

# **Robust Linear Prediction Analysis for Low Bit-Rate Speech Coding**

Nanda Prasetyo Koestoer B. Eng (Hon) (1998)

School of Microelectronic Engineering  
Faculty of Engineering and Information Technology  
Griffith University  
Brisbane, Australia

This dissertation is submitted in fulfilment of the requirements of the degree of  
Doctor of Philosophy

November 2002

## Abstract

Speech coding is a very important area of research in digital signal processing. It is a fundamental element of digital communications and has progressed at a fast pace in parallel to the increase of demands in telecommunication services and capabilities.

Most of the speech coders reported in the literature are based on linear prediction (LP) analysis. Code Excited Linear Predictive (CELP) coder is a typical and popular example of this class of coders. This coder performs LP analysis of speech for extracting LP coefficients and employs an analysis-by-synthesis procedure to search a stochastic codebook to compute the excitation signal. The method used for performing LP analysis plays an important role in the design of a CELP coder. The autocorrelation method is conventionally used for LP analysis. Though this works reasonably well for noise-free (clean) speech, its performance goes down when signal is corrupted by noise.

Spectral analysis of speech signals in noisy environments is an aspect of speech coding that deserves more attention. This dissertation studies the application of recently proposed robust LP analysis methods for estimating the power spectrum envelope of speech signals. These methods are the moving average, moving maximum and average threshold methods. The proposed methods will be compared to the more commonly used methods of LP analysis, such as the conventional autocorrelation method and the Spectral Envelope Estimation Vocoder (SEEVOC) method.

The Linear Predictive Coding (LPC) spectrum calculated from these proposed methods are shown to be more robust. These methods work as well as the conventional methods when the speech signal is clean or has high signal-to-noise ratio.

Also, these robust methods give less quantisation distortion than the conventional methods. The application of these robust methods for speech compression using the CELP coder provides better speech quality when compared to the conventional LP analysis methods.

## **Acknowledgments**

Firstly I wish to express my deepest gratitude to my supervisor Prof. Kuldip Paliwal for all the support and guidance he has offered me. I am very grateful for the knowledge and inspirational wisdom he has shared with me during the course of my study.

I am also thankful to the support I have received from the School of Microelectronic Engineering at Griffith University. The technical support and facilities have been essential in providing a great academic environment for me to complete my study. Specifically I would like to thank everyone at the Signal Processing Laboratory, with which I have had the honour of being associated with. The suggestions, discussions, valuable advice and support provided by the people associated with the laboratory, including visiting researchers, has been crucial during the progression of this work. Special mention goes to Brett Wildermoth, whose assistance in using the laboratory facilities was very beneficial to my research.

Very special thanks go to my closest friend, Shelley Kemp, whose support has been ever-present during my times of need. She is very dear to me and has been responsible for the best times of my life. Finally, I would like to thank my family for everything they have given me during this time. I will forever be grateful to experience their love and support.

## **Statement of Originality**

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

---

Nanda Prasetyo Koestoer

November 2002

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Speech Coding . . . . .	1
1.2	Research Objective . . . . .	3
1.3	Thesis Organisation . . . . .	5
<b>2</b>	<b>Speech Coding and LP Analysis</b>	<b>6</b>
2.1	Speech Production . . . . .	6
2.2	Speech Signal . . . . .	7
2.2.1	Time Domain Representation . . . . .	7
2.2.2	Frequency Domain Representation . . . . .	8
2.3	Properties of Speech . . . . .	10
2.4	Digital Encoding of Speech Signals . . . . .	12
2.4.1	Sampling . . . . .	12
2.4.2	Quantisation . . . . .	13
2.5	Overview of Speech Coding Methods . . . . .	20
2.5.1	Introduction . . . . .	20
2.5.2	LPC-10 . . . . .	21
2.5.3	Multipulse LPC . . . . .	22
2.5.4	CELP . . . . .	26
2.6	LP Analysis . . . . .	26
2.6.1	Background Theory . . . . .	26
2.6.2	Conventional LP Analysis Methods . . . . .	37
2.6.3	Robust Spectral Analysis . . . . .	39

2.6.4	Determination of the LP Parameters . . . . .	45
2.7	Code Excited Linear Prediction Coder . . . . .	47
2.7.1	Background Theory . . . . .	47
2.7.2	Quantisation of Pitch Parameters . . . . .	49
2.7.3	Quantisation of Gain Parameters . . . . .	52
2.7.4	Quantisation of LP Parameters . . . . .	55
2.8	Performance Evaluation Criteria . . . . .	63
2.8.1	Spectral Distortion Measure . . . . .	63
2.8.2	Quantisation of the LP Parameters . . . . .	63
2.8.3	Performance of the CELP Coder . . . . .	64
<b>3</b>	<b>Robust LP Analysis Methods</b>	<b>66</b>
3.1	Introduction . . . . .	66
3.2	Moving Average Method . . . . .	67
3.3	Moving Maximum Method . . . . .	69
3.4	Average Threshold Method . . . . .	71
3.5	Robustness and Accuracy Analysis . . . . .	73
3.5.1	Database . . . . .	73
3.5.2	Procedure . . . . .	74
3.5.3	Results . . . . .	77
<b>4</b>	<b>Quantisation of the LP Parameters</b>	<b>85</b>
4.1	Scalar Quantisation of LP Parameters . . . . .	85
4.2	Split Vector Quantisation of LP Parameters . . . . .	88
<b>5</b>	<b>Low Bit-Rate Speech Coding Application</b>	<b>96</b>
5.1	Application of the Robust LP Analysis Methods in CELP . . . . .	96
5.2	Noise Introduction . . . . .	99
5.2.1	Real World Noise . . . . .	99
5.2.2	Gaussian Noise . . . . .	106
5.3	Variation of the Analysis Window Lengths . . . . .	109
<b>6</b>	<b>Conclusions</b>	<b>117</b>

6.1	Summary . . . . .	117
6.1.1	Observations on Robustness and Accuracy . . . . .	118
6.1.2	Quantisation Performance . . . . .	119
6.1.3	Low Bit-rate Speech Coding Application . . . . .	120
6.2	Future Work . . . . .	121
	<b>Bibliography</b>	<b>123</b>

# List of Figures

1.1	Power spectrum of clean and noise-corrupted speech. . . . .	4
2.1	Basic speech production model. . . . .	7
2.2	Speech signal [ <i>she</i> ] in time domain. . . . .	8
2.3	Speech signal in frequency domain. . . . .	9
2.4	Power spectrum of speech segment [ <i>e</i> ] over 30 ms time frame. . . . .	11
2.5	Basic speech synthesis model of the LPC-10 method. . . . .	22
2.6	Block diagram of the multipulse coder. . . . .	23
2.7	Speech processing model in LP analysis. . . . .	27
2.8	Open-loop AR model. . . . .	28
2.9	Methodology of the search process in SEEVOC. . . . .	41
2.10	SEEVOC power spectrum after allocation of peaks. . . . .	42
2.11	SEEVOC spectral envelope after linear prediction. . . . .	43
2.12	The effect of CP selection with TP=6.8 frequency samples. . . . .	44
2.13	Block diagram of the CELP coder. . . . .	48
2.14	Block diagram of the long term prediction analysis. . . . .	50
2.15	Basic block diagram of the codebook computation procedure. . . . .	52
2.16	SNR performance for different codebook dimension. . . . .	54
3.1	MA spectral envelope. . . . .	68
3.2	MA spectrum after LP analysis. . . . .	68
3.3	MM spectral envelope. . . . .	69
3.4	MM spectrum after LP analysis. . . . .	70
3.5	AT spectral envelope. . . . .	71

3.6	AT spectrum after LP analysis. . . . .	72
3.7	Methodology to simulate robustness. . . . .	74
3.8	Methodology to simulate accuracy for the proposed methods. . . . .	75
3.9	Excitation process to construct synthetic signal. . . . .	76
3.10	Robustness analysis of MA method. . . . .	77
3.11	Accuracy analysis of MA method. . . . .	78
3.12	Robustness analysis of MM method. . . . .	78
3.13	Accuracy analysis of MM method. . . . .	79
3.14	Robustness analysis of AT method. . . . .	80
3.15	Accuracy analysis of AT method. . . . .	80
3.16	Performance of AT for different window lengths. . . . .	81
3.17	Robustness performance of AT for different repetitions. . . . .	82
3.18	Accuracy performance of AT for different repetitions. . . . .	82
3.19	Robustness analysis of speech with added 'restaurant' noise. . . . .	83
3.20	Robustness analysis of speech with added Gaussian noise. . . . .	83
3.21	Accuracy analysis of speech [e] with added Gaussian noise (please refer to Table 3.1). . . . .	84
4.1	Block diagram of the split VQ for 2 partitions. . . . .	90
4.2	Average SD for VQ with no partition. . . . .	93

# List of Tables

2.1	SD performance of mid-level uniform SQ on LP parameters. . . . .	58
2.2	Mid-level uniform SQ on PARCOR coefficients. . . . .	59
2.3	Non-uniform SQ on PARCOR coefficients. . . . .	59
2.4	Non-uniform SQ on ASRC coefficients. . . . .	60
2.5	Non-uniform SQ on LAR coefficients. . . . .	61
2.6	Non-uniform SQ on LSF coefficients. . . . .	62
3.1	Level of noise with respect to SNR. . . . .	81
4.1	Quantisation of LP parameters using AM method. . . . .	87
4.2	Quantisation of LP parameters using the proposed methods. . . . .	87
4.3	Performance of non-uniform SQ using LSF transformation. . . . .	88
4.4	Comparison for quantisation with different VQ selection criterion. . . . .	91
4.5	Quantisation performance for 2 part split VQ. . . . .	92
4.6	Quantisation performance for 3 part split VQ. . . . .	92
4.7	Quantisation performance for 5 part split VQ. . . . .	93
4.8	Performance of the conventional LP analysis methods. . . . .	94
4.9	Performance of the robust LP analysis methods. . . . .	94
5.1	CELP performance for the different LP analysis methods. . . . .	98
5.2	CELP performance for 3 part split VQ on Set 0 sentences. . . . .	98
5.3	CELP performance for 3 part split VQ on Set 1 sentences. . . . .	99
5.4	Performance for 3 part split VQ with babble noise (Set 0). . . . .	100
5.5	Performance for 3 part split VQ with babble noise (Set 1). . . . .	101
5.6	Performance for 5 and 2 part split VQ with babble noise. . . . .	103

5.7	Real world noise on Set 0 at 27 bits/frame. . . . .	104
5.8	Real world noise on Set 1 at 27 bits/frame. . . . .	105
5.9	Performance at 18 bits/frame with Gaussian noise on Set 0. . . . .	106
5.10	Performance at 18 bits/frame with Gaussian noise on Set 1. . . . .	106
5.11	Performance for 3 part split VQ on Set 0 (Gaussian noise). . . . .	107
5.12	Performance for 3 part split VQ on Set 1 (Gaussian noise). . . . .	108
5.13	Performance for 2 part split VQ on Set 0 (Gaussian noise). . . . .	108
5.14	Performance for 5 part split VQ on Set 0 (Gaussian noise). . . . .	109
5.15	Comparison of window lengths for 18 bits/frame on Set 0. . . . .	109
5.16	Comparison of window lengths for 18 bits/frame on Set 1. . . . .	110
5.17	Comparison of window lengths for 21 bits/frame on Set 0. . . . .	110
5.18	Comparison of window lengths for 21 bits/frame on Set 1. . . . .	111
5.19	Comparison of window lengths for 30 bits/frame on Set 0. . . . .	111
5.20	Comparison of window lengths for 30 bits/frame on Set 1. . . . .	112
5.21	Performance at 18 bits/frame with babble noise. . . . .	113
5.22	Performance at 18 bits/frame with street noise. . . . .	114
5.23	Performance at 21 bits/frame with babble noise. . . . .	115
5.24	Performance at 21 bits/frame with street noise. . . . .	116