UNIVERSITI PUTRA MALAYSIA

A SPEECH ENHANCEMENT FRAMEWORK USING DISCRETE KRAWTCHOUK-TCHEBICHEF TRANSFORM

BASHEERA M. MAHMMOD

FK 2018 137

A SPEECH ENHANCEMENT FRAMEWORK USING DISCRETE KRAWTCHOUK-TCHEBICHEF TRANSFORM

By

BASHEERA M. MAHMMOD

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

May 2018

DEDICATION

To my beloved son Mohammed

The bird in Paradise

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

A SPEECH ENHANCEMENT FRAMEWORK USING DISCRETE KRAWTCHOUK-TCHEBICHEF TRANSFORM

By

BASHEERA M. MAHMMOD

May 2018

Chairman: Associate Professor Abd Rahman Ramli, PhD
Faculty: Engineering

Speech is considered the key mode of interaction amongst humans. Speech signals encounter different scenarios during transmission, such as interference and additive noise, which lead to generate noisy signals. Therefore, robust Speech Enhancement Algorithms (SEA) that suppress noise without distorting the original signals are necessary. The removing of noise without causing speech distortion is a challenging task. Moreover, an annoying noise that appears after the enhancement process, called Musical Noise (MN), should be eliminated. Recent SEA approaches tend to enhance speech quality and intelligibility, because improving these two attributes is critical for normal people and hearing impairments. Therefore, this thesis aims to restore speech signals from corrupted signal with minimum MN and best trade-off between Residual Noise (RN) and Signal Distortion (SD). First, a new transform based on new orthogonal polynomials, called the Discrete Krawtchouk–Tchebichef Transform (DKTT), is presented. DKTT exhibits superior compaction and localization properties that affect noise extraction process. Second, a noise classification method is adopted to identify the types of additive noise. Then, three optimum types of parameters are determined based on noise type. The subsequent phase of the developed system involves the proposed non-linear speech estimator. It is based on the Minimum Mean Square Error (MMSE) and the low-distortion approaches. The analytical solution is derived from the assumption that speech and noise components can be modeled based on a combination between Gamma and Laplacian distributions. These types of combination are used first in the developed SEA. Afterward, the second proposed linear estimator has been proposed mainly to reduce the effects of MN. Finally, the inverse of DKTT is applied to regain the clean signal back. To demonstrate the capability of the proposed system, clean speech sentences are selected from the TIMIT dataset. Moreover, eleven types of noise are chosen from the NOISEX-92 dataset, in addition to speech-shaped noises. These noises are the most dominate in the real world. Comparison results reinforce the improvement in quality and intelligibility measurements with reducing of MN level. The objective measurements are including Perceptual Evaluation of Speech Quality (PESQ), Frequency-Weighted Segmental

Signal-to-Noise Ratio (FWSNR), the Coherence Speech Intelligibility Index (CSII), Short-Time Objective Intelligibility measure (STOI), along with three types of composite measures, namely, Signal distortion (SIG), Back-ground intrusiveness (BAK), and Overall quality (OVL). The improved SEA demonstrated an improvement in nearly all the aforementioned quality and intelligibility measures for different types of noise and five levels of signal-to-noise ratio (SNR), i.e., −10, −5, 0, 5, and 10 dB. In white noise, for example, the average absolute improvements and their corresponding percentage values of the system performance in terms of PESQ, OVL, STOI, and FWSNR in (dB) for the five SNR levels are 0.37 (17.3%), 0.37 (24.7%), 0.59 (7.8%), and 0.06 (7.7%), respectively. For cockpit noise, the improvements are 0.22 (10.6%), 0.18 (10.5%), 1.5 (23.3%), and 0.07 (9.5%), respectively. For Speech-Shaped noise, the improvements are 0.23 (11.3%), 0.17 (9.1%), 2.05 (31.6%), and 0.05 (7.8%), respectively. Moreover, the classification accuracy has been reached to 99.44%. This work contributed in developing a new transform, finding a new speech and noise models, introducing new linear and non-linear estimators with their adaptively smoothing parameter to get good noise reduction. As a conclusion, the proposed SEA enhances and improves noisy signals and regain clean signals with less RN and SD, reducing MN level. Moreover, best improvement in quality and intelligibility properties is obtained particularly in high noise levels.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia Sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

RANGKA KERJA PENINGKATAN PERTUTURAN MENGGUNAKAN JELMAAN DISKRIT *KRAWTCHOUK-TCHEBICHEF*

Oleh

BASHEERA M. MAHMMOD

Mei 2018

Pengerusi: Profesor Madya Abd Rahman Ramli, PhD
Fakulti: Kejuruteraan

Pertuturan dianggap sebagai cara utama untuk berinteraksi di kalangan manusia. Isyarat pertuturan menghadapi senario yang sukar semasa penghantaran, seperti gangguan dan hingar tambahan, yang membawa kepada menghasilkan isyarat hingar. Oleh itu, Algoritma Peningkatan Pertuturan (APP) lasak yang menyekat hingar tanpa gangguan kepada isyarat asal adalah perlu. Menyingkirkan hingar tanpa menyebabkan herotan isyarat pertuturan adalah tugas yang mencabar. Selain itu, hingar yang menjengkelkan terhasil selepas proses peningkatan, yang dikenali sebagai Hingar Musik (HM), perlu dibuang. APP terkini cenderung untuk meningkatkan kualiti pertuturan dan kebolehfahaman, kerana menambahbaikan kedua-dua ciri ini adalah kritikal untuk orang normal dan yang terjejas pendengaran. Oleh itu, matlamat tesis ini adalah untuk memulihkan isyarat pertuturan dari isyarat yang rosak dengan HM minimum dan keseimbangan terbaik antara Sisa Hingar (SH) dan Isyarat Herotan (IH). Pertamanya, jelmaan baru berdasarkan polinomial ortogonal baru, dipanggil Jelmaan Diskret Krawtchouk-Tchebichef (TDKT) diperkenalkan. TDKT mempamerkan pemadatan lebih baik dan sifat penyetempatan yang memberi kesan kepada proses penyarian hingar. Keduanya, kaedah pengklasan hingar digunakan untuk mengesalpasti jenis hingar tambahan. Kemudian, tiga parameter yang optimum ditentukan berdasarkan jenis hingar. Fasa seterusnya bagi sistem yang dicadangkan merangkumi penganggar pertuturan tak linear yang berdasarkan pendekatan Min Ralat Kuasa Dua Minimum (MRKDM) dan herotan rendah. Penyelesaian analisis diperolehi berdasarkan andaian bahawa komponen pertuturan boleh dimodelkan sebagai gabungan Gamma dan taburan Laplacean. Jenis gabungan ini digunakan terlebih dahulu dalam APP yang dibangunkan. Selepas itu, penganggar linear yang kedua dicadangkan untuk mengurangkan kesan-kesan HM. Akhirnya, TDKT songsang diaplikasikan untuk mendapatkan kembali isyarat bersih yang asal. Untuk menunjukkan keupayaan sistem yang dicadangkan, ayat-ayat pertuturan bersih dipilih daripada set data TIMIT. Tambahan lagi, sebelas jenis hingar dipilih daripada set data NOISEX-92, sebagai tambahan kepada hingar berbentuk pertuturan. Keputusan perbandingan mengukuhkan peningkatan pengukuran kualiti objektif dan

kebolehfahaman dengan mengurangkan tahap HM. Pengukuran objektif seperti Penilaian Persepsi Kualiti Pertuturan (PPKP), kekerapan berwajaran nisbah isyarat-hingar Segmental (kbSNR), Indeks Kebolehfahaman Pertuturan Koheren (IKPK), pengukuran Objektif Kebolehfahaman Masa-Pendek (OKMP), bersama dengan tiga jenis pengukuran komposit iaitu Isyarat herotan (IH), campur tangan latar-belakang (CTLB), dan Kualiti Keseluruhan (KS). APP yang ditambahbaik menunjukkan peningkatan dalam hampir semua pengukuran kualiti dan kejelasan yang disebut untuk jenis hingar berbeza dan untuk lima tahap SNR, iaitu -10, -5, 0, 5 dan 10 dB. Dalam hingar putih, contohnya, purata peningkatan mutlak dan nilai peratusan yang berkaitan dengan prestasi sistem dari segi PPKP, KS, OKMP, dan kbSNR untuk lima tahap SNR adalah masing-masing 0.37 (17.3%), 0.37 (24.7%), 0.59 (7.8%) dan 0.06 (7.7%). Untuk hingar kokpit, peningkatan adalah masing-masing 0.22 (10.6%), 0.18 (10.5%), 1.5 (23.3%), dan 0.07 (9.5%). Untuk hingar berbentuk pertuturan, peningkatan adalah masing-masing 0.23 (11.3%), 0.17 (9.1%), 2.05 (31.6%), dan 0.05 (7.8%). Malahan, proses ketepatan pengelasan telah mencapai 99.44%. Kerja ini menyumbang dalam membangunkan jelmaan baru, mencari model pertuturan dan hingar baru, memperkenalkan penganggar linear dan tak linear baru dengan parameter pelicinan penyesuaian mereka untuk mendapatkan pengurangan hingar yang baik. Sebagai kesimpulan, APP yang dicadangkan meningkatkan dan menambahbaik isyarat hingar dan mendapatkan isyarat bersih dengan kurang SH dan IH, minimum HM. Tambahan pula, peningkatan terbaik dalam ciri-ciri kualiti dan kebolehfahaman diperolehi terutamanya dalam tahap hingar tinggi.

ACKNOWLEDGEMENTS

Firstly, I would like to express my special gratitude and appreciation to my supervisor Associate Professor Abdul Rahman Ramli for his guidance throughout my Ph.D. study at UPM. His broad knowledge, interest and intuition in my study affected greatly of the way I do my research throughout my work. He is a kind person and made ever thing easy to me by his support and respectful comments during the years of my PhD study. He is a special person and loved by everyone. He supported me to be as independent researcher and gave the opportunity to be what I am today. Indeed, thanks a lot Prof. Ramli for your kindness, wisdom, respect, and support during my difficult moments in my work.

I would also like to thank my supervisory committee members: Associate Prof. Dr. Syed Abdul Rahman Al-Haddad, Prof. Dr. Iqbal M. Saribaan and Associate Prof. Dr. Wissam A. Jassim for served as my best lectures in my work field. They are a critical thinker and excellent academician who believe in high quality standard when analyzing different issues in their fields. And great thanks go to Prof. Dr. Wissam A. Jassim who had a really informative discussion in my defense and his significant ideas and suggestions which are very useful in improving and extending this work. He spent a lot of time to give me valuable comments and instruct me to develop and enhance this thesis during my study period.

Special thanks go to my husband, Sadiq H. Abdulhussain, for his great assistance and support in everything in all of these years of Ph.D. work. Actually, I could not reach what I am today without him. He was very kindness, praiseworthy and kind hearted.

Last but not least, I would like to extend my heart felt appreciation to my beloved family, my children: Mariam, Jafar, and Toqa, for their believing in me and their understanding and encouragement was in the end what made this work possible. My big gratitude to my father and mother which are with me all days and nights of my work and for their asking and praying. And million thanks to my brothers Basheer and Laith, and my sisters Bashaar, for their continuous inspiration, support and trust me to continue this study.

Without any hesitation, I would like to express my overwhelming to ministry of higher education and scientific research in Iraq for making my dream of acquiring higher education a reality and pursuing a PhD degree.

At the end, a special appreciations go to UPM for providing me and all its students a peaceful, complete, and clam academic environment to study and finishing the scientific path.

Basheera M. Mahmmod

I certify that a Thesis Examination Committee has met on 15 May 2018 to conduct the final examination of Basheera M. Mahmmod on her thesis entitled "A Speech Enhancement Framework Using Discrete Krawtchouk-Tchebichef Transform" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Wan Azizun binti Wan Adnan, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

Sharifah Mumtazah Syed Ahmad Abdul Rahman, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Shaiful Jahari bin Hashim, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Internal Examiner)

Dhiya Al-Jumeily, PhD
Professor
Liverpool John Moores University
United Kingdom
(External Examiner)

RUSLI HAJI ABDULLAH, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 30 July 2018

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Abdul Rahman b. Ramli, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Chairman)

M. Iqbal b. Saripan, PhD
Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Syed Abdul Rahman Al-Haddad, PhD
Associate Professor
Faculty of Engineering
Universiti Putra Malaysia
(Member)

Wissam A. Jassim, PhD
Research Fellow
Faculty of Engineering
University of Deblin, Ireland
(Member)

ROBIAH BINTI YUNUS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature:____________________________ Date:_________________

Name and Matric No.: Basheera M. Mahmmod, GS40254

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: ______________________________

Name of Chairman of Supervisory Committee: Associate Professor Dr. Abdul Rahman b. Ramli

Signature: ______________________________

Name of Member of Supervisory Committee: Professor Dr. M. Iqbal b. Saripan

Signature: ______________________________

Name of Member of Supervisory Committee: Associate Professor Dr. Syed Abdul Rahman Al-Haddad

Signature: ______________________________

Name of Member of Supervisory Committee: Dr. Wissam A. Jassim

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
AU	Arbitrary Unit
BAK	Five-point Scale of Background Intrusiveness
CI	Cochlear Implant
COMLSA	Optimally Modified Log-Spectral Amplitude method based on noise Classification
CSII	Coherence Speech Intelligibility Index
DCT	Discrete Cosine Transform
DCP	Distribution Controlling Parameter
DFT	Discrete Fourier Transform
DGW	Dual-Gain Wiener
DKT	Discrete Krawtchouk Transform
DKTT	Discrete Krawtchouk-Tchebichef Transform
DMMSE	Dual-MMSE
DSP	Digital Signal Processing
DTT	Discrete Tchebichef Transform
FIR	Finite Impulse Response
FWSNR	Frequency-Weighted segmental SNR
HDLCMV	Harmonic Decomposition Linearly Constrained Minimum Variance
HMM	Hidden Markov Model
IDFT	Inverse Discrete Fourier Transform
IDKTT	Inverse Discrete Krawtchouk-Tchebichef Transform
IIR	Infinite Impulse Response
ITU-T	International Telecommunication Union-Telecommunication
KLT	Karhunen–Loève Transform
KP	Krawtchouk Polynomial
LBSE	Linear Bilateral Super-Gaussian Estimator
LCMV	linearly Constrained Minimum Variance
LLR	Log-Likelihood Ratio quality measure
LOSB	Length of Sub-Band
LPC	Linear Predictive Coefficient
ML	Maximum-Likelihood
MLE	Maximum Likelihood Estimator
MMSE	Minimum Mean Square Error

MMSE-STSA	Minimum Mean Square Error-Short Time Spectral Amplitude
MN	Musical Noise
MOS	Mean Opinion Score
MSE	Mean Square Error
MVDR	Minimum Variance Distortionless Response
NBSE	Non-Linear Bilateral Super-Gaussian Estimator
NLBSE	Non-Linear Linear Bilateral Super-Gaussian Estimator
NN	Neural Network
ODMVDR	Orthogonal Decomposition Based Minimum Variance Distortion less Response
OP	Orthogonal Polynomial
OVL	Five-point Scale of the Mean Opinion Score of Overall Quality
PDF	Probability Density Function
PESQ	Perceptual Evaluation of Speech Quality
PSD	Power Spectral Density
PSS	Power Spectrum Subtraction
RMSE	Root Mean Square Error
RN	Residual Noise
ROI	Region of Interest
SD	Speech Distortion
SE	Speech Enhancement
SEA	Speech Enhancement Algorithms
SIG	Five-point Scale of Signal Distortion
SII	Speech Intelligibility Index
SNR	Signal to Noise Ratio
SP	speech pauses
SS	Spectral Subtraction
SSA	Spectral-Subtractive algorithms
SSF	Spectral Subtraction Filter
STFT	Short Time Fourier Transform
STOI	Short-Time Objective Intelligibility measure
STSA	Short Time Spectral Amplitude
SVM	Support Vector Machine
TP	Tchebichef Polynomial

CHAPTER 1

INTRODUCTION

1.1 Introduction

Speech, means of interaction among human beings, plays a central role in recent communication fields. However, different types of environmental noise always interfere with speech signals, thereby producing degraded and annoying signals for the human auditory system. To make the exchange of speech possible in a noisy environment, quality and intelligibility attributes must be maintained, which are challenging tasks. Quality and intelligibility are necessary to guarantee human satisfaction. Accordingly, the demand for robust Speech Enhancement Algorithms (SEA) has increased. Accurate noise suppression with minimum levels of Speech Distortion (SD) and Residual Noise (RN) should be considered to fulfill Speech Enhancement (SE) requirements. The optimization of these requirements depends on a robust SEA that can deal with the aforementioned circumstances and contribute to perfect outcomes. This chapter first highlights the basic aspects of SE by presenting different types of noise and SEA. Then, it provides a brief description of different problems that have been experienced recently by SEA. In addition, the main objectives of this thesis are presented. In addition, the methodology and the research scope are briefly described. The outline of the thesis is provided at the end of this chapter.

1.2 Background

The 18th century was the epoch of the Industrial Revolution and great mechanical systems; the 19th century was the era of the steam engine; the 20th century was marked by the development of important technologies for information gathering, exchange, processing, and distribution (Tanenbaum, 2003). During the 20th century, new technologies were occasionally introduced; some of these technologies were immediately accepted, others were rejected, whereas others went through a slow but relentless adoption curve (Mahmood, 2012). Advances in technology aim to increase scientific development and to achieve overall human satisfaction and comfort. Digital Signal Processing (DSP) technology has been an active research area in recent years and has addressed the aforementioned issues. Speech processing, which has numerous applications in daily life, is one of the most important fields of DSP. Speech is considered a vital means of exchanging information among humans and between humans and machines (Upadhyay & Karmakar, 2015).

In various life circumstances, as the distance between the source and the receiver of speech increases, speech signals become increasingly corrupted by surrounding noises, interferences, and echoes, and expose to distortions introduced by communication media. This condition leads to decrease speech intelligibility and reduces the effectiveness of communication quality. Notably, quality refers to "how" a speaker produces an utterance, whereas intelligibility refers to "what" the speaker has said, i.e., the meaning or content of the spoken words (Loizou, 2013). Consequently, SEA is introduced to reduce disturbance that results from such situations. The general types of noise that affect speech signals are shown in Figure 1.1.

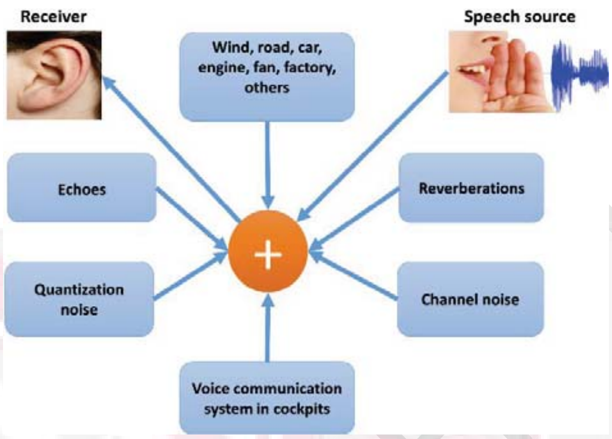

Figure 1.1 : Different types of environment noise (Loizou, 2013)

SEA has many objectives, including increasing intelligibility, improving quality, solving the noise pollution problem, increasing speech recognition accuracy, and reducing listener fatigue. However, its key objective is to regain clean speech signals from noisy signals degraded by additive background noise (Lim & Oppenheim, 1979; Yousheng & Jianwen, 2014). Recently, SEA are considered substantial issues in our daily life, because they are applied in different key applications such as hands-free telephony, speech recognition, military and air-ground communication, VoIP, teleconferencing systems, performance improvement of hearing aids and cochlear implants, and speech coding (Abutalebi & Rashidinejad, 2015; Jassim et al., 2014; Liu et al., 2016; Loizou, 2013; Zoulikha & Djendi, 2016). The importance of SEA is increased by the developments in speech processing applications. All existing SEA cannot completely suppress background noise because suppression is performed in relatively different degrees based on aspects such as algorithm effectiveness, application at hand, noise source characteristics or interference, relationship between clean signal and noise, and the number of available microphones (Loizou, 2013). Practically, speech enhancement task becomes easier with a large number of microphones. In an ideal SEA, quality and intelligibility must be significantly improved to reach the levels of clean speech signals; however, this issue remains unachievable. In particular, noise can be reduced to an acceptable level but at the expense of increasing SD, which degrades speech intelligibility. Therefore, designing an effective SEA without introducing a noticeable distortion to speech with minimum RN is considered a challenging task (Loizou, 2013).

The general degradation and denoising processes are illustrated in Figure 1.2. Additive noise is added to clean speech signal in an uncontrolled manner when speech is transmitted from source to destination. Numerous approaches deal with additive noise because it is the most widely affected noise type in real environments. SEA have different classifications. Several studies have categorized SEA into two main groups namely; supervised and unsupervised methods are based on using a training set. These training datasets are based on a large number of observed samples of a signal (Chehrehsa & Moir, 2016; Mohammadiha et al., 2013).

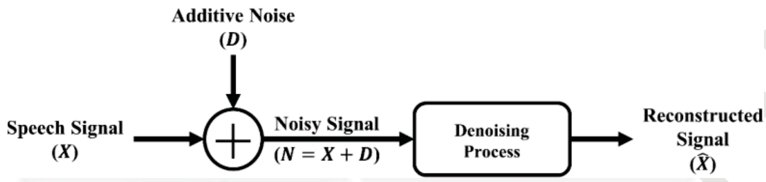

Figure 1.2 : General processes of speech degradation and restoration (Loizou, 2013)

Another classification of SEA depends on the type of processing domain; time domain (Lim et al., 1978) and frequency domain (Jassim et al., 2014). Notably, the frequency domain compresses a substantial amount of information in a signal into specific coefficients that facilitate the removal of corrupting noise. Transform-based SEA have different types, such as Discrete Fourier Transform (DFT) (Ephraim & Malah, 1984), Wavelet Transform (WT) (Jayakumar & Sathidevi, 2016), and Discrete Cosine Transform (DCT) (Ding et al., 2011; Soon, 2003; Soon et al., 1998; Wei et al., 2016).

Research from other perspectives has divided SEA into three main classes based on the techniques used to process information. The first class includes spectral-subtractive algorithms (SSA) which is are considered the simplest and first SEA (Boll, 1979). The second class, which is based on statistical models and optimization criteria, places the problems of speech enhancement in statistical and estimation frameworks. Examples for this class are Wiener Filtering (WF) algorithms (Jassim et al., 2014; Lim & Oppenheim, 1979) and Minimum Mean Square Error (MMSE) algorithms (Chen & Loizou, 2007; Ephraim & Malah, 1984). The third category, called subspace algorithms, is related to linear algebra theory (Ephraim & Van Trees, 1995; Hu & Loizou, 2003).

Other classifications depend on the application and number of input channels. However, most recent SEA has demonstrated good performance and considered promising methods, some problems remain. In general, the performance of SEA is reduced by changing surrounding conditions. Therefore, they become an inadequate choice for dealing with all life situations. SEA performance is characterized by RN and SD levels. For example, SSA is considered the simplest and most effective method but it comes at the price of suffering from Musical Noise (MN), which is a perceptually annoying spectral artifact (Paliwal et al., 2010). Moreover, the performance of modern communication devices in noisy environments is better using SEA; nevertheless, most of these algorithms exhibit a trade-off between the RN and SD levels (Fenghua et al., 2010). These algorithms do not concurrently improve speech perceptual aspects, overall quality, and intelligibility. The aforementioned trade-off remains a serious problem in modern research; hence, it should be considered when designing SEA. An important parameter, called the smoothing parameter, controls the trade-off between the smoothing degree of a priori signal-to-noise ratio (SNR) during speech-absent segments and transient distortion level incurred during signal onsets (Loizou, 2013). This parameter primarily affects the estimated value of MN, an annoying artifact noise, which remains an issue in many state-of-the-art methods.

1.3 Problem Statement

An effective and robust SEA with robust gain function is required to estimate clean speech signals properly. Designing the gain function is one of the most crucial parts in proposing a statistical SEA (Tsao & Lai, 2016). Moreover, the key challenge in designing an effective SEA is to suppress noise without presenting any perceptible distortion in speech signals (Loizou, 2013). The recent speech applications increase the demand for simultaneously enhancing speech quality and intelligibility attributes (Kolbæk et al., 2017). In addition, enhancing noisy signals leads to the generation of artifact MN, which is extremely annoying to the human ears and needs optimum solution. Optimization SEA, which reduces computational cost via a forceful transform with specific properties, must also be considered. The following problems have been addressed as points throughout this thesis.

1- In existing SEA, the adopted Probability Density Functions (PDF) to model speech and noise transform coefficients is assumed Gaussian or super Gaussian functions (Abutalebi & Rashidinejad, 2015; Chen & Loizou, 2007; Jassim et al., 2014). These PDFs provide allowable mapping pattern to speech and noise signals(De Abreu et al., 2017). However, they are not always true in real life situations because of variable characteristics of noise (Yuan & Xia, 2015). Moreover, the RN and SD levels remain unacceptable based on these densities (Lee & Lee, 2016; Wei et al., 2016). Realistically, this issue stills a serious problem because there are no perfect models represents speech and noise coefficient distribution. Therefore, noise types need to be classified accurately to improve the generalizability of SEA performance. This issue remains a controversial topic in different investigations and studies.
2- To obtain an optimum SEA that is suitable for different life circumstances, quality and intelligibility must be improved concurrently (Guang-yan et al., 2009; Loizou, 2013), which is a challenging task (Kolbæk et al., 2017). Conventional SEA have reported improvement in quality but not necessarily intelligibility (Bhat et al., 2017; Gaikwad & Vasekar, 2015). Where, only a few SEA performed enhancement on intelligibility (Goehring et al., 2017). In low SNR condition, noise is reduced but at the expense of introducing damage to speech intelligibility (Loizou, 2013). Since weak harmonic components of speech are distorted because of over-attenuation (Ding, 2011; Ding et al., 2010; Krawczyk-Becker & Gerkmann, 2016). Therefore, simultaneously improving of quality and intelligibility remains a significant topic that must be considered.

3- Although many SEA reduce background noise, a serious problem is generated in the enhanced signal namely, MN, which is one of the most annoying acoustic phenomena for the human ears (Hamon et al., 2017). Most SEA reduce noise effectively but suffer from artificial distortion (Hamon et al., 2017; Saruwatari, 2015). Other studies have attempted to address this phenomenon but complete removal has not been realized. This artificial distortion is more noticeable in the unvoiced speech segments because noise power is comparable with speech power. Therefore, the minimization of MN without affecting speech in a certain manner is extremely difficult (Loizou, 2013; Wei et al., 2016). Therefore, this serious drawback must be addressed.

1.4 Aim and Objectives

To fulfill the requirements of SE trends, the performance of the proposed algorithm should be specified, such that noise will be suppressed perfectly. However, in any enhancement process, a low RN level leads to a high SD level, in addition to the appearance of MN. Therefore, this thesis aims to obtain an optimum SE process by minimizing the RN level without sacrificing the quality and intelligibility of speech signals. Moreover, managing noisy signals in various scenarios of noise effects is vital. Consequently, the objectives of this thesis consist of three specific points, which are listed below.

1- To find perfect models for speech and noise coefficients distribution based on accurate noise classification method and new discrete transform that exhibits robust properties.
2- To propose and derive a new non-linear low-distortion MMSE estimator (first-stage estimator) based on a new composite of super- Gaussian PDF that can improve quality and intelligibility concurrently.
3- To propose and derive a post-processing filtering, a new linear low-distortion MMSE estimator (second-stage estimator), based on a new composite of super-Gaussian PDF that can tracking the MN and eliminate it.

1.5 Brief Methodology

This work is divided into five phases, as shown in Figure 1.3, to provide comprehensive details regarding the proposed SEA. In the first phase, the observation signal is transformed through the proposed transform, which is called the Discrete Krawtchouk-Tchebichef Transform, (DKTT), to obtain the noisy moment coefficients. The observed signal in the transformed domain meets the requirements of signal compaction and localization properties. Furthermore, in the proposed real transform there is no need to consider the phase of the noisy signal to obtain the phase of clean signal, thereby reducing the computational complexity of the proposed SEA. This is because real transform has sign information instead of phase information. The distribution of speech and noise coefficients of the suggested transform provides a better fit with the assumed composite distributions, which will be explained in detail in

Chapter Three. Figure 1.3 illustrates the comprehensive paradigm of the proposed algorithm.

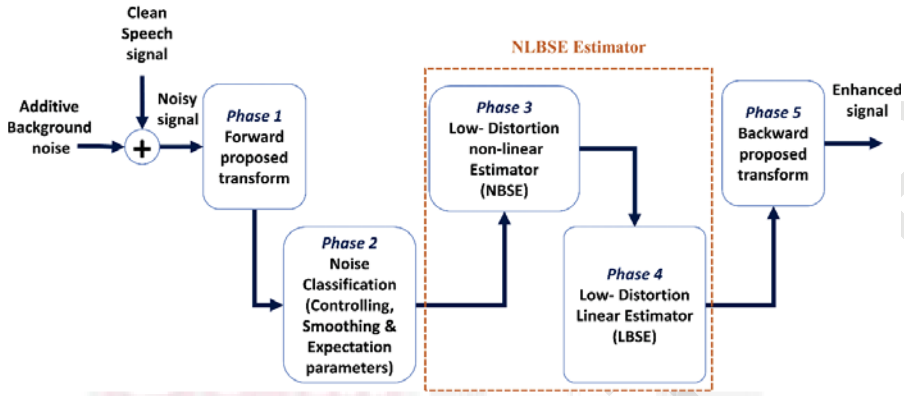

Figure 1.3 : General paradigm of the methodology stages.

In the second phase, a noise classification technique is suggested to optimize the selection of an accurate matching between noise DKTT sample distribution and the assumed density functions. Three sets of significant parameters are determined based on noise type, which are as follows: Distribution Controlling Parameters (DCP) that which determine the external form and appearance characteristic of PDF noise; smoothing parameter ($\propto$), which quantifies the value of the estimated a priori SNR, and consequently, the MN level; and an expectation parameter (p_E), which is suggested to improve the attenuation and amplification levels and further minimize the effect of MN. $\propto$ and p_E lead to lower MN perception by reducing narrow spectral peak perception, through the decrease of spectral excursions. The aforementioned parameters are included in the third and fourth phases, where speech estimators based on low-distortion MMSE are proposed and derived. The first optimum non-linear estimator works in cascade with an optimum linear MMSE estimator (a post-processing filtering). These two cascade estimators work together to suppress noise with minimum distortion and the best performance under different conditions of the underlying speech signal. The derivation of these two estimators is based on a composite Super-Gaussian distribution model for speech and noise signals. The assumption considered in this thesis for speech and noise modeling, which is a new approach in SEA, deals with the constructive and destructive events of noise. The two analytical solutions for the two proposed estimators are termed as Non-Linear Bilateral Super-Gaussian Estimator (NBSE) and Linear Bilateral Super-Gaussian Estimator (LBSE) as shown in Figure 1.3. The cascade combination between NBSE and LBSE is called NLBSE. In the final phase, the inverse of the proposed transform (IDKTT) is applied to the estimated signal to return it back to the time domain. The enhanced speech signal is then obtained. This signal is highly correlated with the original speech signal. The details of each phase are analyzed and explained in Chapter Three.

1.6 Main Contributions of the Thesis

On the basis of the five aforementioned phases, the main contributions of this thesis are presented as follows.

A. Developing a new discrete transform
A new discrete transform, called DKTT, is derived. It has two significant points. First, it is derived from a new set of orthogonal functions. Second, it is more robust against different types of noise and exhibits good energy compaction and localization properties.

B. Building new speech and noise models
One of the main contributions of this work is building a new model for clean and noise signals that accurately matches their transformed coefficient distribution. Noise models relatively change based on noise type. New parameters are suggested to control the appearance characteristics of the assumed distribution by regulating amplitude and standard deviation.

C. Introducing a relatively changeable smoothing parameter
This work proposes a relatively changeable parameter $\propto$ to control the MN level based on noise type because. This parameter is an important factor in the decision-directed approach, where the former is used to estimate a *priori* SNR.

D. Introducing a new non –linear filter formulation (NBSE)
This work contributes an accurate non-linear estimator, i.e., NBSE, with two gain formulas for constructive and destructive events. This estimator is formulated by assuming new appropriate Super-Gaussian distributions for speech and noise signals and is based on MMSE sense and low-distortion approaches. This estimator minimizes the trade-off between quality and intelligibility of speech.

E. Introducing a new linear filter formulation (LBSE)
A new linear estimator, i.e., LBSE, based on a low-distortion approach is derived. This estimator is based on the notion of WF but with dual gain formulas and new PDF for speech and noise signals. It has two essential functions. Its first function is to reinforce the attenuation and amplification levels. Its second function is to minimize MN that appeared in the enhanced signal of the NBSE output. This estimator has a new parameter, p_E, which affects the reduction of MN.

F. Significant reduction in additive noise
Significant noise reduction is applied according to the proposed SEA developed in Chapter Three. First, a significant noise classification technique that produces accurate models for speech and noise signals using Super-Gaussian distribution. Second, two powerful estimators, i.e., NBSE and LBSE, based on low-distortion approaches and new speech and noise models for enhancing noisy signal with minimum level of RN and SD are derived based on a new discrete transform. Moreover, the proposed SEA effectively addresses the problem of MN and over-attenuation. The other significant point is the simultaneous improvement of quality and intelligibility attributes. A flexible estimator will optimize enhancement performance without complex computations and achieve the objectives of SE.

1.7 Thesis Scope

To achieve the objectives of this thesis, linear and non-linear MMSE estimators based on a new model for speech and noise that used zero mean Super-Gaussian (Laplacian and Gamma) distribution, are developed. Hamming window with 75% overlap is used to truncate the noisy signal and improve the proposed SEA performance. Where Hamming window provides the best tradeoff between resolution and spectral leakage.

This thesis is directed toward enhancing speech signals degraded by statistically uncorrelated and independent additive noise. Focus is on suggesting an optimal estimator, NLBSE, with two ideal mutually exclusive interference cases. The estimation of a polarity parameter is beyond the scope of this work because this thesis deals with optimum estimators in the given events. Moreover, the estimation of the noise power spectrum was obtained from the initial six frames segment of each file sentence. Eleven types of noise, including stationary and non-stationary noise, are selected to be implemented in this work. Ten are selected from the NOISEX-92 dataset (Varga & Steeneken, 1993), namely, white, pink, cockpit (F16), cockpit (buccaneer), factory, babble, engine room noise (destroyer), operation room noise (destroyer), military vehicle noise (leopard), and military vehicle noise (M109), in addition to speech-shaped noise (Loizou, 2013). These different types of noise are selected since they are the more dominate in real life. The use of this number of noise adequately verifies the objective and performance of this thesis. Moreover, five levels of SNR, i.e., −10, −5, 0, 5, and 10 dB, for each type of noise are used to confirm the capability of this work to be applied to different noise conditions. Only the noisy signal is assumed to be available, which contains both clean speech and additive noise. Furthermore, a single channel, which is the general case, is assumed to be available. This situation constitutes one of the most challenging problems, and more complex than multiple microphone systems because of the absence of access to the reference microphone that picks up the noise signal. For evaluation purposes, the objective test has been implemented. Moreover, the well-known TIMIT dataset is used, which is a corpus of American English speakers of lexically and phonemically transcribed speech for different dialects and genders.

1.8 Thesis Organization

This thesis is divided into five chapters including this chapter. Chapter Two presents a comprehensive review of existing SEA, different transform-based SEA, various problems and the techniques used to solve them, MN, noise classification, and different objective measures are also included and discussed in this chapter. Chapter Two ends by highlighting the main gaps in recent research that should be considered when proposing a SEA with high specifications.

Chapter Three provides a complete description of the research methodology steps. The work flow of this study is divided into five sections. The first section presents the derivation of the new discrete transform that facilitates and positively improves the noise extraction process. The second section explains the new idea of the noise classification technique and its role in finding the optimum model for noise signals. The third section demonstrates the derivation of the analytical solution for the non-linear proposed low-distortion MMSE estimator based on Super- Gaussian prior. The fourth section describes the derivation of the analytical solution for the linear proposed low-distortion estimator based on Super- Gaussian prior. The final section presents the characteristics of the proposed linear and non-linear estimators.

Chapter Four presents the results and discussion of the proposed algorithm. A comparison is performed based on two aspects. In the first aspect, the proposed algorithm is compared with its different versions. In the second aspect, the proposed algorithm is compared with other state-of-the-art algorithms. Different quality and intelligibility measurements are used in the comparison assessment. Moreover, waveform analysis and spectrogram decomposition techniques are presented to provide comprehensive explanations of the remarkable results and robustness of the proposed SEA.

This thesis ends with a summary and conclusion in Chapter Five. Potential ideas for future work are also suggested.

REFERENCES

Abd El-Fattah, M., Dessouky, M. I., Diab, S. M., & Abd El-Samie, F. E.-S. (2008). Speech enhancement using an adaptive wiener filtering approach. *Progress In Electromagnetics Research M, 4*, 167-184.

Abdulhussain, S. H., Ramli, A. R., Al-Haddad, S. A. R., Mahmmod, B. M., & Jassim, W. A. (2017). On Computational Aspects of Tchebichef Polynomials for Higher Polynomial Order. *IEEE Access, 5*, 2470-2478.

Abdulhussain, S. H., Ramli, A. R., Al-Haddad, S. A. R., Mahmmod, B. M., & Jassim, W. A. (2018). Fast Recursive Computation of Krawtchouk Polynomials. *Journal of Mathematical Imaging and Vision, 60*(3), 285-303.

Abdulhussain, S. H., Ramli, A. R., Mahmmod, B. M., Al-Haddad, S., & Jassim, W. A. (2017). Image edge detection operators based on orthogonal polynomials. *International Journal of Image and Data Fusion, 8*(3), 293-308.

Abdulhussain, S. H., Ramli, A. R., Saripan, M. I., Mahmmod, B. M., Al-Haddad, S. A. R., & Jassim, W. A. (2018). Methods and Challenges in Shot Boundary Detection: A Review. *Entropy, 20*(4), 214.

Abutalebi, H. R., & Rashidinejad, M. (2015). Speech enhancement based on β-order MMSE estimation of Short Time Spectral Amplitude and Laplacian speech modeling. *Speech Communication, 67*, 92-101.

Achirul Nanda, M., Boro Seminar, K., Nandika, D., & Maddu, A. (2018). A Comparison Study of Kernel Functions in the Support Vector Machine and Its Application for Termite Detection. *Information, 9*(1), 5.

Al Banna, T. H. (2008). *A Hybrid Speech Enhancement Method using Optimal Dual Gain Filters and EMD Based Post Processing.* Bangladesh University of Engineering And Technology.

Alam, M. J., & O'Shaughnessy, D. (2011). Perceptual improvement of Wiener filtering employing a post-filter. *Digital Signal Processing, 21*(1), 54-65.

Aroudi, A., Veisi, H., Sameti, H., & Mafakheri, Z. (2015). Speech signal modeling using multivariate distributions. *EURASIP Journal on Audio, Speech, and Music Processing, 2015*(1), 1-14.

Benesty, J., & Chen, J. (2011). *Optimal Time-Domain Noise Reduction Filters*: Springer.

Bhat, G. S., Shankar, N., Reddy, C. K., & Panahi, I. M. (2017). *Formant frequency-based speech enhancement technique to improve intelligibility for hearing aid users with smartphone as an assistive device.* Paper presented at the 2017 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT).

Boll, S. F. (1979). Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech and Signal Processing., 27*(2), 113-120.

Bolner, F., Goehring, T., Monaghan, J., van Dijk, B., Wouters, J., & Bleeck, S. (2016). *Speech enhancement based on neural networks applied to cochlear implant* *coding strategies.* Paper presented at the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery, 2*(2), 121-167.

Cappé, O. (1994). Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor. *IEEE transactions on Speech and Audio Processing, 2*(2), 345-349.

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST), 2*(3), 27.

Chehrehsa, S., & Moir, T. J. (2016). Speech enhancement using Maximum A-Posteriori and Gaussian Mixture Models for Speech and Noise Periodogram estimation. *Computer Speech & Language, 36*, 58-71.

Chen, B., & Loizou, P. C. (2005). *Speech enhancement using a MMSE short time spectral amplitude estimator with Laplacian speech modeling.* Paper presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings.(ICASSP'05). .

Chen, B., & Loizou, P. C. (2007). A Laplacian-based MMSE estimator for speech enhancement. *Speech Communication, 49*(2), 134-143.

De Abreu, C. C. E., Duarte, M. A. Q., & Villarreal, F. (2017). An immunological approach based on the negative selection algorithm for real noise classification in speech signals. *AEU-International Journal of Electronics and Communications, 72*, 125-133.

Deng, F., Bao, C.-c., & Kleijn, W. B. (2015). *Sparse HMM-based speech enhancement method for stationary and non-stationary noise environments.* Paper presented at the 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

Diaz-Ramirez, V. H., & Kober, V. (2013). Robust speech processing using local adaptive non-linear filtering. *IET Signal Processing, 7*(5), 345-359.

Ding, H. (2011). *Speech Enhancement in Transform Domain.* Nanyang Technological University.

Ding, H., Lee, T., Soon, Y., Yeo, C. K., Dai, P., & Dan, G. (2015). Objective measures for quality assessment of noise-suppressed speech. *Speech Communication, 71*, 62-73.

Ding, H., & Soon, Y. (2009). *An adaptive time-shift analysis for DCT based speech enhancement.* Paper presented at the 7th International Conference on Information, Communications and Signal Processing, 2009. (ICICS). .

Ding, H., Soon, Y., & Yeo, C. K. (2010). Over-attenuated components regeneration for speech enhancement. *IEEE Transactions on Audio, Speech, and Language Processing, 18*(8), 2004-2014.

Ding, H., Soon, Y., & Yeo, C. K. (2011). A DCT-based speech enhancement system with pitch synchronous analysis. *IEEE Transactions on Audio, Speech, and Language Processing, 19*(8), 2614-2623.

Dubey, R. K., & Kumar, A. (2015). *Comparison of subjective and objective speech quality assessment for different degradation / noise conditions.* Paper presented at the 2015 International Conference on Signal Processing and Communication (ICSC).

Ephraim, Y., & Malah, D. (1984). Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech and Signal Processing., 32*(6), 1109-1121.

Ephraim, Y., & Malah, D. (1985). Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech and Signal Processing, 33*(2), 443-445.

Ephraim, Y., & Van Trees, H. L. (1995). A signal subspace approach for speech enhancement. *IEEE Transactions on Speech and Audio Processing, 3*(4), 251-266.

Erkelens, J. S., Hendriks, R. C., Heusdens, R., & Jensen, J. (2007). Minimum mean-square error estimation of discrete Fourier coefficients with generalized Gamma priors. *IEEE Transactions on Audio, Speech, and Language Processing, 15*(6), 1741-1752.

Esch, T., & Vary, P. (2009). *Efficient musical noise suppression for speech enhancement system.* Paper presented at the IEEE International Conference on Acoustics, Speech and Signal Processing, 2009. ICASSP 2009.

Fenghua, Z., Le, Y., Jian, W., & Qiang, S. (2010). *Speech signal enhancement through wavelet domain MMSE filtering.* Paper presented at the 2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering.

Fodor, B., & Fingscheidt, T. (2011). *MMSE speech spectral amplitude estimation assuming non-Gaussian noise.* Paper presented at the 19th European Signal Processing Conference. .

Gaikwad, V. M., & Vasekar, S. S. (2015). *Survey on Quality and Intelligibility Offered by Speech Enhancement Algorithms.* Paper presented at the 2015 International Conference on Computing Communication Control and Automation (ICCUBEA).

Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. S., Dahlgren, N. L., & Zue, V. (1993). TIMIT acoustic-phonetic continuous speech corpus. *Linguistic data consortium, Philadelphia.*

Gazor, S. (2004). *Employing Laplacian-Gaussian densities for speech enhancement.* Paper presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04).

Gazor, S., & Zhang, W. (2003a). A soft voice activity detector based on a Laplacian-Gaussian model. *IEEE Transactions on Speech and Audio Processing, 11*(5), 498-505.

Gazor, S., & Zhang, W. (2003b). Speech probability distribution. *IEEE Signal Processing Letters, 10*(7), 204-207.

Gazor, S., & Zhang, W. (2005). Speech enhancement employing Laplacian-Gaussian mixture. *IEEE Transactions on Speech and Audio Processing, 13*(5), 896-904.

George, A. E., Pickersgill, C., Schwerin, B., & So, S. (2016). *A comparison of estimation methods in the discrete cosine transform modulation domain for speech enhancement.* Paper presented at the Signal Processing and Communication Systems (ICSPCS), 2016 10th International Conference on.

Goehring, T., Bolner, F., Monaghan, J. J., van Dijk, B., Zarowski, A., & Bleeck, S. (2017). Speech enhancement based on neural networks improves speech intelligibility in noise for cochlear implant users. *Hearing Research, 344*, 183-194.

Guang-yan, W., Xiao-qun, Z., & Xia, W. (2009). *Musical noise reduction based on spectral subtraction combined with Wiener filtering for speech* communication. Paper presented at the IET International Communication Conference on Wireless Mobile and Computing (CCWMC 2009).

Hamon, R., Emiya, V., Rencker, L., Wang, W., & Plumbley, M. (2017). *Assessment of musical noise using localization of isolated peaks in time-frequency domain.* Paper presented at the 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).

Hasan, M., Zilany, M., & Khan, M. (2002). DCT speech enhancement with hard and soft thresholding criteria. *Electronics letters, 38*(13), 669-670.

Hasan, T., & Hasan, M. K. (2010). MMSE estimator for speech enhancement considering the constructive and destructive interference of noise. *IET Signal Processing, 4*(1), 1-11.

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). A practical guide to support vector classification.

Hsu, C.-W., & Lin, C.-J. (2002). A comparison of methods for multiclass support vector machines. *IEEE transactions on Neural Networks, 13*(2), 415-425.

Hu, Y., & Loizou, P. C. (2003). A generalized subspace approach for enhancing speech corrupted by colored noise. *IEEE Transactions on Speech and Audio Processing, 11*(4), 334-341.

Hu, Y., & Loizou, P. C. (2004). Speech enhancement based on wavelet thresholding the multitaper spectrum. *IEEE Transactions on speech and audio processing, 12*(1), 59-67.

Hu, Y., & Loizou, P. C. (2006). *Evaluation of objective measures for speech enhancement.* Paper presented at the Interspeech.

Hu, Y., & Loizou, P. C. (2008). Evaluation of Objective Quality Measures for Speech Enhancement. *IEEE Transactions on Audio, Speech, and Language Processing., 16*(1), 229-238.

J. S. Garofolo, L. F. L., W. M. Fisher, J. G. Fiscus, D. S. Pallett, N. L. Dahlgren, et al. (1993). TIMIT Acoustic-Phonetic Continuous Speech Corpus.

Jassim, W., & Raveendran, P. (2012). *Face Recognition Using Discrete Tchebichef-Krawtchouk Transform.* Paper presented at the Multimedia (ISM), 2012 IEEE International Symposium on.

Jassim, W., Raveendran, P., & Muhammad, Z. (2014). Enhancing noisy speech signals using orthogonal moments. *IET Signal Processing, 8*(8), 891-905.

Jassim, W., Raveendran, P., & Mukundan, R. (2012). New orthogonal polynomials for speech signal and image processing. *IET Signal Processing, 6*(8), 713-723.

Jassim, W. A. (2012). *A New Discrete Orthogonal Function Based on Tchebichef and Krawtchouk Polynomials and Its Applications to Speech and Image Analysis.* PhD thesis, Jabatan Kejuruteraan Elektrik, Fakulti Kejuruteraan, Universiti Malaya.

Jayakumar, E., & Sathidevi, P. (2016). Speech Enhancement Based on Noise Type and Wavelet Thresholding the Multitaper Spectrum *Advances in Machine Learning and Signal Processing* (pp. 187-200): Springer.

Jeffrey, A., & Zwillinger, D. (2007). *Table of integrals, series, and products*: Academic Press.

Jensen, J. R., Benesty, J., Christensen, M. G., & Jensen, S. H. (2012). Enhancement of single-channel periodic signals in the time-domain. *IEEE Transactions on Audio, Speech, and Language Processing, 20*(7), 1948-1963.

Kamath, S., & Loizou, P. (2002). *A multi-band spectral subtraction method for enhancing speech corrupted by colored noise.* Paper presented at the IEEE international conference on acoustics speech and signal processing.

Kang, S., Cho, S., & Kang, P. (2015). Constructing a multi-class classifier using one-against-one approach with different binary classifiers. *Neurocomputing, 149*, 677-682.

Kates, J. M., & Arehart, K. H. (2005). Coherence and the speech intelligibility index. *The Journal of the Acoustical Society of America, 117*(4), 2224-2237.

Kato, A. (2017). *Hidden Markov model-based speech enhancement.* PhD thesis, University of East Anglia.

Kirubagari, B., Palanivel, S., & Subathra, N. (2014). *Speech enhancement using minimum mean square error filter and spectral subtraction filter.* Paper presented at the International Conference on Information Communication and Embedded Systems (ICICES).

Koekoek, R., Lesky, P. A., & Swarttouw, R. F. (2010). *Hypergeometric orthogonal polynomials and their q-analogues*: Springer Science & Business Media.

Kolbæk, M., Tan, Z.-H., & Jensen, J. (2017). Speech Intelligibility Potential of General and Specialized Deep Neural Network Based Speech Enhancement Systems. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 25*(1), 153-167.

Kondo, K. (2012). Speech Quality *Subjective Quality Measurement of Speech* (pp. 7-20): Springer Berlin Heidelberg.

Krawczyk-Becker, M., & Gerkmann, T. (2016). Fundamental frequency informed speech enhancement in a flexible statistical framework. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), 24*(5), 940-951.

Krishnamurthy, N., & Hansen, J. H. L. (2009). Babble Noise: Modeling, Analysis, and Applications. *IEEE Transactions on Audio, Speech, and Language Processing, 17*(7), 1394-1407.

Kundu, A., Chatterjee, S., & Sreenivas, T. (2008). *Speech enhancement using intra-frame dependency in DCT domain.* Paper presented at the Signal Processing Conference, 2008 16th European.

Le, P. N., Ambikairajah, E., & Choi, E. (2008). *An improved soft threshold method for DCT speech enhancement.* Paper presented at the 2008 Second International Conference on Communications and Electronics.

Lee, S., & Lee, G. (2016). Noise Estimation and Suppression Using Nonlinear Function with A Priori Speech Absence Probability in Speech Enhancement. *Journal of Sensors, 2016*.

Li, R., Liu, Y., Shi, Y., Dong, L., & Cui, W. (2016). ILMSAF based speech enhancement with DNN and noise classification. *Speech Communication, 85*, 53-70.

Lim, C., Lee, S.-R., & Chang, J.-H. (2012). Efficient implementation of an SVM-based speech/music classifier by enhancing temporal locality in support vector references. *IEEE Transactions on Consumer Electronics, 58*(3).

Lim, J., Oppenheim, A., & Braida, L. (1978). Evaluation of an adaptive comb filtering method for enhancing speech degraded by white noise addition. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 26*(4), 354-358.

Lim, J. S., & Oppenheim, A. V. (1979). Enhancement and bandwidth compression of noisy speech. *Proceedings of the IEEE, 67*(12), 1586-1604.

Lin, H.-T., & Lin, C.-J. (2003). A study on sigmoid kernels for SVM and the training of non-PSD kernels by SMO-type methods. *submitted to Neural Computation, 3*, 1-32.

Lin, L., Holmes, W. H., & Ambikairajah, E. (2003). *Subband noise estimation for speech enhancement using a perceptual Wiener filter.* Paper presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). .

Lin, Z., & Goubran, R. (2003). *Musical noise reduction in speech using two-dimensional spectrogram enhancement.* Paper presented at the Haptic, Audio and Visual Environments and Their Applications, 2003. HAVE 2003. Proceedings. The 2nd IEEE Internatioal Workshop on.

Lin, Z., Xiong, Y., Cai, G., Dai, H., Xia, X., Tan, Y., & Lueth, T. C. (2018). Quantification of Parkinsonian Bradykinesia based on Axis-Angle Representation and SVM Multiclass Classification Method. *IEEE Access*, 1-1.

Liu, B., Tao, J., Wen, Z., & Mo, F. (2016). Speech Enhancement Based on Analysis–Synthesis Framework with Improved Parameter Domain Enhancement. *Journal of Signal Processing Systems, 82*(2), 141-150.

Liu, H., Zhang, R., Zhou, Y., Jing, X., & Truong, T.-K. (2017). Speech Denoising Using Transform Domains in the Presence of Impulsive and Gaussian Noises. *IEEE Access, 5*, 21193-21203.

Loizou, P. (2007). *Speech enhancement: theory and practice.*: CRC press.

Loizou, P. C. (2011). Speech quality assessment *Multimedia analysis, processing and communications* (pp. 623-654): Springer.

Loizou, P. C. (2013). *Speech enhancement: theory and practice*: CRC press.

Lu, C.-T. (2007). Reduction of musical residual noise for speech enhancement using masking properties and optimal smoothing. *Pattern recognition letters, 28*(11), 1300-1306.

Lu, Y., & Loizou, P. C. (2008). A geometric approach to spectral subtraction. *Speech Communication, 50*(6), 453-466.

Ma, J., Hu, Y., & Loizou, P. C. (2009). Objective measures for predicting speech intelligibility in noisy conditions based on new band-importance functions. *The Journal of the Acoustical Society of America, 125*(5), 3387-3405.

Mahmood, B. (2012). *Construction of an Infrastructure for RFID System.* M. Sc. Thesis, University of Baghdad, Colleague of Engineering.

Martin, R. (2002). *Speech enhancement using MMSE short time spectral estimation with gamma distributed speech priors.* Paper presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).

Martin, R. (2005a). Speech enhancement based on minimum mean-square error estimation and supergaussian priors. *IEEE transactions on speech and audio processing, 13*(5), 845-856.

Martin, R. (2005b). Statistical methods for the enhancement of noisy speech *Speech Enhancement*: Springer.

Martin, R., & Breithaupt, C. (2003). *Speech enhancement in the DFT domain using Laplacian speech priors.* Paper presented at the International Workshop on Acoustic Echo and Noise Control (IWAENC).

Milgram, J., Cheriet, M., & Sabourin, R. (2006). *"One against one" or "one against all": Which one is better for handwriting recognition with SVMs?* Paper presented at the tenth international workshop on Frontiers in handwriting recognition.

Mohammadiha, N., Smaragdis, P., & Leijon, A. (2013). Supervised and unsupervised speech enhancement using nonnegative matrix factorization. *IEEE Transactions on Audio, Speech, and Language Processing, 21*(10), 2140-2151.

Mukundan, R., Ong, S., & Lee, P. (2001a). Discrete vs. continuous orthogonal moments for image analysis.

Mukundan, R., Ong, S., & Lee, P. A. (2001b). Image analysis by Tchebichef moments. *IEEE Transactions on Image Processing, 10*(9), 1357-1364.

Nabi, W., Aloui, N., & Cherif, A. (2016). Speech enhancement in dual-microphone mobile phones using Kalman filter. *Applied Acoustics, 109*, 1-4.

Paliwal, K., Wójcicki, K., & Schwerin, B. (2010). Single-channel speech enhancement using spectral subtraction in the short-time modulation domain. *Speech Communication, 52*(5), 450-475.

Papoulis, A., & Pillai, S. U. (2002). *Probability, random variables, and stochastic processes*: Tata McGraw-Hill Education.

Quackenbush, S. R., Barnwell, T. P., & Clements, M. A. (1988). *Objective measures of speech quality*: Prentice Hall.

Rashidi-nejad, M., & Abutalebi, H. R. (2012). *Speech enhancement using adaptive MMSE estimator under signal presence uncertainty and Laplacian prior.* Paper presented at the Telecommunications (IST), 2012 Sixth International Symposium on.

Rashidi-nejad, M., Abutalebi, H. R., & Tadaion, A. A. (2010). *Speech enhancement using an Improved MMSE estimator with Laplacian prior.* Paper presented at the 5th International Symposium on Telecommunications (IST).

Rasoolzadeh, S., & Abdipour, R. (2010). *Low power implementation of a wavelet based noise reduction system.* Paper presented at the 2nd International Conference on Signal Processing Systems (ICSPS).

REC, I. (2001). P. 862. *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs.*

Rix, A. W., Beerends, J. G., Hollier, M. P., & Hekstra, A. P. (2001). *Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs.* Paper presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01).

Salahuddin, S., Al Islam, S., Hasan, M. K., & Khan, M. (2002). Soft thresholding for DCT speech enhancement. *Electronics letters, 38*(24), 1605-1607.

Saldanha, J. C., & Shruthi, O. (2016). *Reduction of noise for speech signal enhancement using Spectral Subtraction method.* Paper presented at the Information Science (ICIS), International Conference on.

Sanam, T. F., & Imtiaz, H. (2013). *A DCT-based noisy speech enhancement method using Teager energy operator.* Paper presented at the Knowledge and Smart Technology (KST), 2013 5th International Conference on.

Saruwatari, H. (2015). *Statistical-model-based speech enhancement with musical-noise-free properties.* Paper presented at the IEEE International Conference on Digital Signal Processing (DSP).

Selvakumari, N. S., & Radha, V. (2017). *A voice activity detector using SVM and Naïve Bayes classification algorithm.* Paper presented at the Signal Processing and Communication (ICSPC), 2017 International Conference on.

Shehla, R., & Khan, A. A. (2013). Bayesian analysis of gamma model with Laplace approximation. *International Journal of Innovative Research in Science, Engineering and Technology, 2*(11).

Shekokar, S., & Mali, M. (2013). A brief survey of a DCT-based speech enhancement system. *Int. J. Scientif. Eng. Res, 4*, 1-3.

Soon, I., & Koh, S. (2000). Low distortion speech enhancement. *IEE Proceedings-Vision, Image and Signal Processing, 147*(3), 247-253.

Soon, Y. (2003). *Transform based speech enhancement techniques.* PhD Thesis, Nanyang Technological University.

Soon, Y., & Koh, S. N. (2000). Low distortion speech enhancement. *IEE Proceedings-Vision, Image and Signal Processing, 147*(3), 247-253.

Soon, Y., Koh, S. N., & Yeo, C. K. (1998). Noisy speech enhancement using discrete cosine transform. *Speech Communication, 24*(3), 249-257.

Surendran, S., & Kumar, T. K. (2015). Perceptual Subspace Speech Enhancement with Variance Normalization. *Procedia Computer Science, 54*, 818-828.

Taal, C. H., Hendriks, R. C., Heusdens, R., & Jensen, J. (2011). An algorithm for intelligibility prediction of time–frequency weighted noisy speech. *IEEE Transactions on Audio, Speech, and Language Processing, 19*(7), 2125-2136.

Tanenbaum, A. S. (2003). Computer networks, 4-th edition. *ed: Prentice Hall*.

Thiemann, J. (2001). *Acoustic noise suppression for speech signals using auditory masking effects.* McGill University Montreal, Canada.

Tsao, Y., & Lai, Y.-H. (2016). Generalized maximum a posteriori spectral amplitude estimation for speech enhancement. *Speech Communication, 76*, 112-126.

Upadhyay, N., & Karmakar, A. (2015). Speech Enhancement using Spectral Subtraction-type Algorithms: A Comparison and Simulation Study. *Procedia Computer Science, 54*, 574-584.

Varga, A., & Steeneken, H. J. (1993). Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems. *Speech Communication, 12*(3), 247-251.

Vary, P., & Martin, R. (2006). *Digital speech transmission: Enhancement, coding and error concealment*: John Wiley & Sons.

Viswanathan, M. (2013). Simulation of digital communication systems using Matlab. *Mathuranathan Viswanathan at Smashwords*.

Wang, W., & Zhao, J. (2016). Hiding depth information in compressed 2D image/video using reversible watermarking. *Multimedia Tools and Applications, 75*(8), 4285-4303.

Wang, Y., & Brookes, M. (2018). Model-based speech enhancement in the modulation domain. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 26*(3), 580-594.

Wang, Y., Vilermo, M., & Yaroslavsky, L. (2000). *Energy compaction property of the MDCT in comparison with other transforms.* Paper presented at the Audio Engineering Society Convention 109.

Wei, J., Ou, S., Shen, S., & Gao, Y. (2016). *Laplacian-Gaussian mixture based dual-gain wiener filter for speech enhancement.* Paper presented at the 2016 IEEE International Conference on Signal and Image Processing (ICSIP).

Wilk-Kolodziejczyk, D., Regulski, K., & Gumienny, G. (2016). Comparative analysis of the properties of the nodular cast iron with carbides and the austempered ductile iron with use of the machine learning and the support vector machine. *The International Journal of Advanced Manufacturing Technology, 87*(1-4), 1077-1093.

Xia, B., & Bao, C. (2014). Wiener filtering based speech enhancement with weighted denoising auto-encoder and noise classification. *Speech Communication, 60*, 13-29.

Yann, S. I. (2003). *Transform based speech enhancement techniques.* Nanyang Technological University.

Yap, P.-T., Paramesran, R., & Ong, S.-H. (2003). Image analysis by Krawtchouk moments. *IEEE Transactions on Image Processing, 12*(11), 1367-1377.

Yap, P.-T., Paramesran, R., & Ong, S.-H. (2007). Image analysis using Hahn moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 29*(11), 2057-2062.

Yousheng, X., & Jianwen, H. (2014). *Speech enhancement based on combination of wiener filter and subspace filter.* Paper presented at the International Conference on Audio, Language and Image Processing (ICALIP).

Yuan, W., & Xia, B. (2015). A speech enhancement approach based on noise classification. *Applied Acoustics, 96*, 11-19.

Zhang, W., Benesty, J., & Chen, J. (2016). Single-channel noise reduction via semi-orthogonal transformations and reduced-rank filtering. *Speech Communication, 78*, 73-83.

Zhang, Y., & Wu, L. (2012). Classification of fruits using computer vision and a multiclass support vector machine. *sensors, 12*(9), 12489-12505.

Zhao, D. Y., & Kleijn, W. B. (2006). *HMM-Based Speech Enhancement using Explicit Gain Modeling.* Paper presented at the 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings.

Zhao, G., Zhou, B., Zhang, X., & Lu-ying, S. (2012). *A new speech enhancement algorithm with generalized Gamma speech model.* Paper presented at the International Conference on Wireless Communications & Signal Processing (WCSP).

Zhu, H., Shu, H., Zhou, J., Luo, L., & Coatrieux, J.-L. (2007). Image analysis by discrete orthogonal dual Hahn moments. *Pattern Recognition Letters, 28*(13), 1688-1704.

Zou, X., & Zhang, X. (2007). Speech enhancement using an MMSE short time DCT coefficients estimator with supergaussian speech modeling. *Journal of Electronics (China), 24*(3), 332-337.

Zoulikha, M., & Djendi, M. (2016). A new regularized forward blind source separation algorithm for automatic speech quality enhancement. *Applied Acoustics, 112*, 192-200.