

- 1 Introduction
- Discrete-Time Speech Signal Processing
- The Speech Communication Pathway
- Analysis/Synthesis Based on Speech Production and Perception
- Applications
- Outline of Book
- 2 A Discrete-Time Signal Processing Framework
- Discrete-Time Signals
- Discrete-Time Systems
- Discrete-Time Fourier Transform
- Uncertainty Principle
- z-Transform
- LTI Systems in the Frequency Domain
- Properties of LTI Systems
- Time-Varying Systems
- Discrete-Fourier Transform
- Conversion of Continuous Signals and Systems to Discrete Time
- 3 Production and Classification of Speech Sounds
- Anatomy and Physiology of Speech Production
- Spectrographic Analysis of Speech
- Categorization of Speech Sounds
- Prosody: The Melody of Speech
- Speech Perception
- 4 Acoustics of Speech Production
- Physics of Sound
- Uniform Tube Model
- A Discrete-Time Model Based on Tube Concatenation
- Vocal Fold/Vocal Tract Interaction
- 5 Analysis and Synthesis of Pole-Zero Speech Models
- Time-Dependent Processing
- All-Pole Modeling of Deterministic Signals
- Linear Prediction Analysis of Stochastic Speech Sounds
- Criterion of Goodness
- Synthesis Based on All-Pole Modeling
- Pole-Zero Estimation
- Decomposition of the Glottal Flow Derivative
- Appendix 5.A Properties of Stochastic Processes
- Random Processes
- Ensemble Averages
- Stationary Random Process
- Time Averages
- Power Density Spectrum
- Appendix 5.B Derivation of the Lattice Filter in Linear Prediction Analysis
- 6 Homomorphic Signal Processing
- Concept
- Homomorphic Systems for Convolution

- Complex Cepstrum of Speech-Like Sequences
- Spectral Root Homomorphic Filtering
- Short-Time Homomorphic Analysis of Periodic Sequences
- Short-Time Speech Analysis
- Analysis/Synthesis Structures
- Contrasting Linear Prediction and Homomorphic Filtering
- 7 Short-Time Fourier Transform Analysis and Synthesis
- Short-Time Analysis
- Short-Time Synthesis
- Short-Time Fourier Transform Magnitude
- Signal Estimation from the Modified STFT or STFTM
- Time-Scale Modification and Enhancement of Speech
- Appendix 7.A FBS Method with Multiplicative Modification
- 8 Filter-Bank Analysis/Synthesis
- Revisiting the FBS Method
- Phase Vocoder
- Phase Coherence in the Phase Vocoder
- Constant-Q Analysis/Synthesis
- Auditory Modeling
- 9 Sinusoidal Analysis/Synthesis
- Sinusoidal Speech Model
- Estimation of Sinewave Parameters
- Synthesis
- Source/Filter Phase Model
- Additive Deterministic-Stochastic Model
- Appendix 9.A Derivation of the Sinewave Model
- Appendix 9.B Derivation of Optimal Cubic Phase Parameters
- 10 Frequency-Domain Pitch Estimation
- A Correlation-Based Pitch Estimator
- Pitch Estimation Based on a Comb Filter
- Pitch Estimation Based on a Harmonic Sinewave Model
- Glottal Pulse Onset Estimation
- Multi-Band Pitch and Voicing Estimation
- 11 Nonlinear Measurement and Modeling Techniques
- The STFT and Wavelet Transform Revisited
- Bilinear Time-Frequency Distributions
- Aeroacoustic Flow in the Vocal Tract
- Instantaneous Teager Energy Operator
- 12 Speech Coding
- Statistical Models of Speech
- Scaler Quantization
- Vector Quantization (VQ)
- Frequency-Domain Coding
- Model-Based Coding
- LPC Residual Coding
- 13 Speech Enhancement

- Introduction
- Preliminaries
- Wiener Filtering
- Model-Based Processing
- Enhancement Based on Auditory Masking
- Appendix 13.A Stochastic-Theoretic parameter Estimation
- 14 Speaker Recognition
 - Introduction
 - Spectral Features for Speaker Recognition
 - Speaker Recognition Algorithms
 - Non-Spectral Features in Speaker Recognition
 - Signal Enhancement for the Mismatched Condition
 - Speaker Recognition from Coded Speech