Signals and Communication Technology

For further volumes:
http://www.springer.com/series/4748

Sean A. Fulop

Speech Spectrum Analysis

Springer

Dr. Sean A. Fulop
Department of Linguistics
California State University Fresno
N. Backer Ave. 5245
Fresno CA 93740-8001
USA
e-mail: sfulop@csufresno.edu

ISSN 1860-4862

ISBN 978-3-642-17477-3 e-ISBN 978-3-642-17478-0

DOI 10.1007/978-3-642-17478-0

Springer Heidelberg Dordrecht London New York

Cover design: eStudio Calamar, Berlin/Figueres

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

For Billy, who helped me write this.

Preface

The analysis and measurement of the spectrum of a speech signal is one of the most important areas of sound signal processing for a number of fields, yet it is not an area to which a book has been specifically devoted. The accurate determination of the speech spectrum is commonly pursued in diverse areas including speech processing, recognition, and acoustic phonetics. With this book I hope to make the subject of spectrum analysis understandable to a wide audience, which I imagine could include those with a solid background in general signal processing (but not necessarily in speech), and also speech scientists and students with some acoustic phonetics experience who have limited knowledge of signal processing. In keeping with these goals, this is not a book that replaces or attempts to cover the material found in a general signal processing textbook. Some essential signal processing concepts are presented in Chap. 2, but even there the concepts are presented in a generally understandable fashion as far as is possible. Throughout the book, the focus will be on applications to speech analysis and the measurement of important descriptive speech parameters. No attention is paid to parametrizing speech purely for coding or decorrelation for further processing. Mathematical theory will be provided for completeness, but many of these developments are set off in boxes for the benefit of those readers with sufficient background. Other readers may proceed through the main text, where the key results and applications will be presented in plain language as far as possible, and illustrated with software routines and practical "show-and-tell" discussions of the results.

At some points, the book refers to and uses the implementations in the Praat speech analysis software package, which has the advantages that it is used by many scientists around the world, and it is free and open source software, obtainable on the internet from the Praat homepage. At other points, special software routines have been developed and made available to complement the book, and these are provided in the Matlab programming language. If the reader has the basic Matlab package, he/she will be able to immediately implement most of the programs in that platform—only Chap. 7 requires the extra Signal Processing toolbox. A few other freely available toolboxes are also needed, and all the Matlab code is made available for download at the Springer website for additional materials.

And finally, as was written by Lord Kelvin and Professor Tait in their *Treatise on Natural Philosophy* (1912), "I confidently hope that few erratums of serious note will now be found in the work."

Fresno, October 2010 Sean Fulop

Acknowledgments

I should definitely thank Kelly Fitz for collaborating with me on developing reassignment algorithms; Doug Nelson for sharing some of his knowledge (and code) on that subject; Paul Boersma for always answering Praat-related questions; François Auger for telling me a few things about the Time–Frequency Toolbox; and Sandy Disner for helping to perfect my reassigned spectrograms. I should also thank Steven Lulich and Stefanie Shattuck–Hufnagel for trying to figure out what a voice bar is. The writing of this book was carried out over a two-year period which included summer visits to the Department of Linguistics, University of Calgary in 2009 and 2010; thanks to John Archibald, Betsy Ritter, and the other members of the department who tolerate my perennial intrusion (and get me a library card). The book was finished while I had a sabbatical at Fresno State, but was never a funded project. Thanks to Christoph Baumann at Springer for facilitating its publication.

Contents

1 Introduction 1
References 4

2 Phonetics and Signal Processing 5
2.1 Essentials of Phonetics 5
2.1.1 Speech Production Fundamentals 5
2.1.2 Syllables and Speech Sounds 7
2.1.3 Vowels and Consonants 8
2.1.4 Uses of Vocal Pitch 12
2.2 Essentials of Digital Signal Processing 12
2.2.1 Periodic and Aperiodic Signals 13
2.2.2 Sampling of Analog Signals 15
2.2.3 Autocorrelation 15
2.2.4 Fourier's Series and Transform Spectra 18
2.2.5 Practical Computing of Fourier Spectra 27
2.2.6 Filters 34
2.2.7 Analytic Signals 35
2.2.8 Concepts of Frequency 37
References 39

3 History of Speech Spectrum Analysis 41
3.1 Fourier Analysis of Speech 41
3.1.1 Early History 41
3.1.2 The Physical Reality of Fourier Components 43
3.1.3 Recording Sound Signals 45
3.1.4 Early Methods of Fourier Analysis 49
3.2 History of Speech Spectra 52
3.2.1 Vowels and Formants: Early Years 52
3.2.2 Vowel Spectra: 1915–1960 57
3.2.3 Spectrographic Analysis 63

3.3 Parametric Spectral Analysis . . . 65
References . . . 66

4 The Fourier Power Spectrum and Spectrogram . . . 69
4.1 The Power Spectrum in Speech Analysis . . . 69
4.1.1 Vowel Spectra . . . 70
4.1.2 Obstruent Spectra and Averaging Techniques . . . 75
4.1.3 Phonation Types . . . 78
4.2 Principles of the Spectrogram . . . 80
4.2.1 Definitions of the Spectrogram . . . 80
4.2.2 Development of Spectrogram Theory . . . 87
4.2.3 Uncertainty Principle . . . 87
4.3 Spectrographic Analysis of Speech . . . 88
4.3.1 General Guidelines . . . 89
4.3.2 Short Window (Wideband) Analysis . . . 92
4.3.3 Long Window (Narrowband) Analysis . . . 99
4.4 Appendix: Praat and Matlab Techniques . . . 102
4.4.1 Praat Functions . . . 102
4.4.2 Matlab Code . . . 104
References . . . 105

5 Alternative Time–Frequency Representations . . . 107
5.1 Wigner–Ville Distribution . . . 108
5.1.1 Definition and Theory . . . 108
5.1.2 Discrete Implementation . . . 109
5.1.3 Features of the Wigner–Ville Distribution . . . 111
5.2 Zhao-Atlas-Marks Distribution . . . 113
5.2.1 Quadratic Distributions . . . 114
5.2.2 Discrete Implementation . . . 115
5.2.3 Speech Analysis with ZAM . . . 118
5.3 Appendix: Matlab Routines . . . 122
References . . . 124

6 The Reassigned Spectrogram . . . 127
6.1 Reassignment: History and Definitions . . . 128
6.2 Reassigning the Spectrogram . . . 131
6.2.1 Nelson's Algorithm . . . 131
6.2.2 Reassigned Power Spectrum . . . 135
6.3 Pruning the Reassigned Spectrogram . . . 136
6.3.1 General Definitions . . . 136
6.3.2 Cross-Spectral Method . . . 138
6.3.3 Justifying the Interpretation of the Phase Derivative . . . 139
6.3.4 Separation of Formants from Glottal Impulses . . . 139
6.4 Analyzing Phonation . . . 140

6.4.1 Beyond Source-Filter Theory 140
6.4.2 Observations on Phonation Types 145
6.4.3 Phonation as a Biometric 149
6.5 Dynamics of Formants 149
6.6 Nasals and Nasalization 156
6.7 Stops and Fricatives 158
6.8 Pitch Tracking 160
6.9 Appendix: Matlab Techniques 162
References 164

7 Linear Prediction and ARMA Spectrum Estimation 167
7.1 Preliminaries 168
7.1.1 Linear Filters and the z-Transform 169
7.1.2 Source-Filter Model of Speech 171
7.2 Speech Spectra from Linear Prediction 173
7.2.1 Computing the LP Coefficients 173
7.2.2 Computing the Gain and Power Spectrum 174
7.3 Interpretation as Filter Spectrum 175
7.3.1 Poles and Resonances 176
7.3.2 Picking Peaks Versus Solving Roots 179
7.4 Practical Spectrum Analysis and Formant Extraction 181
7.4.1 Linear Prediction Accuracy Studies 181
7.4.2 Analysis Windows 182
7.4.3 Filter Order and Pre-Emphasis 184
7.4.4 Pitch-Asynchronous Versus Closed Phase 186
7.5 Application to Real Speech 188
7.6 Autoregressive Moving Average Modeling 192
7.6.1 A Little ARMA Theory 192
7.6.2 ARMA Computation 193
7.6.3 Applications to Speech 194
7.7 Appendix: Praat and Matlab Techniques 197
7.7.1 Praat Functions 197
7.7.2 Matlab Functions 198
References 200

Index 203