

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 1.1 | Linear and Nonlinear Techniques in Speech Processing | 1 |
| 1.2 | Applications of Nonlinear Speech Processing | 2 |
| 1.3 | Outline of the Book | 6 |
| 1.4 | Summary | 7 |
| | References | 7 |
| 2 | Nonlinearity Framework in Speech Processing | 11 |
| 2.1 | Introduction | 11 |
| 2.2 | Nonlinear Techniques in Speech Processing | 11 |
| 2.3 | Speech Production Mechanism | 12 |
| 2.4 | Speech Perception Mechanism | 14 |
| 2.5 | Conventional Speech Synthesis Approaches | 16 |
| 2.6 | Nonlinearity in Speech Production | 17 |
| 2.6.1 | Vocal Fold Oscillation | 18 |
| 2.6.2 | The Turbulent Sound Source | 20 |
| 2.6.3 | Interaction Phenomenon | 21 |
| 2.7 | Common Signals of Interest | 21 |
| 2.7.1 | AM Signals | 21 |
| 2.7.2 | FM Signals | 22 |
| 2.7.3 | AM-FM Signals | 23 |
| 2.7.4 | Discrete Versions | 23 |
| 2.8 | Summary | 23 |
| | References | 24 |
| 3 | Linear and Dynamic System Model | 27 |
| 3.1 | Introduction | 27 |
| 3.2 | Linear Model | 27 |
| 3.3 | The Linear Source-Filter Model | 30 |
| 3.3.1 | Linear Speech Production Model | 30 |

| | | |
|----------|---|-----------|
| 3.3.2 | The Vocal Tract Transfer Function | 30 |
| 3.3.3 | Lossless Tube Modeling Assumptions | 32 |
| 3.3.4 | Representations Computed from LPC | 33 |
| 3.3.5 | LP Based Cepstrum. | 33 |
| 3.4 | Time-Varying Linear Model | 33 |
| 3.5 | Dynamic System Model | 35 |
| 3.6 | Time-Varying Dynamic System Model | 36 |
| 3.7 | Nonlinear Dynamic System Model | 36 |
| 3.8 | Nonlinear AR Model with Additive Noise | 37 |
| 3.8.1 | Multi-Layer Perception | 38 |
| 3.8.2 | Radial Basis Function | 39 |
| 3.8.3 | Truncated Taylor Series Approximation. | 40 |
| 3.8.4 | Quasi-Linear Approximation | 41 |
| 3.8.5 | Piecewise Linear Approximation. | 42 |
| 3.9 | Summary | 42 |
| | References | 43 |
| 4 | Nonlinear Measurement and Modeling Using Teager | |
| | Energy Operator | 45 |
| 4.1 | Introduction | 45 |
| 4.2 | Signal Energy | 45 |
| 4.3 | Teager Energy Operator | 46 |
| 4.3.1 | Continuous and Discrete Form of Teager Energy Operator | 47 |
| 4.4 | Energies of Well-Known Signals. | 47 |
| 4.4.1 | Sinusoidal Signal | 47 |
| 4.4.2 | Exponential Signal | 48 |
| 4.4.3 | AM Signal | 48 |
| 4.4.4 | FM Signal | 48 |
| 4.4.5 | AM-FM Signal | 50 |
| 4.5 | Generalization of Teager Energy Operator | 50 |
| 4.6 | Energy Separation | 52 |
| 4.6.1 | Energy Separation for Continuous-Time Signals. | 52 |
| 4.6.2 | Energy Separation for Discrete-Time Signals | 53 |
| 4.7 | Teager Energy Operator in Noise | 57 |
| 4.7.1 | Noise Suppression Using Teager Energy Operator | 57 |
| 4.8 | Summary | 59 |
| | References | 59 |
| 5 | AM-FM: Modulation and Demodulation Techniques | 61 |
| 5.1 | Introduction | 61 |
| 5.2 | Importance of Phase | 62 |
| 5.3 | AM-FM Model | 63 |
| 5.3.1 | Amplitude Modulation and Demodulation | 64 |
| 5.3.2 | Frequency Modulation and Demodulation | 65 |

| | | |
|----------|---|------------|
| 5.4 | Estimation Using the Teager Energy Operator | 67 |
| 5.5 | Estimation Using the Hilbert Transform | 68 |
| 5.6 | Multiband Filtering and Demodulation | 69 |
| 5.7 | Short-Time Estimates | 70 |
| 5.7.1 | Short-Time Estimate: Frequency | 70 |
| 5.7.2 | Short-Time Estimate: Bandwidth | 72 |
| 5.8 | Summary | 74 |
| | References | 75 |
| 6 | Application to Speaker Recognition | 77 |
| 6.1 | Introduction | 77 |
| 6.2 | Speaker Recognition System | 78 |
| 6.3 | Preprocessing of Speech Signal | 79 |
| 6.3.1 | Pre-Emphasis | 79 |
| 6.3.2 | Framing | 80 |
| 6.3.3 | Windowing | 81 |
| 6.4 | Investigating Importance of Different Frequency Bands for Speaker Identification | 82 |
| 6.4.1 | Experiment 1 | 83 |
| 6.4.2 | Experiment 2 | 83 |
| 6.4.3 | Experiment 3 | 84 |
| 6.4.4 | Experiment 4 | 85 |
| 6.5 | Speaker Identification Using TEO | 86 |
| 6.5.1 | Performance Evaluation for Clean Speech | 88 |
| 6.5.2 | Performance Evaluation for Noisy Speech: Speech Corrupted by Car Engine Noise | 88 |
| 6.5.3 | Performance Evaluation for Noisy Speech: Speech Corrupted by Babble Noise | 89 |
| 6.5.4 | Effect of Feature Vector Dimensions | 89 |
| 6.6 | Speaker Identification Using AM-FM Model Based Features . . . | 90 |
| 6.6.1 | Set-up1: Combining Instantaneous Frequency and Amplitude | 90 |
| 6.6.2 | Set-up 2: Combining Instantaneous Frequency and Bandwidth | 93 |
| 6.6.3 | Combining Instantaneous Frequency, Bandwidth and Post Smoothing | 93 |
| 6.7 | Summary | 96 |
| | References | 96 |
| | Index | 101 |

Advances in Non-Linear Modeling for Speech
Processing

Holambe, R.S.; Deshpande, M.S.

2012, XIII, 102 p. 32 illus., Softcover

ISBN: 978-1-4614-1504-6