

# Contents

## Part I Basic Theory

<b>1</b>	<b>Introduction</b>	3
1.1	Definition and History	3
1.2	Speaker Recognition Branches	5
1.2.1	Speaker Verification (Speaker Authentication)	5
1.2.2	Speaker Identification (Closed-Set and Open-Set)	7
1.2.3	Speaker and Event Classification	8
1.2.4	Speaker Segmentation	9
1.2.5	Speaker Detection	11
1.2.6	Speaker Tracking	11
1.3	Speaker Recognition Modalities	12
1.3.1	Text-Dependent Speaker Recognition	12
1.3.2	Text-Independent Speaker Recognition	13
1.3.3	Text-Prompted Speaker Recognition	14
1.3.4	Knowledge-Based Speaker Recognition	15
1.4	Applications	16
1.4.1	Financial Applications	16
1.4.2	Forensic and Legal Applications	18
1.4.3	Access Control (Security) Applications	19
1.4.4	Audio and Video Indexing (Diarization) Applications	19
1.4.5	Surveillance Applications	20
1.4.6	Teleconferencing Applications	21
1.4.7	Proctorless Oral Testing	21
1.4.8	Other Applications	23
1.5	Comparison to Other Biometrics	23
1.5.1	Deoxyribonucleic Acid (DNA)	24
1.5.2	Ear	25
1.5.3	Face	27
1.5.4	Fingerprint and Palm	28
1.5.5	Hand and Finger Geometry	30

1.5.6	Iris . . . . .	30
1.5.7	Retina . . . . .	31
1.5.8	Thermography . . . . .	32
1.5.9	Vein . . . . .	32
1.5.10	Gait . . . . .	33
1.5.11	Handwriting . . . . .	34
1.5.12	Keystroke . . . . .	35
1.5.13	Multimodal . . . . .	35
1.5.14	Summary of Speaker Biometric Characteristics . . . . .	37
	References . . . . .	38
<b>2</b>	<b>The Anatomy of Speech . . . . .</b>	<b>43</b>
2.1	The Human Vocal System . . . . .	44
2.1.1	Trachea and Larynx . . . . .	44
2.1.2	Vocal Folds (Vocal Chords) . . . . .	44
2.1.3	Pharynx . . . . .	47
2.1.4	Soft Palate and the Nasal System . . . . .	48
2.1.5	Hard Palate . . . . .	48
2.1.6	Oral Cavity Exit . . . . .	48
2.2	The Human Auditory System . . . . .	48
2.2.1	The Ear . . . . .	50
2.3	The Nervous System and the Brain . . . . .	51
2.3.1	Neurons – Elementary Building Blocks . . . . .	52
2.3.2	The Brain . . . . .	54
2.3.3	Function Localization in the Brain . . . . .	59
2.3.4	Specializations of the Hemispheres of the Brain . . . . .	62
2.3.5	Audio Production . . . . .	64
2.3.6	Auditory Perception . . . . .	66
2.3.7	Speaker Recognition . . . . .	71
	References . . . . .	72
<b>3</b>	<b>Signal Representation of Speech . . . . .</b>	<b>75</b>
3.1	Sampling The Audio . . . . .	77
3.1.1	The Sampling Theorem . . . . .	78
3.1.2	Convergence Criteria for the Sampling Theorem . . . . .	84
3.1.3	Extensions of the Sampling Theorem . . . . .	84
3.2	Quantization and Amplitude Errors . . . . .	85
3.3	The Speech Waveform . . . . .	87
3.4	The Spectrogram . . . . .	87
3.5	Formant Representation . . . . .	89
3.6	Practical Sampling and Associated Errors . . . . .	92
3.6.1	Ideal Sampler . . . . .	98
3.6.2	Aliasing . . . . .	99
3.6.3	Truncation Error . . . . .	102
3.6.4	Jitter . . . . .	103

3.6.5	Loss of Information . . . . .	104
References	. . . . .	105
<b>4</b>	<b>Phonetics and Phonology</b> . . . . .	107
4.1	Phonetics . . . . .	107
4.1.1	Initiation . . . . .	109
4.1.2	Phonation . . . . .	109
4.1.3	Articulation . . . . .	110
4.1.4	Coordination . . . . .	111
4.1.5	Vowels . . . . .	112
4.1.6	Pulmonic Consonants . . . . .	115
4.1.7	Whisper . . . . .	119
4.1.8	Whistle . . . . .	119
4.1.9	Non-Pulmonic Consonants . . . . .	120
4.2	Phonology and Linguistics . . . . .	122
4.2.1	Phonemic Utilization Across Languages . . . . .	122
4.2.2	Whisper . . . . .	125
4.2.3	Importance of Vowels in Speaker Recognition . . . . .	127
4.2.4	Evolution of Languages toward Discriminability . . . . .	129
4.3	Suprasegmental Features of Speech . . . . .	131
4.3.1	Prosodic Features . . . . .	132
4.3.2	Metrical features of Speech . . . . .	138
4.3.3	Temporal features of Speech . . . . .	140
4.3.4	Co-Articulation . . . . .	140
References	. . . . .	141
<b>5</b>	<b>Signal Processing of Speech and Feature Extraction</b> . . . . .	143
5.1	Auditory Perception . . . . .	144
5.1.1	Pitch . . . . .	146
5.1.2	Loudness . . . . .	149
5.1.3	Timbre . . . . .	151
5.2	The Sampling Process . . . . .	152
5.2.1	Anti-Aliasing . . . . .	153
5.2.2	Hi-Pass Filtering . . . . .	153
5.2.3	Pre-Emphasis . . . . .	153
5.2.4	Quantization . . . . .	155
5.3	Spectral Analysis and Direct Method Features . . . . .	157
5.3.1	Framing the Signal . . . . .	160
5.3.2	Windowing . . . . .	162
5.3.3	Discrete Fourier Transform (DFT) and Spectral Estimation . . . . .	167
5.3.4	Frequency Warping . . . . .	169
5.3.5	Magnitude Warping . . . . .	172
5.3.6	Mel Frequency Cepstral Coefficients (MFCC) . . . . .	173
5.3.7	Mel Cepstral Dynamics . . . . .	175
5.4	Linear Predictive Cepstral Coefficients (LPCC) . . . . .	176

5.4.1	Autoregressive (AR) Estimate of the PSD . . . . .	177
5.4.2	LPC Computation . . . . .	184
5.4.3	Partial Correlation (PARCOR) Features . . . . .	185
5.4.4	Log Area Ratio (LAR) Features . . . . .	189
5.4.5	Linear Predictive Cepstral Coefficient (LPCC) Features . . . . .	189
5.5	Perceptual Linear Predictive (PLP) Analysis . . . . .	190
5.5.1	Spectral Analysis . . . . .	191
5.5.2	Bark Frequency Warping . . . . .	191
5.5.3	Equal-Loudness Pre-emphasis . . . . .	192
5.5.4	Magnitude Warping . . . . .	193
5.5.5	Inverse DFT . . . . .	193
5.6	Other Features . . . . .	193
5.6.1	Wavelet Filterbanks . . . . .	194
5.6.2	Instantaneous Frequencies . . . . .	197
5.6.3	Empirical Mode Decomposition (EMD) . . . . .	198
5.7	Signal Enhancement and Pre-Processing . . . . .	199
	References . . . . .	199
<b>6</b>	<b>Probability Theory and Statistics . . . . .</b>	<b>205</b>
6.1	Set Theory . . . . .	205
6.1.1	Equivalence and Partitions . . . . .	208
6.1.2	R-Rough Sets (Rough Sets) . . . . .	210
6.1.3	Fuzzy Sets . . . . .	211
6.2	Measure Theory . . . . .	211
6.2.1	Measure . . . . .	212
6.2.2	Multiple Dimensional Spaces . . . . .	216
6.2.3	Metric Space . . . . .	217
6.2.4	Banach Space (Normed Vector Space) . . . . .	218
6.2.5	Inner Product Space (Dot Product Space) . . . . .	219
6.2.6	Infinite Dimensional Spaces (Pre-Hilbert and Hilbert) . . . . .	219
6.3	Probability Measure . . . . .	221
6.4	Integration . . . . .	227
6.5	Functions . . . . .	228
6.5.1	Probability Density Function . . . . .	229
6.5.2	Densities in the Cartesian Product Space . . . . .	232
6.5.3	Cumulative Distribution Function . . . . .	235
6.5.4	Function Spaces . . . . .	236
6.5.5	Transformations . . . . .	238
6.6	Statistical Moments . . . . .	239
6.6.1	Mean . . . . .	239
6.6.2	Variance . . . . .	242
6.6.3	Skewness (skew) . . . . .	245
6.6.4	Kurtosis . . . . .	246
6.7	Discrete Random Variables . . . . .	247
6.7.1	Combinations of Random Variables . . . . .	250

6.7.2	Convergence of a Sequence	250
6.8	Sufficient Statistics	251
6.9	Moment Estimation	253
6.9.1	Estimating the Mean	253
6.9.2	Law of Large Numbers (LLN)	254
6.9.3	Different Types of Mean	257
6.9.4	Estimating the Variance	258
6.10	Multi-Variate Normal Distribution	259
	References	261
<b>7</b>	<b>Information Theory</b>	<b>265</b>
7.1	Sources	266
7.2	The Relation between Uncertainty and Choice	269
7.3	Discrete Sources	269
7.3.1	Entropy or Uncertainty	270
7.3.2	Generalized Entropy	278
7.3.3	Information	279
7.3.4	The Relation between Information and Entropy	280
7.4	Discrete Channels	282
7.5	Continuous Sources	284
7.5.1	Differential Entropy (Continuous Entropy)	284
7.6	Relative Entropy	286
7.6.1	Mutual Information	291
7.7	Fisher Information	294
	References	299
<b>8</b>	<b>Metrics and Divergences</b>	<b>301</b>
8.1	Distance (Metric)	301
8.1.1	Distance Between Sequences	302
8.1.2	Distance Between Vectors and Sets of Vectors	302
8.1.3	Hellinger Distance	304
8.2	Divergences and Directed Divergences	304
8.2.1	Kullback-Leibler's Directed Divergence	305
8.2.2	Jeffreys' Divergence	305
8.2.3	Bhattacharyya Divergence	306
8.2.4	Matsushita Divergence	307
8.2.5	F-Divergence	308
8.2.6	$\delta$ -Divergence	309
8.2.7	$\chi^\alpha$ Directed Divergence	310
	References	310
<b>9</b>	<b>Decision Theory</b>	<b>313</b>
9.1	Hypothesis Testing	313
9.2	Bayesian Decision Theory	316
9.2.1	Binary Hypothesis	320

9.2.2	Relative Information and Log Likelihood Ratio . . . . .	321
9.3	Bayesian Classifier . . . . .	322
9.3.1	Multi-Dimensional Normal Classification . . . . .	326
9.3.2	Classification of a Sequence . . . . .	328
9.4	Decision Trees . . . . .	331
9.4.1	Tree Construction . . . . .	332
9.4.2	Types of Questions . . . . .	333
9.4.3	Maximum Likelihood Estimation (MLE) . . . . .	336
	References . . . . .	338
<b>10</b>	<b>Parameter Estimation . . . . .</b>	<b>341</b>
10.1	Maximum Likelihood Estimation . . . . .	342
10.2	Maximum A-Posteriori (MAP) Estimation . . . . .	344
10.3	Maximum Entropy Estimation . . . . .	345
10.4	Minimum Relative Entropy Estimation . . . . .	346
10.5	Maximum Mutual Information Estimation (MMIE) . . . . .	348
10.6	Model Selection . . . . .	349
10.6.1	Akaike Information Criterion (AIC) . . . . .	350
10.6.2	Bayesian Information Criterion (BIC) . . . . .	353
	References . . . . .	354
<b>11</b>	<b>Unsupervised Clustering and Learning . . . . .</b>	<b>357</b>
11.1	Vector Quantization (VQ) . . . . .	358
11.2	Basic Clustering Techniques . . . . .	359
11.2.1	Standard k-Means (Lloyd) Algorithm . . . . .	360
11.2.2	Generalized Clustering . . . . .	363
11.2.3	Overpartitioning . . . . .	364
11.2.4	Merging . . . . .	364
11.2.5	Modifications to the k-Means Algorithm . . . . .	365
11.2.6	k-Means Wrappers . . . . .	368
11.2.7	Rough k-Means . . . . .	375
11.2.8	Fuzzy k-Means . . . . .	377
11.2.9	k-Harmonic Means Algorithm . . . . .	378
11.2.10	Hybrid Clustering Algorithms . . . . .	380
11.3	Estimation using Incomplete Data . . . . .	381
11.3.1	Expectation Maximization (EM) . . . . .	381
11.4	Hierarchical Clustering . . . . .	388
11.4.1	Agglomerative (Bottom-Up) Clustering (AHC) . . . . .	389
11.4.2	Divisive (Top-Down) Clustering (DHC) . . . . .	389
11.5	Semi-Supervised Learning . . . . .	390
	References . . . . .	390

<b>12 Transformation</b> .....	393
12.1 Principal Component Analysis (PCA) .....	394
12.1.1 Formulation .....	394
12.2 Generalized Eigenvalue Problem .....	397
12.3 Nonlinear Component Analysis .....	399
12.3.1 Kernel Principal Component Analysis (Kernel PCA) .....	400
12.4 Linear Discriminant Analysis (LDA) .....	401
12.4.1 Integrated Mel Linear Discriminant Analysis (IMELDA) .....	404
12.5 Factor Analysis .....	404
References .....	409
<b>13 Hidden Markov Modeling (HMM)</b> .....	411
13.1 Memoryless Models .....	413
13.2 Discrete Markov Chains .....	415
13.3 Markov Models .....	416
13.4 Hidden Markov Models .....	418
13.5 Model Design and States .....	421
13.6 Training and Decoding .....	423
13.6.1 Trellis Diagram Representation .....	428
13.6.2 Forward Pass Algorithm .....	430
13.6.3 Viterbi Algorithm .....	432
13.6.4 Baum-Welch (Forward-Backward) Algorithm .....	433
13.7 Gaussian Mixture Models (GMM) .....	442
13.7.1 Training .....	444
13.7.2 Tractability of Models .....	449
13.8 Practical Issues .....	451
13.8.1 Smoothing .....	451
13.8.2 Model Comparison .....	453
13.8.3 Held-Out Estimation .....	456
13.8.4 Deleted Estimation .....	461
References .....	462
<b>14 Neural Networks</b> .....	465
14.1 Perceptron .....	466
14.2 Feedforward Networks .....	466
14.2.1 Auto Associative Neural Networks (AANN) .....	469
14.2.2 Radial Basis Function Neural Networks (RBFNN) .....	469
14.2.3 Training (Learning) Formulation .....	470
14.2.4 Optimization Problem .....	473
14.2.5 Global Solution .....	474
14.3 Recurrent Neural Networks (RNN) .....	476
14.4 Time-Delay Neural Networks (TDNNs) .....	477
14.5 Hierarchical Mixtures of Experts (HME) .....	479
14.6 Practical Issues .....	479
References .....	481

<b>15 Support Vector Machines</b> .....	485
15.1 Risk Minimization .....	488
15.1.1 Empirical Risk Minimization .....	492
15.1.2 Capacity and Bounds on Risk .....	493
15.1.3 Structural Risk Minimization .....	493
15.2 The Two-Class Problem .....	494
15.2.1 Dual Representation .....	497
15.2.2 Soft Margin Classification .....	500
15.3 Kernel Mapping .....	503
15.3.1 The Kernel Trick .....	504
15.4 Positive Semi-Definite Kernels .....	506
15.4.1 Linear Kernel .....	506
15.4.2 Polynomial Kernel .....	506
15.4.3 Gaussian Radial Basis Function (GRBF) Kernel .....	507
15.4.4 Cosine Kernel .....	508
15.4.5 Fisher Kernel .....	508
15.4.6 GLDS Kernel .....	509
15.4.7 GMM-UBM Mean Interval (GUMI) Kernel .....	510
15.5 Non Positive Semi-Definite Kernels .....	511
15.5.1 Jeffreys Divergence Kernel .....	511
15.5.2 Fuzzy Hyperbolic Tangent (tanh) Kernel .....	512
15.5.3 Neural Network Kernel .....	513
15.6 Kernel Normalization .....	513
15.7 Kernel Principal Component Analysis (Kernel PCA) .....	514
15.8 Nuisance Attribute Projection (NAP) .....	516
15.9 The multiclass ( $\Gamma$ -Class) Problem .....	518
References .....	519

## Part II Advanced Theory

<b>16 Speaker Modeling</b> .....	525
16.1 Individual Speaker Modeling .....	526
16.2 Background Models and Cohorts .....	527
16.2.1 Background Models .....	528
16.2.2 Cohorts .....	529
16.3 Pooling of Data and Speaker Independent Models .....	529
16.4 Speaker Adaptation .....	530
16.4.1 Factor Analysis (FA) .....	530
16.4.2 Joint Factor Analysis (JFA) .....	531
16.4.3 Total Factors (Total Variability) .....	532
16.5 Audio Segmentation .....	532
16.6 Model Quality Assessment .....	534
16.6.1 Enrollment Utterance Quality Control .....	534
16.6.2 Speaker Menagerie .....	536
References .....	538



<b>17 Speaker Recognition</b> .....	543
17.1 The Enrollment Task .....	543
17.2 The Verification Task .....	544
17.2.1 Text-Dependent .....	546
17.2.2 Text-Prompted .....	546
17.2.3 Knowledge-Based .....	548
17.3 The Identification Task .....	548
17.3.1 Closed-Set Identification .....	548
17.3.2 Open-Set Identification .....	549
17.4 Speaker Segmentation .....	549
17.5 Speaker and Event Classification .....	550
17.5.1 Gender and Age Classification (Identification) .....	551
17.5.2 Audio Classification .....	553
17.5.3 Multiple Codebooks .....	553
17.5.4 Farfield Speaker Recognition .....	553
17.5.5 Whispering Speaker Recognition .....	554
17.6 Speaker Diarization .....	554
17.6.1 Speaker Position and Orientation .....	555
References .....	555
<b>18 Signal Enhancement and Compensation</b> .....	561
18.1 Silence Detection, Voice Activity Detection (VAD) .....	561
18.2 Audio Volume Estimation .....	564
18.3 Echo Cancellation .....	564
18.4 Spectral Filtering and Cepstral Liftering .....	565
18.4.1 Cepstral Mean Normalization (Subtraction) – CMN (CMS) .....	567
18.4.2 Cepstral Mean and Variance Normalization (CMVN) .....	569
18.4.3 Cepstral Histogram Normalization (Histogram Equalization) .....	570
18.4.4 RelATive SpecTrAl (RASTA) Filtering .....	571
18.4.5 Other Lifters .....	571
18.4.6 Vocal Tract Length Normalization (VTLN) .....	573
18.4.7 Other Normalization Techniques .....	576
18.4.8 Steady Tone Removal (Narrowband Noise Reduction) .....	579
18.4.9 Adaptive Wiener Filtering .....	580
18.5 Speaker Model Normalization .....	581
18.5.1 Z-Norm .....	581
18.5.2 T-Norm (Test Norm) .....	582
18.5.3 H-Norm .....	582
18.5.4 HT-Norm .....	582
18.5.5 AT-Norm .....	582
18.5.6 C-Norm .....	582
18.5.7 D-Norm .....	583
18.5.8 F-Norm (F-Ratio Normalization) .....	583
18.5.9 Group-Specific Normalization .....	583

18.5.10 Within Class Covariance Normalization (WCCN) . . . . .	583
18.5.11 Other Normalization Techniques . . . . .	583
References . . . . .	584

### Part III Practice

<b>19 Evaluation and Representation of Results . . . . .</b>	<b>589</b>
19.1 Verification Results . . . . .	589
19.1.1 Equal-Error Rate . . . . .	589
19.1.2 Half Total Error Rate . . . . .	590
19.1.3 Receiver Operating Characteristic (ROC) Curve . . . . .	590
19.1.4 Detection Error Trade-Off (DET) Curve . . . . .	592
19.1.5 Detection Cost Function (DCF) . . . . .	593
19.2 Identification Results . . . . .	593
References . . . . .	594
<b>20 Time Lapse Effects (Case Study) . . . . .</b>	<b>595</b>
20.1 The Audio Data . . . . .	598
20.2 Baseline Speaker Recognition . . . . .	598
References . . . . .	600
<b>21 Adaptation over Time (Case Study) . . . . .</b>	<b>601</b>
21.1 Data Augmentation . . . . .	601
21.2 Maximum A Posteriori (MAP) Adaptation . . . . .	603
21.3 Eigenvoice Adaptation . . . . .	605
21.4 Minimum Classification Error (MCE) . . . . .	605
21.5 Linear Regression Techniques . . . . .	606
21.5.1 Maximum Likelihood Linear Regression (MLLR) . . . . .	606
21.6 Maximum a-Posteriori Linear Regression (MAPLR) . . . . .	607
21.6.1 Other Adaptation Techniques . . . . .	607
21.7 Practical Perspectives . . . . .	607
References . . . . .	608
<b>22 Overall Design . . . . .</b>	<b>611</b>
22.1 Choosing the Model . . . . .	611
22.1.1 Phonetic Speaker Recognition . . . . .	612
22.2 Choosing an Adaptation Technique . . . . .	613
22.3 Microphones . . . . .	613
22.4 Channel Mismatch . . . . .	615
22.5 Voice Over Internet Protocol (VoIP) . . . . .	615
22.6 Public Databases . . . . .	616
22.6.1 NIST . . . . .	616
22.6.2 Linguistic Data Consortium (LDC) . . . . .	616
22.6.3 European Language Resources Association (ELRA) . . . . .	619
22.7 High Level Information . . . . .	620
22.7.1 Choosing Basic Segments . . . . .	622

22.8	Numerical Stability	623
22.9	Privacy	624
22.10	Biometric Encryption	625
22.11	Spoofing	625
22.11.1	Text-Prompted Verification Systems	625
22.11.2	Text-Independent Verification Systems	626
22.12	Quality Issues	627
22.13	Large-Scale Systems	628
22.14	Useful Tools	628
	References	629

## Part IV Background Material

<b>23</b>	<b>Linear Algebra</b>	635
23.1	Basic Definitions	635
23.2	Norms	636
23.3	Gram-Schmidt Orthogonalization	641
23.3.1	Ordinary Gram-Schmidt Orthogonalization	641
23.3.2	Modified Gram-Schmidt Orthogonalization	641
23.4	Sherman-Morrison Inversion Formula	642
23.5	Vector Representation under a Set of Normal Conjugate Direction	642
23.6	Stochastic Matrix	643
23.7	Linear Equations	643
	References	646
<b>24</b>	<b>Integral Transforms</b>	647
24.1	Complex Variable Theory in Integral Transforms	648
24.1.1	Complex Variables	648
24.1.2	Limits	651
24.1.3	Continuity and Forms of Discontinuity	652
24.1.4	Convexity and Concavity of Functions	658
24.1.5	Odd, Even and Periodic Functions	661
24.1.6	Differentiation	663
24.1.7	Analyticity	665
24.1.8	Integration	672
24.1.9	Power Series Expansion of Functions	683
24.1.10	Residues	686
24.2	Relations Between Functions	688
24.2.1	Convolution	688
24.2.2	Correlation	689
24.3	Orthogonality of Functions	690
24.4	Integral Equations	694
24.5	Kernel Functions	696
24.5.1	Hilbert's Expansion Theorem	698
24.5.2	Eigenvalues and Eigenfunctions of the Kernel	700

24.6	Fourier Series Expansion . . . . .	708
24.6.1	Convergence of the Fourier Series . . . . .	713
24.6.2	Parseval's Theorem . . . . .	714
24.7	Wavelet Series Expansion . . . . .	716
24.8	The Laplace Transform . . . . .	717
24.8.1	Inversion . . . . .	720
24.8.2	Some Useful Transforms . . . . .	721
24.9	Complex Fourier Transform (Fourier Integral Transform) . . . . .	722
24.9.1	Translation . . . . .	724
24.9.2	Scaling . . . . .	724
24.9.3	Symmetry Table . . . . .	724
24.9.4	Time and Complex Scaling and Shifting . . . . .	725
24.9.5	Convolution . . . . .	725
24.9.6	Correlation . . . . .	726
24.9.7	Parseval's Theorem . . . . .	726
24.9.8	Power Spectral Density . . . . .	728
24.9.9	One-Sided Power Spectral Density . . . . .	728
24.9.10	PSD-per-unit-time . . . . .	729
24.9.11	Wiener-Khintchine Theorem . . . . .	729
24.10	Discrete Fourier Transform (DFT) . . . . .	731
24.10.1	Inverse Discrete Fourier Transform (IDFT) . . . . .	732
24.10.2	Periodicity . . . . .	734
24.10.3	Plancherel and Parseval's Theorem . . . . .	734
24.10.4	Power Spectral Density (PSD) Estimation . . . . .	735
24.10.5	Fast Fourier Transform (FFT) . . . . .	736
24.11	Discrete-Time Fourier Transform (DTFT) . . . . .	738
24.11.1	Power Spectral Density (PSD) Estimation . . . . .	739
24.12	Complex Short-Time Fourier Transform (STFT) . . . . .	740
24.12.1	Discrete-Time Short-Time Fourier Transform DTSTFT . . . . .	744
24.12.2	Discrete Short-Time Fourier Transform DSTFT . . . . .	746
24.13	Discrete Cosine Transform (DCT) . . . . .	748
24.13.1	Efficient DCT Computation . . . . .	749
24.14	The z-Transform . . . . .	750
24.14.1	Translation . . . . .	756
24.14.2	Scaling . . . . .	756
24.14.3	Shifting – Time Lag . . . . .	757
24.14.4	Shifting – Time Lead . . . . .	757
24.14.5	Complex Translation . . . . .	757
24.14.6	Initial Value Theorem . . . . .	758
24.14.7	Final Value Theorem . . . . .	758
24.14.8	Real Convolution Theorem . . . . .	759
24.14.9	Inversion . . . . .	760
24.15	Cepstrum . . . . .	762
	References . . . . .	769

<b>25 Nonlinear Optimization</b> .....	773
25.1 Gradient-Based Optimization .....	775
25.1.1 The Steepest Descent Technique .....	775
25.1.2 Newton's Minimization Technique .....	777
25.1.3 Quasi-Newton or Large Step Gradient Techniques .....	779
25.1.4 Conjugate Gradient Methods .....	793
25.2 Gradient-Free Optimization .....	803
25.2.1 Search Methods .....	804
25.2.2 Gradient-Free Conjugate Direction Methods .....	804
25.3 The Line Search Sub-Problem .....	809
25.4 Practical Considerations .....	810
25.4.1 Large-Scale Optimization .....	810
25.4.2 Numerical Stability .....	813
25.4.3 Nonsmooth Optimization .....	814
25.5 Constrained Optimization .....	814
25.5.1 The Lagrangian and Lagrange Multipliers .....	817
25.5.2 Duality .....	831
25.6 Global Convergence .....	835
References .....	836
<b>26 Standards</b> .....	841
26.1 Standard Audio Formats .....	842
26.1.1 Linear PCM (Uniform PCM) .....	842
26.1.2 $\mu$ -Law PCM (PCMU) .....	843
26.1.3 A-Law (PCMA) .....	843
26.1.4 MP3 .....	843
26.1.5 HE-AAC .....	844
26.1.6 OGG Vorbis .....	844
26.1.7 ADPCM (G.726) .....	845
26.1.8 GSM .....	845
26.1.9 CELP .....	847
26.1.10 DTMF .....	848
26.1.11 Others Audio Formats .....	848
26.2 Standard Audio Encapsulation Formats .....	849
26.2.1 WAV .....	849
26.2.2 SPHERE .....	850
26.2.3 Standard Audio Format Encapsulation (SAFE) .....	850
26.3 APIs and Protocols .....	854
26.3.1 SVAPI .....	855
26.3.2 BioAPI .....	855
26.3.3 VoiceXML .....	856
26.3.4 MRCP .....	857
26.3.5 Real-time Transport Protocol (RTP) .....	858
26.3.6 Extensible MultiModal Annotation (EMMA) .....	858
References .....	859

<b>Bibliography</b> .....	861
Bibliography .....	861
<b>Solutions</b> .....	901
<b>Index</b> .....	909



<http://www.springer.com/978-0-387-77591-3>

Fundamentals of Speaker Recognition

Beigi, H.

2011, LXI, 942 p. 177 illus., Hardcover

ISBN: 978-0-387-77591-3