

Contents

- 1 Introduction** 1
 - 1.1 Authentication 1
 - 1.2 Biometric-Based Authentication 3
 - 1.3 Information-Based Authentication 5
 - 1.4 Speaker Authentication 6
 - 1.4.1 Speaker Recognition 7
 - 1.4.2 Verbal Information Verification 9
 - 1.5 Historical Perspective and Further Reading 11
 - 1.6 Book Organization 13
 - References 18

- 2 Multivariate Statistical Analysis and One-Pass Vector Quantization** 23
 - 2.1 Multivariate Gaussian Distribution 23
 - 2.2 Principal Component Analysis 25
 - 2.3 Vector Quantization 27
 - 2.4 One-Pass VQ 28
 - 2.4.1 The One-Pass VQ Algorithm 28
 - 2.4.2 Steps of the One-Pass VQ Algorithm 32
 - 2.4.3 Complexity Analysis 33
 - 2.4.4 Codebook Design Examples 34
 - 2.4.5 Robustness Analysis 38
 - 2.5 Segmental *K*-Means 39
 - 2.6 Conclusions 40
 - References 40

- 3 Principal Feature Networks for Pattern Recognition** 43
 - 3.1 Overview of the Design Concept 43
 - 3.2 Implementations of Principal Feature Networks 46
 - 3.3 Hidden Node Design 48
 - 3.3.1 Gaussian Discriminant Node 49

3.3.2	Fisher’s Node Design	51
3.4	Principal Component Hidden Node Design	52
3.4.1	Principal Component Discriminant Analysis	53
3.5	Relation between PC Node and the Optimal Gaussian Classifier	53
3.6	Maximum Signal-to-Noise-Ratio (SNR) Hidden Node Design ..	55
3.7	Determining the Thresholds from Design Specifications	56
3.8	Simplification of the Hidden Nodes	56
3.9	Application 1 – Data Recognition	56
3.10	Application 2 – Multispectral Pattern Recognition	58
3.11	Conclusions	59
	References	59
4	Non-Stationary Pattern Recognition	61
4.1	Introduction	61
4.2	Gaussian Mixture Models (GMM) for Stationary Process	62
4.2.1	An Illustrative Example	63
4.3	Hidden Markov Model (HMM) for Non-Stationary Process	66
4.4	Speech Segmentation	68
4.5	Bayesian Decision Theory	68
4.6	Statistical Verification	70
4.7	Conclusions	71
	References	72
5	Robust Endpoint Detection	75
5.1	Introduction	76
5.2	A Filter for Endpoint Detection	78
5.3	Real-Time Endpoint Detection and Energy Normalization	81
5.3.1	A Filter for Both Beginning- and Ending-Edge Detection	82
5.3.2	Decision Diagram	82
5.3.3	Real-Time Energy Normalization	83
5.3.4	Database Evaluation	85
5.4	Conclusions	88
	References	89
6	Detection-Based Decoder	93
6.1	Introduction	94
6.2	Change-Point Detection	96
6.3	HMM State Change-Point Detection	97
6.4	HMM Search-Space Reduction	100
6.4.1	Concept of Search-Space Reduction	100
6.4.2	Algorithm Summary and Complexity Analysis	102
6.5	Experiments	104
6.5.1	An Example of State Change-Point Detection	104
6.5.2	Application to Speaker Verification	104
6.6	Conclusions	108

References	109
7 Auditory-Based Time Frequency Transform	111
7.1 Introduction	112
7.1.1 Observing Problems with the Fourier Transform	112
7.1.2 Brief Introduction of the Ear	114
7.1.3 Time-Frequency Analyses	117
7.2 Definition of the Auditory-Based Transform	118
7.3 The Inverse Auditory Transform	120
7.4 The Discrete-Time and Fast Transform	123
7.5 Experiments and Discussions	124
7.5.1 Verifying the Inverse Auditory Transform	124
7.5.2 Applications	126
7.6 Comparisons to Other Transforms	127
7.7 Conclusions	131
References	131
8 Auditory-Based Feature Extraction and Robust Speaker Identification	135
8.1 Introduction	135
8.2 Auditory-Based Feature Extraction Algorithm	138
8.2.1 Forward Auditory Transform and Cochlea Filter Bank	138
8.2.2 Cochlear filter cepstral coefficients (CFCC)	140
8.2.3 Analysis and Comparison	141
8.3 Speaker Identification and Experimental Evaluation	142
8.3.1 Experimental Datasets	142
8.3.2 The Baseline Speaker Identification System	143
8.3.3 Experiments	145
8.3.4 Further Comparison with PLP and RASTA-PLP	146
8.4 Conclusions	147
References	149
9 Fixed-Phrase Speaker Verification	151
9.1 Introduction	151
9.2 A Fixed-Phrase System	152
9.3 An Evaluation Database and Model Parameters	154
9.4 Adaptation and Reference Results	155
9.5 Conclusions	155
References	156
10 Robust Speaker Verification with Stochastic Matching	157
10.1 Introduction	157
10.2 A Fast Stochastic Matching Algorithm	158
10.3 Fast Estimation for a General Linear Transform	160
10.4 Speaker Verification with Stochastic Matching	161

10.5 Database and Experiments	163
10.6 Conclusions	164
References	164
11 Randomly Prompted Speaker Verification	165
11.1 Introduction	165
11.2 Normalized Discriminant Analysis	168
11.3 Applying NDA in the Hybrid Speaker-Verification System	169
11.3.1 Training of the NDA System	169
11.3.2 Training of the HMM System	171
11.3.3 Training of the Data Fusion Layer	173
11.4 Speaker Verification Experiments	173
11.4.1 Experimental Database	173
11.4.2 NDA System Results	174
11.4.3 Hybrid Speaker-Verification System Results	174
11.5 Conclusions	175
References	176
12 Objectives for Discriminative Training	179
12.1 Introduction	179
12.2 Error Rates vs. Posterior Probability	180
12.3 Minimum Classification Error vs. Posterior Probability	181
12.4 Maximum Mutual Information vs. Minimum Classification Error	183
12.5 Generalized Minimum Error Rate vs. Other Objectives	185
12.6 Experimental Comparisons	186
12.7 Discussion	186
12.8 Relations between Objectives and Optimization Algorithms	187
12.9 Conclusions	188
References	189
13 Fast Discriminative Training	191
13.1 Introduction	191
13.2 Objective for Fast Discriminative Training	193
13.3 Derivation of Fast Estimation Formulas	195
13.3.1 Estimation of Covariance Matrices	196
13.3.2 Determination of Weighting Scalar	196
13.3.3 Estimation of Mean Vectors	197
13.3.4 Estimation of Mixture Parameters	198
13.3.5 Discussions	198
13.4 Summary of Practical Training Procedure	200
13.5 Experiments	200
13.5.1 Continuing the Illustrative Example	200
13.5.2 Application to Speaker Identification	204
13.6 Conclusions	204

References 205

14 Verbal Information Verification 207

14.1 Introduction 207

14.2 Single Utterance Verification 209

 14.2.1 Normalized Confidence Measures 211

14.3 Sequential Utterance Verification 212

 14.3.1 Examples in Sequential-Test Design 214

14.4 VIV Experimental Results 215

14.5 Conclusions 219

References 220

15 Speaker Authentication System Design 223

15.1 Introduction 223

15.2 Automatic Enrollment by VIV 224

15.3 Fixed-Phrase Speaker Verification 226

15.4 Experiments 227

 15.4.1 Features and Database 227

 15.4.2 Experimental Results on Using VIV for SV Enrollment 228

15.5 Conclusions 229

References 230

Index 231



<http://www.springer.com/978-3-642-23730-0>

Speaker Authentication

Li, Q.P.

2012, XXVI, 238 p., Hardcover

ISBN: 978-3-642-23730-0