



# Table of Contents

## Chapter 1

Introduction .....	1
1.1 Understanding the Enemy: Noise .....	2
1.1.1 Noise Sources .....	2
1.1.2 Noise and Speech Levels in Various Environments .....	5
1.2 Classes of Speech Enhancement Algorithms .....	6
1.3 Book Organization .....	7
References .....	9

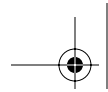
## Part 1

Fundamentals .....	11
--------------------	----

## Chapter 2

Discrete-Time Signal Processing and Short-Time Fourier Analysis .....	13
2.1 Discrete-Time Signals .....	13
2.2 Linear Time-Invariant Discrete-Time Systems .....	16
2.2.1 Difference Equations .....	16
2.2.2 Linear Convolution .....	17
2.3 The z-Transform .....	18
2.3.1 Properties .....	18
2.3.2 The z-Domain Transfer Function .....	19
2.4 Discrete-Time Fourier Transform .....	21
2.4.1 DTFT Properties .....	22
2.4.2 Discrete Fourier Transform .....	24
2.4.3 Windowing .....	27
2.5 Short-Time Fourier Transform .....	32
2.5.1 Definition .....	32
2.5.2 Interpretations of the STFT .....	33
2.5.3 Sampling the STFT in Time and Frequency .....	35
2.5.4 Short-Time Synthesis of Speech .....	36
2.5.4.1 Filterbank Summation for Short-Time Synthesis of Speech .....	37
2.5.4.2 Overlap-and-Add Method for Short-Time Synthesis .....	39
2.6 Spectrographic Analysis of Speech Signals .....	42
2.7 Summary .....	43
References .....	44





**Chapter 3**

Speech Production and Perception .....45

3.1 The Speech Signal.....45

3.2 The Speech Production Process .....45

    3.2.1 Lungs .....46

    3.2.2 Larynx and Vocal Folds.....47

    3.2.3 Vocal Tract.....51

3.3 Engineering Model of Speech Production.....54

3.4 Classes of Speech Sounds.....55

3.5 Acoustic Cues in Speech Perception .....57

    3.5.1 Vowels and Diphthongs.....57

    3.5.2 Semivowels .....60

    3.5.3 Nasals.....61

    3.5.4 Stops .....62

    3.5.4 Fricatives.....64

3.6 Summary.....66

References .....66

**Chapter 4**

Noise Compensation by Human Listeners .....69

4.1 Intelligibility of Speech in Multiple-Talker Conditions .....70

    4.1.1 Effect of Masker’s Spectral/Temporal Characteristics  
        and Number of Talkers: Monaural Hearing.....70

    4.1.2 Binaural Hearing .....73

4.2 Acoustic Properties of Speech Contributing to Robustness.....78

    4.2.1 Shape of the Speech Spectrum .....78

    4.2.2 Spectral Peaks.....80

    4.2.3 Periodicity.....82

    4.2.4 Rapid Spectral Changes Signaling Consonants.....84

4.3 Perceptual Strategies for Listening in Noise .....85

    4.3.1 Auditory Streaming .....85

    4.3.2 Listening in the Gaps and Glimpsing .....86

    4.3.3 Use of F0 Differences .....87

    4.3.4 Use of Linguistic Knowledge .....89

    4.3.5 Use of Spatial and Visual Cues.....89

4.4 Summary.....90

References .....90

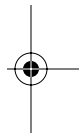
**Part 2**

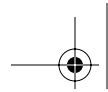
Algorithms.....95

**Chapter 5**

Spectral-Subtractive Algorithms .....97

5.1 Basic Principles of Spectral Subtraction.....97

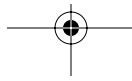
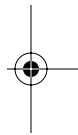


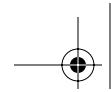


5.2	A Geometric View of Spectral Subtraction .....	101
5.2.1	Upper Bounds on the Difference Between the Noisy and Clean Signals' Phases .....	102
5.2.2	Alternate Spectral-Subtractive Rules and Theoretical Limits .....	104
5.3	Shortcomings of the Spectral Subtraction Method .....	110
5.4	Spectral Subtraction Using Oversubtraction .....	112
5.5	Nonlinear Spectral Subtraction .....	119
5.6	Multiband Spectral Subtraction .....	120
5.7	MMSE Spectral Subtraction Algorithm .....	125
5.8	Extended Spectral Subtraction .....	128
5.9	Spectral Subtraction Using Adaptive Gain Averaging .....	130
5.10	Selective Spectral Subtraction .....	133
5.11	Spectral Subtraction Based on Perceptual Properties .....	135
5.12	Performance of Spectral Subtraction Algorithms .....	136
5.13	Summary .....	138
	References .....	139

**Chapter 6**

	Wiener Filtering .....	143
6.1	Introduction to Wiener Filter Theory .....	143
6.2	Wiener Filters in the Time Domain .....	144
6.3	Wiener Filters in the Frequency Domain .....	146
6.4	Wiener Filters and Linear Prediction .....	148
6.5	Wiener Filters for Noise Reduction .....	150
6.5.1	Square-Root Wiener Filter .....	158
6.5.2	Parametric Wiener Filters .....	158
6.6	Iterative Wiener Filtering .....	163
6.6.1	Speech Production Model .....	164
6.6.2	Statistical Parameter Estimation of the All-Pole Model in Noise .....	165
6.7	Imposing Constraints on Iterative Wiener Filtering .....	172
6.7.1	Across-Time Spectral Constraints .....	172
6.7.2	Across-Iterations Constraints .....	176
6.8	Constrained Iterative Wiener Filtering .....	177
6.9	Constrained Wiener Filtering .....	180
6.9.1	Definitions of Speech and Noise Distortions .....	180
6.9.2	Limiting the Noise Distortion Level .....	184
6.10	Estimating the Wiener Gain Function .....	187
6.10.1	Implementation .....	189
6.11	Incorporating Psychoacoustic Constraints in Wiener Filtering .....	192
6.11.1	Shaping the Noise Distortion in the Frequency Domain .....	192
6.11.2	Using Masking Thresholds as Constraints .....	195

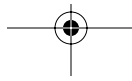


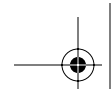


6.12 Codebook-Driven Wiener Filtering.....198  
 6.13 Audible Noise Suppression Algorithm .....202  
 6.14 Summary.....208  
 References .....209

**Chapter 7**

Statistical-Model-Based Methods .....213  
 7.1 Maximum-Likelihood Estimators .....213  
 7.2 Bayesian Estimators .....219  
 7.3 MMSE Estimator.....219  
     7.3.1 MMSE Magnitude Estimator .....222  
     7.3.2 MMSE Complex Exponential Estimator .....227  
     7.3.3 Estimating the *A Priori* SNR.....228  
         7.3.3.1 Maximum-Likelihood Method.....229  
         7.3.3.2 Decision-Directed Approach.....230  
 7.4 Improvements to the Decision-directed Approach .....231  
     7.4.1 Reducing the Bias .....232  
     7.4.2 Improving the Adaptation Speed .....233  
 7.5 Implementation and Evaluation  
     of the MSSE Estimator .....237  
 7.6 Elimination of Musical Noise .....238  
 7.7 Log-MMSE Estimator .....240  
 7.8 MMSE Estimation of the *p*th-Power Spectrum.....242  
 7.9 MMSE Estimators Based on Non-Gaussian Distributions.....247  
 7.10 Maximum *A Posteriori* (MAP) Estimators .....251  
 7.11 General Bayesian Estimators .....254  
 7.12 Perceptually Motivated Bayesian Estimators .....256  
     7.12.1 Psychoacoustically Motivated Distortion Measure .....256  
     7.12.2 Weighted Euclidean Distortion Measure .....257  
     7.12.3 Itakura–Saito Measure.....262  
     7.12.4 Cosh Measure.....263  
     7.12.5 Weighted Likelihood Ratio .....266  
     7.12.6 Modified IS Distortion Measure .....266  
 7.13 Incorporating Speech Absence Probability in Speech  
     Enhancement .....269  
     7.13.1 Incorporating Speech-Presence Uncertainty  
         in Maximum-Likelihood Estimators.....270  
     7.13.2 Incorporating Speech-Presence Uncertainty  
         in MMSE Estimators .....272  
     7.13.3 Incorporating Speech-Presence Uncertainty  
         in Log-MMSE Estimators.....277  
     7.13.4 Implementation Issues Regarding *A Priori*  
         SNR Estimation.....279

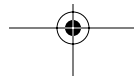
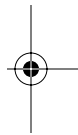


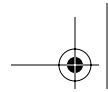


7.14	Methods for Estimating the <i>A Priori</i> Probability of Speech Absence .....	279
7.15	Summary .....	285
	References .....	286

## Chapter 8

	Subspace Algorithms .....	291
8.1	Introduction .....	291
8.1.1	Definitions .....	292
8.1.2	Projections .....	293
8.1.3	Low-Rank Modeling .....	298
8.2	Using SVD for Noise Reduction: Theory .....	300
8.2.1	SVD Analysis of “Noisy” Matrices .....	300
8.2.2	Least-Squares and Minimum-Variance Estimates of the Signal Matrix .....	303
8.3	SVD-Based Algorithms: White Noise .....	306
8.3.1	SVD Synthesis of Speech .....	306
8.3.2	Determining the Effective Rank .....	311
8.3.4	Noise Reduction Algorithm .....	315
8.4	SVD-Based Algorithms: Colored Noise .....	316
8.5	SVD-Based Methods: A Unified View .....	320
8.6	EVD-Based Methods: White Noise .....	320
8.6.1	Eigenvalue Analysis of “Noisy” Matrices .....	320
8.6.2	Subspace Methods Based on Linear Estimators .....	325
8.6.2.1	Linear Minimum Mean-Square Estimator (LMMSE) .....	326
8.6.2.2	Time-Domain-Constrained Estimator .....	328
8.6.2.3	Spectral-Domain-Constrained Estimator .....	332
8.6.3	Implementation .....	338
8.6.3.1	Covariance Estimation .....	338
8.6.3.2	Estimating the Lagrange Multiplier .....	340
8.6.3.3	Estimating the Signal Subspace Dimension .....	342
8.7	EVD-Based Methods: Colored Noise .....	344
8.7.1	Prewhitening Approach .....	345
8.7.2	Signal/Noise KLT-Based Method .....	349
8.7.3	Adaptive KLT Approach .....	352
8.7.4	Subspace Approach with Embedded Prewhitening .....	354
8.7.4.1	Time-Domain-Constrained Estimator .....	354
8.7.4.2	Spectrum-Domain-Constrained Estimator .....	356
8.7.4.3	Implementation .....	359
8.7.4.4	Relationship Between Subspace Estimators and Prewhitening .....	361
8.8	EVD-Based Methods: A Unified View .....	366
8.9	Perceptually Motivated Subspace Algorithms .....	367

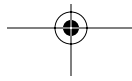


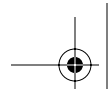


8.9.1	Fourier to Eigen-Domain Relationship.....	368
8.9.2	Incorporating Psychoacoustic Model Constraints .....	372
8.9.3	Incorporating Auditory Masking Constraints .....	374
8.10	Subspace-Tracking Algorithms .....	376
8.10.1	Block Algorithms .....	377
8.10.2	Recursive Algorithms.....	383
8.10.2.1	Modified Eigenvalue Problem Algorithms.....	384
8.10.2.2	Adaptive Algorithms .....	385
8.10.3	Using Subspace-Tracking Algorithms in Speech Enhancement .....	392
8.11	Summary.....	393
	References .....	394

## Chapter 9

	Noise Estimation Algorithms.....	399
9.1	Voice Activity Detection Vs. Noise Estimation.....	399
9.2	Introduction to Noise Estimation Algorithms .....	401
9.3	Minimal-Tracking Algorithms.....	403
9.3.1	Minimum Statistics (MS) Noise Estimation.....	403
9.3.1.1	Principles .....	403
9.3.1.2	Derivation of the Bias Factor.....	405
9.3.1.3	Derivation of Optimal Time- and Frequency-Dependent Smoothing Factor .....	411
9.3.1.4	Searching for the Minimum.....	414
9.3.1.5	Minimum Statistics Algorithm.....	415
9.3.2	Continuous Spectral Minima Tracking .....	417
9.4	Time-Recursive Averaging Algorithms for Noise Estimation.....	420
9.4.1	SNR-Dependent Recursive Averaging .....	421
9.4.2	Weighted Spectral Averaging .....	423
9.4.3	Recursive Averaging Algorithms Based on Signal-Presence Uncertainty .....	429
9.4.3.1	Likelihood Ratio Approach.....	430
9.4.3.2	Minima-Controlled Recursive Averaging (MCRA) Algorithms.....	434
9.5	Histogram-Based Techniques .....	446
9.6	Other Noise Estimation Algorithms.....	453
9.7	Objective Comparison of Noise Estimation Algorithms .....	455
9.8	Summary.....	459
	References .....	460





**Part 3**  
**Evaluation .....463**

**Chapter 10**

Evaluating Performance of Speech Enhancement Algorithms .....465

10.1 Quality vs. Intelligibility.....465

10.2 Evaluating Intelligibility of Processed Speech .....466

    10.2.1 Nonsense Syllable Tests .....467

    10.2.2 Word Tests .....472

        10.2.2.1 Phonetically Balanced Word Tests.....472

        10.2.2.2 Rhyming Word Tests .....473

    10.2.3 Sentence Tests .....476

    10.2.4 Measuring Speech Intelligibility .....478

        10.2.4.1 Speech Reception Threshold .....478

        10.2.4.2 Using Statistical Tests to Assess Significant  
                Differences: Recommended Practice.....480

10.3 Evaluating Quality of Processed Speech.....486

    10.3.1 Relative Preference Methods .....486

    10.3.2 Absolute Category Rating Methods.....489

        10.3.2.1 Mean Opinion Scores .....490

        10.3.2.2 Diagnostic Acceptability Measure .....492

        10.3.2.3 The ITU-T P.835 Standard .....495

10.4 Evaluating Reliability of Quality Judgments:  
Recommended Practice.....498

    10.4.1 Intrarater Reliability Measures.....498

    10.4.2 Interrater Reliability Measures.....500

10.5 Objective Quality Measures .....502

    10.5.1 Segmental SNR Measures.....503

    10.5.2 Spectral Distance Measures Based on LPC .....506

    10.5.3 Perceptually Motivated Measures .....507

        10.5.3.1 Weighted Spectral Slope (WSS)Distance  
                Measure .....508

        10.5.3.2 Bark Distortion Measures .....509

        10.5.3.3 Perceptual Evaluation of Speech  
                Quality (PESQ) Measure .....514

    10.5.4 Composite Measures .....525

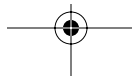
10.6 Nonintrusive Objective Quality Measures.....527

10.7 Figures of Merit of Objective Quality Measures .....528

10.8 Challenges and Future Directions in Objective Quality  
Evaluation .....530

10.9 Summary .....534

References .....535





**Chapter 11**

Comparison of Speech Enhancement Algorithms.....541

11.1 NOIZEUS: A Noisy Speech Corpus for Quality Evaluation of  
Speech Enhancement Algorithms .....542

11.2 Comparison of Enhancement Algorithms: Speech Quality .....543

11.2.1 Quality Evaluation: Procedure .....544

11.2.2 Subjective Quality Evaluation: Results.....545

11.2.3 Within-Class Algorithm Comparisons .....545

11.2.4 Across-Class Algorithm Comparisons .....550

11.2.5 Comparisons in Reference to Noisy Speech .....554

11.2.6 Contribution of Speech and Noise Distortion  
to Judgment of Overall Quality .....558

11.2.7 Summary of Findings.....559

11.3 Comparison of Enhancement Algorithms: Speech Intelligibility .....560

11.3.1 Listening Tests: Procedure .....561

11.3.2 Intelligibility Evaluation: Results .....562

11.3.3 Intelligibility Comparison Among Algorithms .....564

11.3.4 Intelligibility Comparison Against Noisy Speech .....564

11.3.5 Summary of Findings.....567

11.4 Comparison of Objective Measures for Quality Evaluation.....568

11.4.1 Objective Measures .....568

11.4.2 Correlations of Objective Measures with Quality .....572

11.4.3 Summary of Findings.....576

11.5 Summary .....579

References .....579

**Appendix A: Derivation of the MMSE Estimator .....583**

**Appendix B: Special Functions and Integrals .....587**

**Appendix C: Speech Databases and MATLAB Code.....591**

**Index .....603**

