

Contents

1	Pattern Recognition for Speech Detection	1
1.1	Introduction	1
1.2	Back-propagation Neural Network	3
1.2.1	Back-propagation Algorithm	5
1.2.2	ANN Illustration	7
1.3	Support Vector Machine	10
1.3.1	Dual Problem to Solve (1.25)–(1.28)	14
1.3.2	“Kernel-Trick” for Nonlinear Separation in SVM	15
1.3.3	Illustration for Support Vector Machine	16
1.4	Hidden Markov Model	23
1.4.1	Baum–Welch Technique to Obtain the Unknown Parameters in HMM	26
1.4.2	Steps to Compute the Unknown Parameters of HMM Using Expectation–Maximization Algorithm	28
1.4.3	Viterbi Algorithm to Compute the Generating Probability of the Arbitrary Speech Segment	29
1.4.4	Isolated Word Recognition Using HMM	30
1.4.5	Alignment Method to Model HMM	31
1.4.6	Illustration of Hidden Markov Model	31
1.5	Gaussian Mixture Model	37
1.5.1	Steps Involved to Model GMM Using Expectation–Maximization Algorithm	39
1.5.2	Isolated Word Recognition Using GMM	39
1.5.3	Illustration of GMM	40
1.6	Unsupervised Learning System	43
1.6.1	Need for Unsupervised Learning System	43
1.6.2	k-Means Algorithm	43
1.6.3	Illustration of k-Means Algorithm	44
1.6.4	Fuzzy k-Means Algorithm	44
1.6.5	Steps Involved in Fuzzy k-Means Clustering	46
1.6.6	Illustration of Fuzzy k-Means Algorithm	46
1.6.7	Kohonen Self-Organizing Map	48
1.6.8	Illustration of KSOM	51

1.7	Dimensionality Reduction Techniques	53
1.7.1	Principal Component Analysis	53
1.7.2	Illustration of PCA Using 2D to 1D Conversion	54
1.7.3	Illustration of PCA	54
1.7.4	Linear Discriminant Analysis	55
1.7.5	Small Sample Size Problem in LDA	57
1.7.6	Null-Space LDA	58
1.7.7	Kernel LDA	58
1.7.8	Kernel-Trick to Execute LDA in the Higher-Dimensional Space	59
1.7.9	Illustration of Dimensionality Reduction Using LDA	60
1.8	Independent Component Analysis	64
1.8.1	Solving ICA Bases Using Kurtosis Measurement	66
1.8.2	Steps to Obtain the ICA Bases	68
1.8.3	Illustration of Dimensionality Reduction Using ICA	68
2	Speech Production Model	73
2.1	Introduction	73
2.2	1-D Sound Waves	74
2.2.1	Physics on Sound Wave Travelling Through the Tube with Uniform Cross-Sectional Area A	74
2.2.2	Solution to (2.9) and (2.18)	76
2.3	Vocal Tract Model as the Cascade Connections of Identical Length Tubes with Different Cross-Sections	78
2.4	Modelling the Vocal Tract from the Speech Signal	82
2.4.1	Autocorrelation Method	82
2.4.2	Auto Covariance Method	87
2.5	Lattice Structure to Obtain Excitation Source for the Typical Speech Signal	88
2.5.1	Computation of Lattice Co-efficient from LPC Co-efficients	89
3	Feature Extraction of the Speech Signal	93
3.1	Endpoint Detection	93
3.2	Dynamic Time Warping	96
3.3	Linear Predictive Co-efficients	100
3.4	Poles of the Vocal Tract	103
3.5	Reflection Co-efficients	105
3.6	Log Area Ratio	105
3.7	Cepstrum	105
3.8	Line Spectral Frequencies	108
3.9	Mel-Frequency Cepstral Co-efficients	113
3.9.1	Gibbs Phenomenon	117
3.9.2	Discrete Cosine Transformation	118

3.10	Spectrogram	120
3.10.1	Time Resolution Versus Frequency Resolution in Spectrogram	124
3.11	Discrete Wavelet Transformation	124
3.12	Pitch Frequency Estimation	126
3.12.1	Autocorrelation Approach	126
3.12.2	Homomorphic Filtering Approach	127
3.13	Formant Frequency Estimation	129
3.13.1	Formant Extraction Using Vocal Tract Model	129
3.13.2	Formant Extraction Using Homomorphic Filtering	132
4	Speech Compression	135
4.1	Uniform Quantization	135
4.2	Nonuniform Quantization	136
4.3	Adaptive Quantization	139
4.4	Differential Pulse Code Modulation	140
4.4.1	Illustrations of the Prediction of Speech Signal Using lpc	140
4.5	Code-Excited Linear Prediction	142
4.5.1	Estimation of the Delay Constant D	145
4.5.2	Estimation of the Gain Constants G_1 and G_2	146
4.6	Assessment of the Quality of the Compressed Speech Signal	150
Appendix A: Constrained Optimization Using Lagrangian Techniques		151
Appendix B: Expectation–Maximization Algorithm		157
Appendix C: Diagonalization of the Matrix		161
Appendix D: Condition Number		163
Appendix E: Spectral Flatness		169
Appendix F: Functional Blocks of the Vocal Tract and the Ear		175
About the Author		177
About the Book		179
Index		181



<http://www.springer.com/978-81-322-1676-6>

Digital Speech Processing Using Matlab

Gopi, E.S.

2014, XVI, 182 p. 103 illus., Hardcover

ISBN: 978-81-322-1676-6