Robust Linear Prediction Analysis for Low Bit-Rate Speech Coding

Nanda Prasetiyo 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 Prasetiyo Koestoer
November 2002

Contents

List of Figures

List of Tables