. Thiyaga and the rest of the dedicated nurses at the newborn ward of Hospital Sultanah Bahiyah for their help in collecting the clinical database. A special thanks also to Sarojini, Vinod and Ramesh for their kindness and assist in collecting the database.

Moreover, I would like to thank the University Malaysia Perlis (UniMAP) for the financial support (MyBrain15) during my postgraduate study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

t-f	Time-frequency
QTFDs	Quadratic time-frequency distributions
SPEC	Spectrogram
WVD	Wigner-Ville distribution
SWVD	Smoothed Wigner-Ville distribution
CWD	Choi-William distribution
MBD	Modified B-distribution
WPT	Wavelet packet transform
Wpspectrum	Wavelet packet spectrum
MFCCs	Mel-frequency cepstral coefficients
LPCs	Prediction coefficients
WPs	Wavelet packets
CWT	Continuous wavelet transform
DWT	Discrete wavelet transform
PNN	Probabilistic neural network
GRNN	General regression neural network
FS	Feature selection
LRS	Plus-1-minus-r feature selection
IGS	Information gain based feature selection
Exp 1	Experiment 1
Exp 2	Experiment 2
Exp 3	Experiment 3
Exp 4	Experiment 4
Exp 5	Experiment 5
Exp 6	Experiment 6
Exp 7	Experiment 7
Exp 8	Experiment 8
MHL	Moderate hearing loss
MSHL	Moderately severe hearing loss
SHL	Severe hearing loss

dmey	Finite impulse response based approximation of Meyer
F0	Fundamental frequency
RF1	First forming resonant frequency
RF2	Second forming resonant frequency
CNS	Central nervous system
F dom	Dominant frequency
DSP	Digital signal processing
STE	Short time energy
ZCR	Zero crossing rate
PCA	Principal component analysis
T-MFCC	Teager energy based MFCC
TEO	Teager energy operator
PMF	Probability mass function
CZT	Chirp Z-transform
ARX	Auto-regressive with exogenous
SSM	Smoothed spectrum method
SIFT	Simple inverse filtering tracking
FFT	Fast fourier transform
HPS	Harmonic product spectrum
EMD	Empirical mode decomposition
BPSO	Binary particle swarm optimization
PSO	Particle swarm optimization
OLS	Orthogonal least square
MSE	Mean squared error
DMPSO	Discrete mutative particle swarm optimization
AA_MFCCs	Arithmetic average of MFCCs
GFS	Genetic feature selections
DRSO	Data reduction through statistical operations
ES	Evolutionary strategies
FRRC	Fuzzy relational feature compression
RSD	Respiratory distress syndrome
ADEL	Ankyloglossia with deviation of the epiglottis and larynx
IUGR	Intrautererine growth restriction

RSD	Reflex sympathetic dystrophy syndrome
ASD	Autism spectrum disorder
Laryngomalacia	Larynx not developed
HIE	Hypoxic ischemic encephalopathy
URTI	Upper respiratory tract infection
SID	Sudden infant death syndrome
ANN	Artificial neural network
FFNNs	Feed-forward neural networks
GDBP	Gradient descent back-propagation
SCG	Scaled conjugate gradient
TDNN	Time delay neural network
MLP	Multilayer perceptron
PDF	Probability density function
SVM	Support vector machine
HMMs	Hidden markov models
PHMM	Parallel hidden markov model
CD-HMM	Continuous density-hidden markov model
GMM	Gaussian mixture model
BML	Adapted boosting mixture learning
LDA	Linear discriminant analysis
DCT	Discrete cosine transform
FT	Fourier transform
STFT	Short-time fourier transform
CWT	Continuous wavelet transform
DWT	Discrete wavelet transform
NNE	Nearest neighbour error
ANOVA	Analysis of variance
IG	Information gain

LIST OF SYMBOLS

%	Percent
ms	Milliseconds
s	Seconds
kHz	Kilohertz
cm	Centimetre

Penyiasatan Kaedah Analisis berasaskan Masa-Frekuensi dalam Pengelasan Isyarat Bayi Baru Lahir

ABSTRAK

Pengelasan tangisan bayi bayangkan kaedah objektif bukan-invasif, klasifikasi pola tangisan bayi yang berbeza dan penggunaan teknik pemprosesan isyarat tiruan dan digital. Ia telah dimulakan pada dekad yang lalu, untuk mengatasi kekurangan kaedah subjektif terutamanya persepsi auditori dan analisis spectrografik manusia yang bergantung kepada kadar pengalaman dan kepakaran pakar klinikal. Tesis ini membincangkan pembangunan kaedah objektif bagi pengkelasan tangisan bayi yang baru lahir terutamanya dengan menggunakan kaedah masa-frekuensi (m-f). Siasatan novel dengan menggunakan dua jenis m-f isyarat berasaskan pendekatan pemprosesan yang berbeza telah dijalankan: (a) 'Quadratic time-frequency distributions' (QTFDs): 'Spectrogram' (SPEC), 'Wigner-Ville distribution' (WVD), 'Smoothed-Wigner Ville distribution' (SWVD), 'Choi-William distribution' (CWD) dan 'Modified B-distribution' (MBD), and (b) 'Wavelet packet transform' (WPT) based method: 'wavelet packet spectrum' (Wpspectrum). Keberkesanan kaedah m-f yang disyorkan telah dianalisis dengan menggunakan isyarat tangisan bayi sihat dan patologi. Isyarat tangisan telah diakses daripada tiga pangkalan data yang berbeza asal (Mexico, Hungary dan Malaysia (pangkalan data yang dibangunkan sendiri). Dalam usaha untuk menyiasat keberkesanan kaedah m-f yang dicadangkan, lapan set data atau eksperimen yang berbeza telah dicadangkan, ia termasuk masalah 'binary' dan 'multiclass'. Pada 'binary' domain, analisis terhadap isyarat tangisan dari asal yang berbeza dan tahap keadahan patologi telah dijalankan. Rangka kerja ini telah direka kepada dua fasa untuk membandingkan penilaian prestasi kaedah t-f yang dicadangkan dengan keadaan sifat-sifat seni dalam bidang klasifikasi ('Mel frequency cepstral coefficients' (MFCCs) dan 'Linear prediction coefficients' (LPCs)). Pada mulanya penilaian prestasi kaedah m-f yang dicadangkan, MFCCs dan LPCs pada set data tangisan yang berbeza telah dijalankan. Dalam kes ini, satu kelompok m-f berasaskan ciri statistik dipetik daripada kaedah m-f yang disyorkan. Penilaian prestasi dari segi tugas pengelasan telah ditangani dengan menggunakan dua rangkaian berbeza diselia neural iaitu 'Probabilistic Neural Network' (PNN) dan 'General Regression Neural Network' (GRNN). Selepas itu, dengan mempertimbangkan prestasi klasifikasi, kaedah yang terbaik daripada QTFDs telah dipilih. Dalam fasa kedua, satu set ciri, gabungan MFCCs, LPCs dan ciri-ciri statistik diekstrak daripada QTFD terbaik dan kaedah Wpspectrum telah dicipta. Teknik pemilihan ciri yang berbeza seperti 'Plus-1-tolak-r' (LRS) dan 'Information Gain' (IGS) telah digunakan pada set ciri yang dicipta untuk mendapatkan subset yang berguna daripada ciri-ciri tersebut. Keupayaan diskriminasi vektor ciri yang dipilih dari segi ketepatan pengelasan telah dinilai dengan menggunakan PNN dan GRNN. Keputusan empirikal terbaik 100 %, dan 90 % ke atas telah diperolehi dalam kebanyakkan kes penyiasatan. Ia telah disimpulkan bahawa, sumbangan ciri statistik m-f berpangkalan dalam pembentukan keputusan klasifikasi yang baik adalah jauh lebih besar berbanding dengan ciri-ciri yang sedia ada (MFCCs dan LPCs) dalam kebanyakan kes eksperimen. Hasil kajian ini menyokong ketara penggunaan kaedah m-f dalam konteks klasifikasi tangisan bayi sebagai asas objektif untuk alat sokongan keputusan praktikal klinikal.

Investigation of Time-Frequency Analysis based Methods for Classification of Newborn Cry Signals

ABSTRACT

The infant cry classification implies non invasive objective methods, classification of different patterns of infant cry utterances and adoption of artificial and digital signal processing techniques. It has been commenced past decades ago to overcome the limitations of subjective methods in particularly auditory perception and human spectrographic analysis, which are relying on clinical rater‘s experience and expertise. This thesis addresses the development of an objective method for classification of newborn cries primarily using time-frequency (t-f) methods. Towards this aim, a novel investigation using two different t-f based signal processing approaches was performed: (a) Quadratic time-frequency distributions (QTFDs): Spectrogram (SPEC), Wigner-Ville distribution (WVD), Smoothed-Wigner Ville distribution (SWVD), Choi-William distribution (CWD) and Modified B-distribution (MBD), and (b) Wavelet packet transform (WPT) based method: wavelet packet spectrum (Wpspectrum). The effectiveness of the suggested t-f methods was analyzed using normal and different pathological cry signals. The investigational cry signals were accessed from three different origins of databases (Mexico, Hungary and Malaysia (self-developed database). In order to investigate the effectiveness of the suggested t-f methods, eight different cry experiments were suggested, including binary and multiclass problems. In the binary domain, analysis of cry signals from different origin and the severity level of pathological cry signals were considered for investigation. The framework of this work was designed in two phases in order to compare the performance evaluation of the suggested t-f methods with the state of the art attributes in the infant cry classification area (Mel frequency cepstral coefficients (MFCCs) and Linear prediction coefficients (LPCs)). Initially, the performance evaluation of the individual suggested t-f methods, MFCCs and LPCs on different proposed cry datasets were performed. In this case, a cluster of t-f based statistical features was extracted from the suggested t-f methods. The performance evaluation in term of classification task was tackled using two different supervised neural networks, namely Probabilistic Neural Network (PNN) and General Regression Neural Network (GRNN). Subsequently, by considering the classification performance, the best distribution from the QTFDs was selected. In the second phase, a feature set, combination of MFCCs, LPCs and the extracted statistical features from the best QTFDs and Wpspectrum was formed. Different feature selection techniques, such as Plus-1-minus-r (LRS) and Information Gain (IGS) were applied on the formed feature set to obtain a parsimonious subset of those features. The discrimination capability of the selected feature vector in terms of classification accuracy was evaluated using PNN and GRNN. The best empirical result of 100 % and, above 90 % in most of the cases was attained. It was inferred that, the contribution of the t-f based statistical features in the demonstration of the promising classification results was significantly greater compared to the conventional features in most of the experimental cases. The findings of this study significantly support the use of t-f methods in the context of infant cry classification as an objective basis for practical clinical decision support tools.

CHAPTER 1

INTRODUCTION

1.1 Exordium

Crying is a form of biological magnetic siren for an infant. It is their only means of communication and infants generally attract the attention of their external vicinity by crying. In naturally it is a chaotic, high pitched acoustic signal with greater fundamental frequency than adults and a multimodal behavior, which involves limb movement, facial expressions which inconstant over time (Varallyay, 2006). Cry is a valuable acoustic wave since it transmits meaningful information about a baby's physical and physiological status as shown in Fig. 1.1 (Patil, 2010). This fact led towards the emergence of infant cry analysis and classification.

In the early 1960s, the first seminal works with infant cry were initiated by a team of Scandinavian researchers lead by Wasz-Hokert. In 1964, the research group of Wasz-Hockert showed that the four basic types of cry (pain, hunger, pleasure and birth) can be identified by listening. Results of cry analysis reported by this team proved that the acoustic characteristics of an infant are different in each context; it depends on pathological conditions (Lederman, 2002; Wasz-Hockert, Partanen, Vuorenkoski, Michelsson & Valanne, 1964).

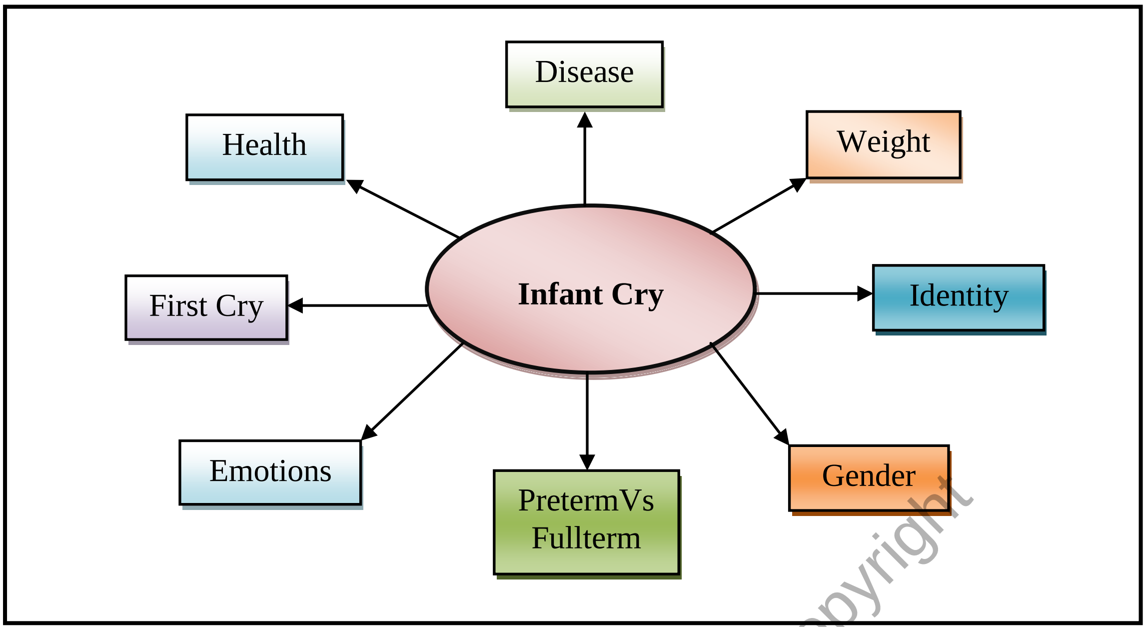

Figure 1.1: Different levels of information conveyed in infant cry (Patil, 2010).

Subsequently, numerous studies have investigated the relationship between various pathological conditions and acoustical features of the cry. Nevertheless, most of the analysis used time and/or frequency human analysis, especially by means of spectrograms which were strongly dependent on subjective evaluation and not suitable for clinical use. Therefore, by considering on these limitations, objective methods which allow automatic classification of the infant cry were desired. In 1984, a system of this type has been proposed and was shown to be successful in classification of hunger and pain cries of healthy full-term infants (Cohen & Zmora, 1984; Lederman, 2002).

These initiatives have encouraged existence of significant cry analyses in present which have successfully detected certain pathological conditions such as hearing impaired babies (Barajas & Reyes, 2005; Jam & Sadjedi, 2009a; Garcia & Reyes-Garcia, 2003a; Orozco-Garcia & Reyes-Garcia, 2003a; Reyes-Galaviz, Cano-Ortiz, & Reyes-Garcia, 2008a), sudden infant death syndrome (SIDS) (Robb, Crowell, & Dunn-Rankin, 2013), cleft palate (Lederman et al., 2002; Lederman, Zmora, Hauschildt, Stellzig-

Eisenhauer, & Wermke, 2008), brain damage (Lester & Boukydis, 1985), hydrocephalus (Michelsson, Kaskinen, Aulanko, & Rinne, 1984), asphyxia (Zabidi, Khuan, Mansor, Yassin, & Sahak, 2010a), hyperbilirubinemia (Kheddache & Tadj, 2013a; Santiago-Sanchez, Reyes-Garcia, & Gomez-Gil, 2009), autism (Orlandi, Manfredi, Bocchi, & Scattoni, 2012; Sheinkopf, Iverson, Rinaldi, & Lester, 2012), hypothyroidism (Zabidi, Khuan, Mansor, Yassin, & Sahak, 2010b; Zabidi, Mansor, Khuan, Yassin, & Sahak, 2010a), cri du chat (Garcia & Reyes García, 2002; Lederman, 2010), Respiratory Distress Syndrome (RSD) (Lederman et al., 2002; Lederman, Zmora, Hauschildt, Stellzig-Eisenhauer, & Wermke, 2008), gastroschisis (Kheddache & Tadj, 2013b), Ankyloglossia with deviation of the epiglottis and larynx (ADEL) (Okada, Fukuta, & Nagashima, 2011), X-Chromosomal abnormalities, Intra-uterine growth retardation, meningitis, lingual frenum, peritonitis, bovine protein allergy, heart defects (Cardio complex, Coarctation of aorta, Tetralogy of fallot, and Thrombosis in vena cava) (Alaie & Tadj, 2012) and neurological disorder (hypoxia-based Central Nervous System (CNS)) (Kheddache & Tadj, 2012). Considering its non-invasive aspect, the analysis of infant cry still has been used intensively in studies with newborn babies by implementing a plethora of signal processing approaches and pattern recognition models using different newborn cry signals to achieve reliable cry-based clinical routines for diagnosis.

1.2 Problem Statement

Human health and quality of life have indeed improved with the advance of medical science and health technology, which is used to evade illness onset, reduce the risk of occurrence and limit impact. However, the mortality and morbidity rates of infants are still at a greater level and declining at a slower pace. According to statistics, 2.9 million babies vanishing within the first month of life accounting for 44 % of under five years old infants. Fig. 1.2 illustrates the causes with the respective percentage which contributed for the fatality of infants under five years. Newborn (< 28 days of age) deaths account for 43 % of all deaths among children less than five years old and 75 % of neonatal deaths occurs during the first week of life, between 25 - 45 % occur within the first 24 hours. It can be deduced from the Fig. 1.2 that, the preponderance of infants is suffering and dying due to different pathological status (IGME, 2013; World Health Organization, 2014). Yet, World Health Organization (WHO) predicts that up to two thirds of newborn deaths could be prevented if led to timely diagnosis and treatments (IGME, 2013; World Health Organization, 2014). Diagnosis using conventional methods is subjective, strenuous and time consuming (Kheddache 2014; Lederman, 2002). Hence, to offset the associated morbidity and mortality rates, significant effort is being made to develop infant cry based non invasive, objective tools which may assist the medical experts in the diagnosis of neonatal pathological status. Moreover, the needs for the non invasive classification of infant cry signals in the infant cry analysis are emerging rapidly due to its significant benefits namely:

- objective and fully automatic system,
- not require the manual inspections of experts,
- the diagnosis judgment or result is accurate and fast,
- not limited to the quantity of infant cry signal which are under diagnosis,
- non invasive and harmless to infants

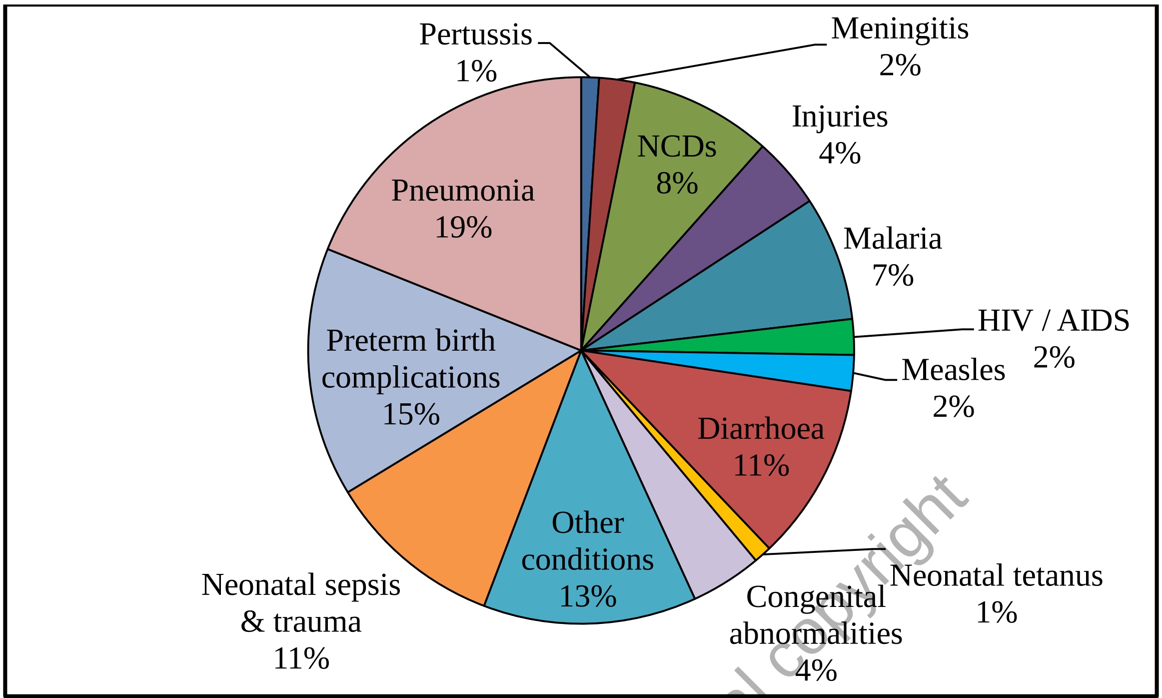

Figure 1.2: Causes of fatality in children under five years of age (World Health Organization, 2014)

In the development of infant cry classification systems, simple signal processing methods such as pitch information, harmonic analysis, noise analysis, Mel frequency cepstral coefficients and linear prediction coding (Orozco-Garcia & Reyes-Garcia, 2003a; Lester & Boukydis, 1985; Reyes-Galaviz, Cano-Ortiz, & Reyes-Garcia, 2008a; Santiago-Sanchez, Reyes-Garcia, & Gomez-Gil, 2009) have been proposed to characterize the multimodal cry patterns and it was observed that most of the proposed methods were from the time or frequency domains. However, generally in analyzing highly non-stationary signals the time-frequency (t-f) based techniques have been shown to outperform classical techniques based on either time or frequency domains in many applications (Avendano-Valencia, Gondino-Llorente, Blanco-Velasco, & Castellanos-Dominguez, 2010; Abdulla & Wong, 2011; Boashash & Boubchir, 2012; Boashasha, Khlif, Ben-Jabeur, East, & Colditz, 2014; Chen & Ling, 1999). Nevertheless the applications of t-f based approaches in the development of infant cry based diagnostic tools are not greatly highlighted and is being as a knowledge gap in this line of research area. Hence, the present study primarily highlights on investigation of t-f based methods for characterization of newborn cry signals. This current work addresses the development of a neonatal pathological status classification or detection tool using t-f based signal processing algorithms and supervised classification methods that can successfully classify different kinds of cries, with the objective of identifying different pathologies in newborns particularly asphyxia, deaf and jaundice (due to their availability) using their cry signals.

1.3 Research Objectives

This thesis embarks the following objectives:

i. To access and study different origins of newborn cry databases.

ii. To develop t-f based signal processing methods for characterizing different patterns of cry utterances.

a. To develop quadratic time-frequency distributions (QTFDs) based signal processing techniques for investigating the newborn cry signals.

b. To develop wavelet packet transform (WPT) based signal processing method for analyzing the newborn cry signals.

iii. To evaluate the effectiveness of the suggested t-f methods on discrimination of different cry datasets.

1.4 Scope of the Work

Since, the t-f analysis is a good approach for analyzing the highly non-stationary characteristic, in time and frequency scale simultaneously without eliminating any salient information, this research work address the development of an objective method for classifying different infant cry signals predominantly using two different t-f methods namely QTFDs (spectrogram (SPEC), Wigner-Ville distribution (WVD), Smoothed-Wigner Ville distribution (SWVD), Choi-William distribution (CWD) and Modified B-distribution (MBD)) and WPT based method (wavelet packet spectrum (Wpspectrum)). A cluster of t-f based features was extracted from the suggested t-f methods and their efficacy was examined using two supervised neural networks, namely probabilistic neural network (PNN) and general regression neural network (GRNN). Particularly, the suggested approaches were applied to three different databases from different origins namely Mexico (asphyxia, deaf, pain, hunger and normal cry signals), Hungary (different severity levels of deaf cry signals) and Malaysia (jaundice cry signals) in order to classify the following groups of datasets:

Binary Class Problem

- Experiment 1 (Exp 1): asphyxia Vs normal (Mexico-Mexico),
- Experiment 2 (Exp 2): deaf Vs normal (Mexico-Mexico),
- Experiment 3 (Exp 3): hunger Vs pain (Mexico-Mexico),
- Experiment 4 (Exp 4): deaf Vs normal (Hungary-Mexico),
- Experiment 5 (Exp 5): jaundice Vs normal (Malaysia-Mexico),
- Experiment 6 (Exp 6): moderate hearing loss (MHL) Vs moderately severe hearing loss (MSHL) (Hungary-Hungary),
- Experiment 7 (Exp 7): moderately severe hearing loss (MSHL) Vs severe hearing loss (SHL) (Hungary-Hungary),