IMPLEMENTATION OF FEATURE EXTRACTION AND CLASSIFICATION FOR SPEECH DYSFLUENCIES

LIM SIN CHEE

UNIVERSITI MALAYSIA PERLIS

2011

Implementation of Feature Extraction and Classification for Speech Dysfluencies

by

Lim Sin Chee
(0830610303)

A thesis submitted
In fulfilment of the requirements for the degree of
Master of Science (Mechatronic Engineering)

School of Mechatronic Engineering
UNIVERSITI MALAYSIA PERLIS

2011

UNIVERSITI MALAYSIA PERLIS

DECLARATION OF THESIS

Author's full name : Lim Sin Chee

Date of birth : 03 September 1984

Title : Implementation of Feature Extraction and Classification for Speech Dysfluencies

Academic Session : 2010 / 2011

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)
- [ ] **RESTICTED** (Contains restricted information as specified by the organization where research was done)
- [x] **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**
840903-10-5324	Prof. Dr. Sazali Yaacob
(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.

Dedicated to my dearest and supportive parents,

siblings and friends,

and my beloved husband Ong Wai Chong...

Thanks for being there and no words could

describe your supports and sacrifices...

I Love You...

ACKNOWLEDGEMENT

First and foremost, I would like to take this opportunity to express my deep appreciation and gratitude to my supervisor Prof. Dr. Sazali Yaacob for his kindness, guidance, constant motivation and encouragement to make this journey.

In addition, I would like to extend my heartfelt appreciation to my co-supervisor Dr. Hariharan Muthusamy for his helpful suggestions, ideas and assistances that have deeply helped me in completing this research.

Apart from that, I would like to thank my former supervisor Madam Ooi Chia Ai who offers me the opportunity to further my study at UniMAP by providing me a financial support from February 2009 to May 2010 through Research Assistantship (RA) supported by E-Science Fund from Ministry of Science, Technology and Innovation (MOSTI).

Furthermore, I would like to thank Dr. Mohd. Zambri Bin Ibrahim and Cik Fairus Binti Mukthar from Hospital Tuanku Fauziah, Kangar, Perlis for their opinions and experiences in speech dysfluencies.

Special thanks to all my friends and members from Intelligent Signal Processing Research Cluster (formerly known as Acoustic Application Cluster) who have directly and indirectly gave me a helping hand. Their kind-heartedness is much appreciated and welcomed. Indeed this is a wonderful research experience. I enjoyed every moment we have shared together and I appreciated everything that I have gained from it.

Lastly, I would like to dedicate my truthful thankfulness to my dearest parents and siblings for their financial and moral support as well as their continuous motivation. Last but not least, special thanks to my husband Ong Wai Chong who acted as my main pillar and supports me all the way in my study, without him this research would never reached to this stage.

TABLE OF CONTENTS

TOPICS		PAGE
TITLE		i
THESIS DECLARATION		ii
ACKNOWLEDGEMENT		iii
TABLE OF CONTENTS		vi
LIST OF TABLES		ix
LIST OF FIGURES		x
LIST OF ABBREVIATIONS		xiii
LIST OF SYMBOLS		xv
ABSTRAK		xvii
ABSTRACT		xviii
CHAPTER 1 INTRODUCTION		
1.1	Importance of the work	4
1.2	Problem Statements	4
1.3	Objectives	5
1.4	Scope of work	6
1.5	Thesis Outline	7
CHAPTER 2 AN OVERVIEW OF DYSFLUENCIES AND RELATED RESEARCH WORKS		
2.1	Introduction	9
2.2	Speech disorder, stuttering and dysfluencies	10
2.3	Types of dysfluencies	11
2.4	Basic features of repetition and prolongation classification	14

2.5 Related research works 16

2.6 Methodology of this work 22

2.7 Conclusion 24

CHAPTER 3 SIGNAL PROCESSING AND FEATURE EXTRACTION METHODS

3.1 Introduction 26

3.2 Database preparation 26

3.3 Signal pre-processing 28

3.3.1 Signal normalization 29

3.3.2 Pre-emphasis 31

3.3.3 Frame-blocking 34

3.3.4 Windowing 34

3.4 Feature Extraction 36

3.4.1 Short time Fourier transform (STFT) 37

3.4.2 Mel-frequency Cepstral Coefficient (MFCC) 43

3.4.3 Linear Predictive Coding based parameterization 44

3.5 Conclusion 49

CHAPTER 4 CLASSIFICATION TECHNIQUES

4.1 Introduction 50

4.2 Linear Discriminant analysis (LDA) 51

4.3 Least Squares Support Vector Machines (LSSVM) 52

4.3.1 Parameter selection: regularization parameter and standardization parameter 55

4.4 *k*-nearest Neighbor (*k*NN) 60

4.4.1 Parameter selection: *k*-values 62

4.5	Techniques for classification performance estimation	65
	4.5.1 Conventional Validation	65
	4.5.2 *Ten*-fold cross validation	66
4.6	Conclusion	66

CHAPTER 5 RESULTS AND DISCUSSION

5.1	Introduction	68
5.2	Experimental Setup	68
5.3	Selection of order *p* for LPC based parameterization features	71
5.4	Comparison between un-normalized data and normalized data	75
5.5	Parameter variation for acoustical features	77
	5.5.1 $\tilde{a}$-values	78
	5.5.2 Frame length	81
	5.5.3 Overlap percentages	84
5.6	Development of graphical user interface for speech dysfluencies classification tool using MATLAB®	88
5.7	Conclusion	95

CHAPTER 6 CONCLUSION AND RECOMMENDATIONS FOR FUTURE WORKS

6.1	Summary	97
6.2	Recommendations for future works	100

REFERENCES	103
APPENDIX A (The Mel filter bank)	108
APPENDIX B (Surface plot of parameter variation for each feature extraction)	109
LIST OF PUBLICATIONS	120
LIST OF AWARDS	121

LIST OF TABLES

NO.		PAGE
2.1	Summary of several research works on stuttering or speech dysfluencies classification, detailing the number of subjects, the features used and the classifiers employed.	16
3.1	Age and gender distribution of the reading recordings in the chosen subset of UCLASS databases.	27
5.1	Parameters setting of experiments.	70
5.2	Comparison of classification accuracy (%) for different feature extraction algorithms and classification techniques by varying the order *p* between 8 and 20.	72
5.3	Comparison of classification performance (%) for different feature extraction algorithms and classification techniques before and after normalization.	76
5.4	Comparison of classification accuracy (%) for different feature extraction algorithms and classification techniques by varying the $\tilde{a}$ values from 0.91 – 0.99.	79
5.5	Comparison of classification performance (%) for different feature extraction algorithms and classification techniques by varying the frame length from 10 – 50 ms.	82
5.6	Comparison of classification performance (%) for different feature extraction algorithms and classification techniques by varying the overlap percentages between 0% and 75%.	85
5.7	Correlation between validation schemas, classifiers, features and parameters that obtained the highest accuracy among others.	87

LIST OF FIGURES

NO.		PAGE
2.1	Graphical representation of the utterance “may”: speech signal (top), spectrogram (bottom). (a) Repetition (b) Prolongation.	15
2.2	Methodology of speech dysfluencies classification.	24
3.1	Block diagram of speech signal pre-processing.	29
3.2	Comparison between original speech waveform and normalized signal for dysfluent events namely (a) repetition; (b) prolongation.	30
3.3	Pre-emphasized signal for dysfluent events namely (a) repetition; (b) prolongation.	32
3.4	Amplitude spectral plot of original speech and pre-emphasis speech for stuttered events namely (a) repetition; (b) prolongation.	32
3.5	Figure 3.5: Magnitude responses of the pre-emphasis filter (a) Main Shot (b) Zoom in version with range of 0-800 Hz.	33
3.6	Common window functions namely Kaiser, Hamming, Bartlett, Blackman and Hanning in time domain.	35
3.7	Several feature extraction methods.	36
3.8	Block diagram of the STFT features	38
3.9	Spectrogram of dysfluent events (a) Repetition (b) Prolongation.	39
3.10	Repetition signal: (a) Time-maximum amplitude plot. (b) Frequency-maximum amplitude plot. (c) Frequency-standard deviation plot.	40
3.11	Prolongation signal: (a) Time-maximum amplitude plot. (b) Frequency-maximum amplitude plot. (c) Frequency-standard deviation plot.	40
3.12	Block diagram of MFCCs.	44
3.13	Block diagram of Linear Predictive Coding based parameterization.	45
4.1	Classification accuracy achieved by correlation between γ and σ^2 for MFCC features.	57
4.2	Classification accuracy achieved by correlation between γ and σ^2 for STFT features.	57

4.3	Classification accuracy achieved by correlation between γ and σ^2 for LPC features.	58
4.4	Classification accuracy achieved by correlation between γ and σ^2 for LPCC features.	58
4.5	Classification accuracy achieved by correlation between γ and σ^2 for WLPCC features.	59
4.6	Classification accuracy achieved by correlation between γ and σ^2 for FOTD features.	59
4.7	Classification accuracy achieved by correlation between γ and σ^2 for SOTD features.	60
4.8	Scatter plot of the acoustical feature classify by *k*NN when k-value equal to one.	62
4.9	Scatter plot of the acoustical feature classify by *k*NN when k-value equal to nine.	63
4.10	Classification accuracy of the two types dysfluencies achieved by the seven acoustical features by varying the *k*-values from one to ten.	64
5.1	Experimental setup of speech dysfluencies classification.	69
5.2	Classification accuracy achieved by correlation between order *p* and classifiers with respective validation schema for WLPCC features.	74
5.3	Classification accuracy achieved by correlation between $\tilde{a}$ values and classifiers with respective validation schema for STFT features.	81
5.4	Classification accuracy achieved by correlation between frame length and classifiers with respective validation schema for STFT features.	84
5.5	Classification accuracy achieved by correlation between overlap percentages and classifiers with respective validation schema for STFT features.	87
5.6	The front page of the system.	89
5.7	The main page of the system.	90
5.8	List of features is displayed in the main menu page of the system.	91
5.9	List of classifiers is illustrated in the main menu page of the system.	91
5.10	Buttons in Input Signal panel are enabled after both drop-down lists are selected.	92

5.11	The main menu page of the system with a dialog box.	93
5.12	The classified type of dysfluency (prolongation) is shown.	94
5.13	The classified type of dysfluency (repetition) is shown.	94

LIST OF ABBREVIATIONS

Symbol		Definition
ANN	-	Artificial Neural Network
CRV	-	Cross Validation
CV	-	Conventional Validation
DC	-	Direct-Current
DCT	-	Discrete Cosine Transform
FIR	-	Finite Impules Response
FFT	-	Fast Fourier Transform
FmA	-	Frequency-Maximum Amplitude
FOTD	-	First Order Temporal Derivative
Fstd	-	Frequency-Standard Deviation
GUI	-	Graphical User Interface
HMM	-	Hidden Markov Model
kHz	-	Kilo Hertz
*k*NN	-	*k*-Nearest Neighbor
LD	-	Lexical Dysfluencies
LDA	-	Linear Discriminant Analysis
LPC	-	Linear Predictive Coefficient
LPCC	-	Linear Predictive Cepstral Coefficient
LSSVM	-	Least Squares Support Vector Machines
MFCC	-	Mel-Frequency Cepstral Coefficient
MLP	-	Multilayer Perceptron
ms	-	milliseconds
PSD	-	Power Spectral Density

RBF	-	Radial Basis Function
SAM	-	STFT- amplitude matrix
SD	-	Supra-lexical Dysfluencies
SLIN	-	Least Squares Support Vector Machines with Linear Kernel
SLP	-	Speech Language Pathologist
SN	-	Signal Normalization
SOTD	-	Second Order Temporal Derivatives
SRBF	-	Least Squares Support Vector Machines with Radial Basis Function Kernel
STFT	-	Short Time Fourier Transform
SVM	-	Support Vector Machines
TmA	-	Time-Maximum Amplitude
TF	-	Time Frequency
UCLASS	-	University College London’s Acrhive of Stuttered Speech
WLPCC	-	Weighted Linear Predictive Cepstral Coefficient

LIST OF SYMBOLS

Symbol		Definition
$\tilde{a}$	-	Positive parameter used to control the degree of pre-emphasis filtering
a_m	-	Linear Predictive Coefficients
α	-	Vector of support values
b	-	Bias term
cov	-	Covariance
c_m	-	Linear Predictive Cepstral Coefficients
$\hat{c}_m$	-	Weighted Linear Predictive Cepstral Coefficients
D	-	Number of data values
f	-	Frequency
F	-	FFT length
G^2	-	Gain term in the LPC model
$K(.,.)$	-	Kernel function
L	-	Lagrangian
M	-	The number of sample overlap
N	-	The number of sample in each frame
P	-	Probability
p	-	Order of LPC
s_i	-	Signal
s_n	-	Normalized signal
s_p	-	Pre-emphasized signal
S_b	-	Scatter between classes
S_w	-	Scatter within classes
t	-	Time location

$w(n)$	-	Window function
w_m	-	Weighting function
$\bar{x}$	-	Mean /Average
$\Delta c_m(t)$	-	First Order Temporal Derivatives
$\Delta^2 c_m(t)$	-	Second Order Temporal Derivatives
%overlap	-	Overlap percentage
γ	-	Regularization Parameter (gamma)
ξ	-	Slack variable of support vector machine
σ	-	Width of RBF kernel
σ^2	-	Standardization Parameter (sigma2)
∂	-	Standard deviation
∂^2	-	Variances
∂_{skew}	-	Skewness
∂_{kurt}	-	Kurtosis

PERLAKSANAAN PENGEKSTRAKAN CIRI-CIRI DAN KLASIFIKASI UNTUK KETIDAKFASIHAN PERTUTURAN

ABSTRAK

Pertuturan sentiasa akan terganggu oleh elemen ketidakfasihan yang tidak diingini terutamanya pengulangan dan pemanjangan suara, suku kata dan perkataan yang mengakibatkan ketidakfasihan semasa berkomunikasi. Secara tradisi, ahli patologi pertuturan menghitung dan mengklasifikasikan kejadian ketidakfasihan dalam pertuturan secara manual. Namun demikian, jenis penilaian ini adalah subjektif, tidak konsisten, memakan masa dan cenderung kepada kesilapan. Pada tiga dekad yang lalu, banyak kerja penyelidikan dilaksanakan untuk automasi penilaian konvensional dengan pelbagai pendekatan seperti analisis isyarat pertuturan, pembolehubah peribadi, analisis akustik dari isyarat pertuturan dan teknik kecerdasan buatan. Berdasarkan kerja penyelidikan yang sebelum ini, ia dapat disimpulkan bahawa kaedah pengekstrakan ciri-ciri dan teknik klasifkasi memainkan peranan yang amat penting dalam bidang ini. Oleh kerana itu, dalam penyelidikan ini, beberapa kaedah pengekstrakan ciri-ciri, iaitu, Jemahan Fourier Masa Pendek (STFT), Pekali Cepstral Frequensi Mel (MFCC) dan Pengekodan Jakaan Lelurus Berdasarkan Pemparameteran telah dicadangkan untuk mengekstrakan ciri-ciri penting dari dua jenis ketidakfasihan pertuturan. Dengan mengaplikasikan kaedah pengekstrakan ciri-ciri pada setiap isyarat, sebanyak tujuh ciri-ciri akustik akan diekstrakan iaitu, STFT, MFCC dan lima ciri-ciri akustik diekstrakan dengan menggunakan Pengekodan Jakaan Lelurus Berdasarkan Pemparameteran, iaitu, Pekali Jakaan Lelurus (LPC), Pekali Cepstral Jakaan Lelurus (LPCC), Pekali Pemberat Cepstral Jakaan Lelurus (WLPCC), Terbitan Masa Tertib Pertama (FOTD) dan Terbitan Masa Tertib Kedua (SOTD). Ciri-ciri akustik yang diekstrakan itu akan digunakan sebagai masukan parameter bagi klasifikasi process seterusnya. Kedua-dua pengklasifikasi linear dan tidak linear iaitu Analisis Pembeza Layan Lelurus, Tetangga Terdekat-*k* dan Mesin Vektor Sokongan Kuadrat Terkecil dengan Kernel Lelurus (SLIN) dan Kernel Fungsi Asas Jejari (SRBF) telah disarankan untuk mengklasifikasikan dua jenis ketidakfasihan pertuturan iaitu pemanjangan dan pengulangan. Bagi menilai keberkesanan kaedah pengekstrakan ciri-ciri dan teknik klasifikasi yang berlainan, sebuah pengkalan data yang standard bernama *University College London Archive Stuttered Speech* (UCLASS) digunakan. Kebolehpercayaan terhadap ketepatan klassifikasi dapat dicapai dengan menerapkan dua skim pengesahsahihan, iaitu, konvensional pengesahsahihan dan sepuluh kali ganda silang pengesahsahihan. Bagi analisis selanjutnya, parameter pilihan dari setiap teknik klasifikasi dan variasi parameter, iaitu, tertib bagi Pengekodan Jakaan Lelurus Berdasarkan Pemparameteran, parameter yang digunakan untuk mengawal darjah penapisan pra-penekanan, kepanjangan rangka and peratusan bertindih dalam teknik pra-pemprosesan telah diselidik. Keputusan analisis melaporkan bahawa ketepatan klasifikasi tertinggi dicapai oleh ciri-ciri STFT dan pengklasifikasi SLIN. Dengan memerhati ketepatan klasifikasi yang diperolehi daripada ciri-ciri akustik dan pengklasifikasi yang berlainan, ia dapat disimpulkan bahawa hubungan antara ciri-ciri akustik dan pengklasifikasi perlu dinilai untuk mencapai ketepatan klasifikasi yang terbaik. Kesimpulannya, kaedah pengektrakan ciri-ciri dan pengklasifikasi yang dicadangkan boleh digunakan dalam klasifikasi ketidakfasihan pertuturan. Akhirnya, sebuah Antara Muka Pengguna Grafik bagi penyelidikan ini telah dibangunkan dengan menggunakan *MATLAB*® berdasarkan keputusan yang dicapai dalam eksperimen.

IMPLEMENTATION OF FEATURE EXTRACTION AND CLASSIFICATION FOR SPEECH DYSFLUENCIES

ABSTRACT

Speech is prone to disruption of involuntary dysfluent events especially repetitions and prolongations of sounds, syllables and words which lead to dysfluency in communication. Traditionally, speech language pathologists count and classify occurrence of dysfluencies in flow of speech manually. However, these types of assessment are subjective, inconsistent, time-consuming and prone to error. In the last three decades, many research works have been developed to automate the conventional assessments with various approaches such as speech signal analysis, personal variables, acoustic analysis of speech signal and artificial intelligence techniques. From the previous works, it can be concluded that feature extraction methods and classification techniques play important roles in this research field. Therefore, in this work, there are few feature extraction methods, namely, Short Time Fourier Transform (STFT), Mel-frequency Cepstral Coefficient (MFCC) and Linear Predictive Coding (LPC) based parameterization were proposed to extract the salient feature of the two types of dysfluencies. By applying the feature extraction methods on each signal, there are total of seven acoustical features extracted namely STFT, MFCC and five acoustical features from Linear Predictive Coding based parameterization, that is, Linear Predictive Coefficient (LPC), Linear Predictive Cepstral Coefficient (LPCC), Weighted Linear Predictive Cepstral Coefficient(WLPCC), First Order Temporal Derivatives (FOTD) and Second Order Temporal Derivatives (SOTD). Acoustical features are extracted from the signal are use as input parameters for classifiers. Both linear and nonlinear classifiers namely Linear Discriminant Analysis (LDA), k-Nearest Neighbor (kNN) and Least-Squares Support Vector Machines (LSSVM) with linear kernel (SLIN) and Radial Basis Function kernel (SRBF) were suggested to classify the two types of dysfluencies. In order to evaluate the effectiveness of the different feature extraction methods and classification techniques, a standard database named as University College London's Archive of Stuttered Speech (UCLASS) is used. The reliability of the classification accuracy is achieved by adopting the two validation schemas, namely, conventional validation and ten-fold cross-validation. For further analysis, parameters selections of the respective classifiers and parameter variation namely order of Linear Predictive Coding based parameterization, parameter used to control the degree of pre-emphasis filtering, frame length and overlap percentages on the signal pre-processing techniques are investigated. Analysis results reported that the highest classification accuracy is achieved by STFT features and SLIN classifier. By observing the classification accuracy obtained from different acoustical features and classifiers, it can be concluded that it is necessary to evaluate correlation between acoustical features and different classifiers in order to achieve the best classification accuracy. As a conclusion, the proposed feature extraction methods and classifiers can be used in speech dysfluencies classification. Finally, a Graphical User Interface of this work is developed by using MATLAB® based on the results achieved in the experiments.

CHAPTER 1

INTRODUCTION

Speech is a verbal means used by humans to express their feelings, ideas and thoughts in their communication. Speech can be used to transfer information, not only the content of statement but also speaker's emotional state, gender, age, intentions and others factors. In addition, speech consists of three components; articulation, voice and fluency skills to produce a fluent speech. Articulation is the ability to produce sound correctly such as how the speech organs are involved in producing a sound; the vocal cords or vocal fold and breathing are used to produce sound which also known as voice; fluency is the rhythm of speech. In order to produce fluent speech, the three components are closely related. For instance, a speech is produced by the movement of human's larynx muscles that controlling activities of the vocal cords.

However, speech will not occur without disruptions, which lead to dysfluency while communicate. Dysfluency in speech can be normal or pathological. Pathological speech may involve particular defections of the speech motor control in the brain such as the malfunction of larynx's muscle thus resulting in making normal speech impossible. According to Awad (1997), the muscle failure is linked to the brain and it is also noted that there is small amount of blood flow and either increased or decreased electrical activity in the regions of the brain involved in speech production.

Stuttering is one of the serious problems found in speech pathology (Awad, 1997; Czyzewski, Kaczmarek, & Kostek, 2003). It is a type of speech disorder and it can be divided into two groups of dysfluencies, such as supralexical dysfluencies (SD) and lexical dysfluencies (LD) (Howell, Au-Yeung, Sackin, Glenn, & Rustin, 1997; Howell, Sackin, & Au-Yeung, 1998). There is 1% of the population have noticeable speech stuttering problem and it is found to affect female to male with ratio 1: 3 or 4 times (Awad, 1997; Chia Ai & Yunus, 2006; Van Borsel, Achten, Santens, Lahorte, & Voet, 2003). Stuttering is defined as a flow of speech which is disrupted by unintentionally of dysfluencies such as repetition, prolongation, interjection of syllables, sounds, words or phrases and involuntary silent pauses or blocks in communication (Awad, 1997; Chia Ai & Yunus, 2006; Tian-Swee, Helbin, Ariff, Chee-Ming, & Salleh, 2007). The taxonomy of stuttered dysfluencies will be further discussed in Chapter 2.

Stuttering cannot be completely cured; it may go into remission for sometime (Awad, 1997). Stutterers can learn to shape their speech into fluent speech with appropriate speech pathology treatments. Therefore, stuttering assessments are needed to evaluate performance of stutterers before and after therapy. Traditionally, speech language pathologist (SLP) counts and classifies incidence of dysfluencies such as repetition and prolongation in speech manually. However, these types of stuttering assessment are subjective, inconsistent, time consuming and prone to error (Awad, 1997; Howell, Sackin, & Glenn, 1997a, 1997b; Nöth, et al., 2000; Ravikumar, Reddy, Rajagopal, & Nagaraj, 2008; Ravikumar, Rajagopal, & Nagaraj, 2009).

In the past three decades, there has been an increasing interest in stuttering. Researchers have focused on developing objective methods to facilitate the SLP during stuttering assessment. Although, some researchers have several attempts to use objective techniques to assess patients' performances before and after treatment, however those techniques do not easily allow for automatic assessment of stuttering severity (Czyzewski, et al., 2003; Howell, Au-Yeung, & Pilgrim, 1999; Howell et al., 1997a).

Consequently, employing SLP as an expert to count the incidence of dysfluencies in speech samples is the most ordinary way (Howell, et al., 1999; Howell, et al., 1998). In contrast, based on an acoustic point of view, it is possible to analyze electrical signal that representing the dysfluent speech signal. For some stutterers, the algorithms used in Digital Signal Aid (DSA) is signal transposition in the spectrum domain, has shown a high reduction of stuttering, especially in case of text reading (Czyzewski, et al., 2003). According to Czyzewski et al. (2003), in stuttered speech, articulation muscles contractions caused the alteration in the speech articulation system that are visible in the results of cepstral and spectral analysis. A number of researchers have used the concept to analyze stuttered speech. Some of the significant research works that related to this work and conducted in the past decades will be depicted in Chapter 2.

From the previous works, it can be concluded that feature extraction algorithms and classification techniques performs an important role in the field of dysfluent events classification. However, their achievements are not easily comparable since they used different database and lack of uniformity in computing the results. In this work, there are seven feature extraction algorithms and both linear and nonlinear classifiers are applied on the dysfluent sample speech signals. This work is designed to evaluate the effectiveness of various feature extraction methods in classification of two types of speech dysfluencies, namely, prolongation and repetition.

1.1 Importance of the work

Normally, children are exhibited with dysfluent speech, but they can be easily recovered during young ages with the appropriate therapy techniques. Therefore, objective method of speech dysfluencies classification system is necessary to measure the incidence of dysfluencies in speech and to help the SLP in diagnosing the clients. It is also used to assess the client performances before and after speech therapy and thus, to evaluate the efficacy of a therapy techniques for each stutterer. On the other hand, based on the incidence of dysfluencies, it helps SLP to decide appropriate therapy techniques for each client.

Moreover, Hancock et al. (1998) showed that the computer-based system were the most effective tool in treating stuttering followed by the parents' guidance in improving the stuttering conditions and the least is by the effort of the SLP. Therefore, this encourages the implementation of computer-based assessment system.

In this work, the design of signal pre-processing techniques, feature extraction methods and classification techniques those play the vital role in speech dysfluencies classification. This work is designed to evaluate the effectiveness of the techniques based on a standard database so that the techniques can be compared easily.

1.2 Problem Statements

Clearly, the incidence of dysfluencies in speech is used to differentiate between stutterer and non-stutterers. Normally, SLP diagnose their client through manual techniques which is based on client's stuttering severity such as measuring frequency of the dysfluent portion in the speech. However, these types of diagnosis method are time-consuming and prone to error. Recently, several research works have been done in speech stuttering field. Most of the research works are implemented to distinguish stutterers from non-stutterers and only few research works have been implemented and concentrated on dysfluencies classification. Thus, in this work, dysfluencies classification of prolongation and repetition are concentrated. A recent research by Howell et al. (1997b) illustrated that the dysfluency classification, namely, prolongation and repetition are relatively crude due to the input parameters such as spectral similarity to the classifiers may need to be improved. Thus, improvement on the previous works is done by implementing various feature extraction methods and artificial intelligence techniques. Furthermore, the reliability of the results which seldom been addressed can be achieved by applying validation schemas.

1.3 Objectives

The objectives of this research are described below:

- To obtain stuttered speech signals from the database and data processing.

- To extract the salient features from the stuttered speech by implementing feature extraction algorithms.