Springer Handbook of Speech Processing

Bearbeitet von
Jacob Benesty, M. M Sondhi, Yiteng Huang

ISBN 978 3 540 49125 5
Format (B x L): 19,3 x 24,2 cm

Weitere Fachgebiete > EDV, Informatik > Informationsverarbeitung > Spracherkennung, Sprachverarbeitung
Zu Leseprobe
schnell und portofrei erhältlich bei

beck-shop.de
DIE FACHBUCHHANDLUNG

Die Online-Fachbuchhandlung beck-shop.de ist spezialisiert auf Fachbücher, insbesondere Recht, Steuern und Wirtschaft. Im Sortiment finden Sie alle Medien (Bücher, Zeitschriften, CDs, eBooks, etc.) aller Verlage. Ergänzt wird das Programm durch Services wie Neuerscheinungsdienst oder Zusammenstellungen von Büchern zu Sonderpreisen. Der Shop führt mehr als 8 Millionen Produkte.
Foreword

Over the past three decades digital signal processing has emerged as a recognized discipline. Much of the impetus for this advance stems from research in representation, coding, transmission, storage and reproduction of speech and image information. In particular, interest in voice communication has stimulated central contributions to digital filtering and discrete-time spectral transforms.

This dynamic development was built upon the convergence of three then-evolving technologies: (i) sampled-data theory and representation of information signals (which led directly to digital telecommunication that provides signal quality independent of transmission distance); (ii) electronic binary computation (aided in early implementation by pulse-circuit techniques from radar design); and, (iii) invention of solid-state devices for exquisite control of electronic current (transistors – which now, through microelectronic materials, scale to systems of enormous size and complexity). This timely convergence was soon followed by optical fiber methods for broadband information transport.

These advances impact an important aspect of human activity – information exchange. And, over man’s existence, speech has played a principal role in human communication. Now, speech is playing an increasing role in human interaction with complex information systems. Automatic services of great variety exploit the comfort of voice exchange, and, in the corporate sector, sophisticated audio/video teleconferencing is reducing the necessity of expensive, time-consuming business travel. In each instance an overarching target is a user environment that captures some of the naturalness and spatial realism of face-to-face communication. Again, speech is a core element, and new understanding from diverse research sectors can be brought to bear.

Editors-in-Chief Benesty, Sondhi and Huang have organized a timely engineering handbook to answer this need. They have assembled a remarkable compendium of current knowledge in speech processing. And, this accumulated understanding can be focused upon enlarging the human capacity to deal with a world ever increasing in complexity. Benesty, Sondhi and Huang are renowned researchers in their own right, and they have attracted an international cadre of over 80 fellow authors and collaborators who constitute a veritable Who’s Who of world leaders in speech processing research. The resulting book provides under one cover authoritative treatments that commence with the basic physics and psychophysics of speech and hearing, and range through the related topics of computational tools, coding, synthesis, recognition, and signal enhancement, concluding with discussions on capture and projection of sound in enclosures. The book can be expected to become a valuable resource for researchers, engineers and speech scientists throughout the global community. It should equally serve teachers and students in human communication, especially delimiting knowledge frontiers where graduate thesis research may be appropriate.

Warren, New Jersey

Jim Flanagan

October 2007
Preface

The achievement of this Springer Handbook is the result of a wonderful journey that started in March 2005 at the 30th International Conference on Acoustics, Speech, and Signal Processing (ICASSP). Two of the editors-in-chief (Benesty and Huang) met in one of the long corridors of the Pennsylvania Convention Center in Philadelphia with Dr. Dieter Merkle from Springer. Together we had a very nice discussion about the conference and immediately an idea came up for a handbook. After a short discussion we converged without too much hesitation on a handbook of speech processing. It was quite surprising to see that, even after 30 years of ICASSP and more than half a century of research in this fundamental area, there was still no major book summarizing the important aspects of speech processing. We thought that the time was ripe for such a large project. Soon after we got home, a third editor-in-chief (Sondhi) joined the efforts.

We had a very clear objective in our minds: to summarize, in a reasonable number of pages, the most important and useful aspects of speech processing. The content was then organized accordingly. This task was not easy since we had to find a good balance between feasible ideas and new trends. As we all know, practical ideas can be viewed as old stuff while emerging ideas can be criticized for not having passed the test of time; we hope that we have succeeded in finding a good compromise. For this we relied on many authors who are well established and are recognized as experts in their field, from all over the world, and from academia as well as from industry.

From simple consumer products such as cell phones and MP3 players to more sophisticated projects such as human-machine interfaces and robots that can obey orders, speech technologies are now everywhere. We believe that it is just a matter of time before more applications of the science of speech become impossible to miss in our daily life. So we believe that this Springer Handbook will play a fundamental role in the sustainable progress of speech research and development.

This handbook is targeted at three categories of readers: graduate students of speech processing, professors and researchers in academia and research labs who are active in this field, and engineers in industry who need to understand or implement specific algorithms for their speech-related products. The handbook could also be used as a text for one or more graduate courses on signal processing for speech and various aspects of speech processing and applications.

For the completion of such an ambitious project we have many people to thank. First, we would like to thank the many authors who did a terrific job in delivering very high-quality chapters. Second, we are very grateful to the members of the editorial board who helped us so much in organizing the content and structure of this book, taking part in all phases of this project from conception to completion. Third, we would like to thank all the reviewers, who helped us to improve the quality of the material. Last, but not least, we would like to thank the Springer team for their availability and very professional work. In particular, we appreciated the help of Dieter Merkle, Christoph Baumann, Werner Skolaut, Petra Jantzen, and Claudia Rau.

We hope this Springer Handbook will inspire many great minds to find new research ideas or to implement algorithms in products.

Montreal, Basking Ridge, Murray Hill
October 2007

Jacob Benesty
M. Mohan Sondhi
Yiteng Huang
**Contents**

**List of Abbreviations** ........................................................................................................... XXXI

1 Introduction to Speech Processing  
*J. Benesty, M. M. Sondhi, Y. Huang* ................................................................. 1  
1.1 A Brief History of Speech Processing ................................................................. 1  
1.2 Applications of Speech Processing ...................................................................... 2  
1.3 Organization of the Handbook ............................................................................ 4  
References ......................................................................................................................... 4

Part A Production, Perception, and Modeling of Speech

2 Physiological Processes of Speech Production  
*K. Honda* ..................................................................................................................... 7  
2.1 Overview of Speech Apparatus ............................................................................ 7  
2.2 Voice Production Mechanisms ........................................................................... 8  
2.3 Articulatory Mechanisms ................................................................................... 14  
2.4 Summary ............................................................................................................. 24  
References ......................................................................................................................... 25

3 Nonlinear Cochlear Signal Processing and Masking in Speech Perception  
*J. B. Allen* .................................................................................................................. 27  
3.1 Basics ................................................................................................................... 27  
3.2 The Nonlinear Cochlea ....................................................................................... 35  
3.3 Neural Masking .................................................................................................... 45  
3.4 Discussion and Summary .................................................................................... 55  
References ......................................................................................................................... 56

4 Perception of Speech and Sound  
*B. Kollmeier, T. Brand, B. Meyer* ............................................................................ 61  
4.1 Basic Psychoacoustic Quantities ......................................................................... 62  
4.2 Acoustical Information Required for Speech Perception ................................... 70  
4.3 Speech Feature Perception .................................................................................. 74  
References ......................................................................................................................... 81

5 Speech Quality Assessment  
*V. Grancharov, W. B. Kleijn* .................................................................................. 83  
5.1 Degradation Factors Affecting Speech Quality ................................................... 84  
5.2 Subjective Tests .................................................................................................... 85  
5.3 Objective Measures ............................................................................................ 90  
5.4 Conclusions ......................................................................................................... 95  
References ......................................................................................................................... 96
### Part B Signal Processing for Speech

#### 6 Wiener and Adaptive Filters

*J. Benesty, Y. Huang, J. Chen*

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Overview</td>
<td>103</td>
</tr>
<tr>
<td>6.2 Signal Models</td>
<td>104</td>
</tr>
<tr>
<td>6.3 Derivation of the Wiener Filter</td>
<td>106</td>
</tr>
<tr>
<td>6.4 Impulse Response Tail Effect</td>
<td>107</td>
</tr>
<tr>
<td>6.5 Condition Number</td>
<td>108</td>
</tr>
<tr>
<td>6.6 Adaptive Algorithms</td>
<td>110</td>
</tr>
<tr>
<td>6.7 MIMO Wiener Filter</td>
<td>116</td>
</tr>
<tr>
<td>6.8 Conclusions</td>
<td>119</td>
</tr>
</tbody>
</table>

**References** .................................................................................................................. 119

#### 7 Linear Prediction

*J. Benesty, J. Chen, Y. Huang*

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1 Fundamentals</td>
<td>121</td>
</tr>
<tr>
<td>7.2 Forward Linear Prediction</td>
<td>122</td>
</tr>
<tr>
<td>7.3 Backward Linear Prediction</td>
<td>123</td>
</tr>
<tr>
<td>7.4 Levinson–Durbin Algorithm</td>
<td>124</td>
</tr>
<tr>
<td>7.5 Lattice Predictor</td>
<td>126</td>
</tr>
<tr>
<td>7.6 Spectral Representation</td>
<td>127</td>
</tr>
<tr>
<td>7.7 Linear Interpolation</td>
<td>128</td>
</tr>
<tr>
<td>7.8 Line Spectrum Pair Representation</td>
<td>129</td>
</tr>
<tr>
<td>7.9 Multichannel Linear Prediction</td>
<td>130</td>
</tr>
<tr>
<td>7.10 Conclusions</td>
<td>133</td>
</tr>
</tbody>
</table>

**References** ..................................................................................................................... 133

#### 8 The Kalman Filter

*S. Gannot, A. Yeredor*

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 Derivation of the Kalman Filter</td>
<td>136</td>
</tr>
<tr>
<td>8.2 Examples: Estimation of Parametric Stochastic Process from Noisy Observations</td>
<td>141</td>
</tr>
<tr>
<td>8.3 Extensions of the Kalman Filter</td>
<td>144</td>
</tr>
<tr>
<td>8.4 The Application of the Kalman Filter to Speech Processing</td>
<td>149</td>
</tr>
<tr>
<td>8.5 Summary</td>
<td>157</td>
</tr>
</tbody>
</table>

**References** ..................................................................................................................... 157

#### 9 Homomorphic Systems and Cepstrum Analysis of Speech

*R. W. Schafer*

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1 Definitions</td>
<td>161</td>
</tr>
<tr>
<td>9.2 Z–Transform Analysis</td>
<td>164</td>
</tr>
<tr>
<td>9.3 Discrete–Time Model for Speech Production</td>
<td>165</td>
</tr>
<tr>
<td>9.4 The Cepstrum of Speech</td>
<td>166</td>
</tr>
<tr>
<td>9.5 Relation to LPC</td>
<td>169</td>
</tr>
<tr>
<td>9.6 Application to Pitch Detection</td>
<td>171</td>
</tr>
</tbody>
</table>

**References** ..................................................................................................................... 171
## 9.7 Applications to Analysis/Synthesis Coding

172

## 9.8 Applications to Speech Pattern Recognition

176

## 9.9 Summary

180

## References

180

---

## 10 Pitch and Voicing Determination of Speech with an Extension Toward Music Signals

*W. J. Hess*

181

### 10.1 Pitch in Time-Variant Quasiperiodic Acoustic Signals

182

### 10.2 Short-Term Analysis PDAs

185

### 10.3 Selected Time-Domain Methods

192

### 10.4 A Short Look into Voicing Determination

195

### 10.5 Evaluation and Postprocessing

197

### 10.6 Applications in Speech and Music

201

### 10.7 Some New Challenges and Developments

203

### 10.8 Concluding Remarks

207

## References

208

---

## 11 Formant Estimation and Tracking

*D. O'Shaughnessy*

213

### 11.1 Historical

213

### 11.2 Vocal Tract Resonances

215

### 11.3 Speech Production

216

### 11.4 Acoustics of the Vocal Tract

218

### 11.5 Short-Time Speech Analysis

221

### 11.6 Formant Estimation

223

### 11.7 Summary

226

## References

226

---

## 12 The STFT, Sinusoidal Models, and Speech Modification

*M. M. Goodwin*

229

### 12.1 The Short-Time Fourier Transform

230

### 12.2 Sinusoidal Models

242

### 12.3 Speech Modification

253

## References

256

---

## 13 Adaptive Blind Multichannel Identification

*Y. Huang, J. Benesty, J. Chen*

259

### 13.1 Overview

259

### 13.2 Signal Model and Problem Formulation

260

### 13.3 Identifiability and Principle

261

### 13.4 Constrained Time-Domain Multichannel LMS and Newton Algorithms

262

### 13.5 Unconstrained Multichannel LMS Algorithm with Optimal Step-Size Control

266

### 13.6 Frequency-Domain Blind Multichannel Identification Algorithms

268

### 13.7 Adaptive Multichannel Exponentiated Gradient Algorithm

276
Part C Speech Coding

14 Principles of Speech Coding
W. B. Kleijn

14.1 The Objective of Speech Coding .............................................................. 283
14.2 Speech Coder Attributes ........................................................................ 284
14.3 A Universal Coder for Speech ................................................................. 286
14.4 Coding with Autoregressive Models ......................................................... 293
14.5 Distortion Measures and Coding Architecture .......................................... 296
14.6 Summary .................................................................................................. 302

References .................................................................................................... 303

15 Voice over IP: Speech Transmission over Packet Networks
J. Skoglund, E. Kozica, J. Linden, R. Hagen, W. B. Kleijn

15.1 Voice Communication ............................................................................... 307
15.2 Properties of the Network ........................................................................ 308
15.3 Outline of a VoIP System .......................................................................... 313
15.4 Robust Encoding ...................................................................................... 317
15.5 Packet Loss Concealment ........................................................................ 326
15.6 Conclusion ............................................................................................... 327

References .................................................................................................... 328

16 Low-Bit-Rate Speech Coding
A. V. McCree

16.1 Speech Coding .......................................................................................... 331
16.2 Fundamentals: Parametric Modeling of Speech Signals ......................... 332
16.3 Flexible Parametric Models ...................................................................... 337
16.4 Efficient Quantization of Model Parameters ............................................ 344
16.5 Low-Rate Speech Coding Standards ......................................................... 345
16.6 Summary .................................................................................................. 347

References .................................................................................................... 347

17 Analysis-by-Synthesis Speech Coding
J.-H. Chen, J. Thyssen

17.1 Overview .................................................................................................. 351
17.2 Basic Concepts of Analysis-by-Synthesis Coding ...................................... 353
17.3 Overview of Prominent Analysis-by-Synthesis Speech Coders ................. 357
17.4 Multipulse Linear Predictive Coding (MPLPC) ........................................ 360
17.5 Regular-Pulse Excitation with Long-Term Prediction (RPE-LTP) ............. 362
17.6 The Original Code Excited Linear Prediction (CELP) Coder .................. 363
17.7 US Federal Standard FS1016 CELP ........................................................... 367
17.8 Vector Sum Excited Linear Prediction (VSELP) ........................................ 368
17.9 Low-Delay CELP (LD-CELP) ................................................................ 370
Part E Speech Recognition

26 Historical Perspective of the Field of ASR/NLU
L. Rabiner, B.-H. Juang .................................................. 521
26.1 ASR Methodologies .................................................. 521
26.2 Important Milestones in Speech Recognition History .......... 523
26.3 Generation 1 – The Early History of Speech Recognition .... 524
26.4 Generation 2 – The First Working Systems for Speech Recognition... 524
26.5 Generation 3 – The Pattern Recognition Approach to Speech Recognition ........................................ 525
26.6 Generation 4 – The Era of the Statistical Model ................. 530
26.7 Generation 5 – The Future ........................................... 534
26.8 Summary .................................................................... 534
References ........................................................................ 535

27 HMMs and Related Speech Recognition Technologies
S. Young ........................................................................... 539
27.1 Basic Framework ....................................................... 539
27.2 Architecture of an HMM-Based Recognizer ..................... 540
27.3 HMM-Based Acoustic Modeling .................................... 547
27.4 Normalization ............................................................ 550
27.5 Adaptation .................................................................. 551
27.6 Multipass Recognition Architectures ............................... 554
27.7 Conclusions ............................................................... 554
References ........................................................................ 555

28 Speech Recognition with Weighted Finite-State Transducers
M. Mohri, F. Pereira, M. Riley ............................................. 559
28.1 Definitions ................................................................. 559
28.2 Overview .................................................................... 560
28.3 Algorithms ............................................................... 567
28.4 Applications to Speech Recognition ............................... 574
28.5 Conclusion ............................................................... 582
References ........................................................................ 582

29 A Machine Learning Framework for Spoken-DIALOG Classification
C. Cortes, P. Haffner, M. Mohri ......................................... 585
29.1 Motivation ................................................................. 585
29.2 Introduction to Kernel Methods .................................... 586
29.3 Rational Kernels ......................................................... 587
29.4 Algorithms ............................................................... 589
29.5 Experiments ............................................................. 591
29.6 Theoretical Results for Rational Kernels ......................... 593
29.7 Conclusion ........................................................................................................ 594
References ............................................................................................................... 595

30 Towards Superhuman Speech Recognition
M. Picheny, D. Nahamoo ......................................................................................... 597
30.1 Current Status .................................................................................................. 597
30.2 A Multidomain Conversational Test Set ......................................................... 598
30.3 Listening Experiments ..................................................................................... 599
30.4 Recognition Experiments ............................................................................... 601
30.5 Speculation ...................................................................................................... 607
References ............................................................................................................... 614

31 Natural Language Understanding
S. Roukos ................................................................................................................. 617
31.1 Overview of NLU Applications ....................................................................... 618
31.2 Natural Language Parsing .............................................................................. 620
31.3 Practical Implementation .................................................................................. 623
31.4 Speech Mining .................................................................................................. 623
31.5 Conclusion ....................................................................................................... 625
References ............................................................................................................... 626

32 Transcription and Distillation of Spontaneous Speech
S. Furui, T. Kawahara .............................................................................................. 627
32.1 Background ...................................................................................................... 627
32.2 Overview of Research Activities on Spontaneous Speech ......................... 628
32.3 Analysis for Spontaneous Speech Recognition .............................................. 632
32.4 Approaches to Spontaneous Speech Recognition .......................................... 635
32.5 Metadata and Structure Extraction of Spontaneous Speech ....................... 640
32.6 Speech Summarization .................................................................................... 644
32.7 Conclusions ..................................................................................................... 647
References ............................................................................................................... 647

33 Environmental Robustness
J. Droppo, A. Acero .................................................................................................. 653
33.1 Noise Robust Speech Recognition .................................................................. 653
33.2 Model Retraining and Adaptation .................................................................. 656
33.3 Feature Transformation and Normalization ..................................................... 657
33.4 A Model of the Environment ......................................................................... 664
33.5 Structured Model Adaptation ......................................................................... 667
33.6 Structured Feature Enhancement .................................................................... 671
33.7 Unifying Model and Feature Techniques ........................................................ 675
33.8 Conclusion ....................................................................................................... 677
References ............................................................................................................... 677

34 The Business of Speech Technologies
J. Wilpon, M. E. Gilbert, J. Cohen ............................................................................ 681
34.1 Introduction ..................................................................................................... 682
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>34.2</td>
<td>Network-Based Speech Services</td>
</tr>
<tr>
<td>34.3</td>
<td>Device-Based Speech Applications</td>
</tr>
<tr>
<td>34.4</td>
<td>Vision/Predications of Future Services – Fueling the Trends</td>
</tr>
<tr>
<td>34.5</td>
<td>Conclusion</td>
</tr>
<tr>
<td>References</td>
<td></td>
</tr>
</tbody>
</table>

### Spoken Dialogue Systems

*V. Zue, S. Seneff*

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>35.1</td>
<td>Technology Components and System Development</td>
</tr>
<tr>
<td>35.2</td>
<td>Development Issues</td>
</tr>
<tr>
<td>35.3</td>
<td>Historical Perspectives</td>
</tr>
<tr>
<td>35.4</td>
<td>New Directions</td>
</tr>
<tr>
<td>35.5</td>
<td>Concluding Remarks</td>
</tr>
<tr>
<td>References</td>
<td></td>
</tr>
</tbody>
</table>

### Part F Speaker Recognition

#### 36 Overview of Speaker Recognition

*A. E. Rosenberg, F. Bimbot, S. Parthasarathy*

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>36.1</td>
<td>Speaker Recognition</td>
</tr>
<tr>
<td>36.2</td>
<td>Measuring Speaker Features</td>
</tr>
<tr>
<td>36.3</td>
<td>Constructing Speaker Models</td>
</tr>
<tr>
<td>36.4</td>
<td>Adaptation</td>
</tr>
<tr>
<td>36.5</td>
<td>Decision and Performance</td>
</tr>
<tr>
<td>36.6</td>
<td>Selected Applications for Automatic Speaker Recognition</td>
</tr>
<tr>
<td>36.7</td>
<td>Summary</td>
</tr>
<tr>
<td>References</td>
<td></td>
</tr>
</tbody>
</table>

#### 37 Text-Dependent Speaker Recognition

*M. Hébert*

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>37.1</td>
<td>Brief Overview</td>
</tr>
<tr>
<td>37.2</td>
<td>Text-Dependent Challenges</td>
</tr>
<tr>
<td>37.3</td>
<td>Selected Results</td>
</tr>
<tr>
<td>37.4</td>
<td>Concluding Remarks</td>
</tr>
<tr>
<td>References</td>
<td></td>
</tr>
</tbody>
</table>

#### 38 Text-Independent Speaker Recognition

*D. A. Reynolds, W. M. Campbell*

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>38.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>38.2</td>
<td>Likelihood Ratio Detector</td>
</tr>
<tr>
<td>38.3</td>
<td>Features</td>
</tr>
<tr>
<td>38.4</td>
<td>Classifiers</td>
</tr>
<tr>
<td>38.5</td>
<td>Performance Assessment</td>
</tr>
<tr>
<td>38.6</td>
<td>Summary</td>
</tr>
<tr>
<td>References</td>
<td></td>
</tr>
</tbody>
</table>
Part G Language Recognition

39 Principles of Spoken Language Recognition
C.-H. Lee .......................................................... 785
39.1 Spoken Language .................................................. 785
39.2 Language Recognition Principles .............................. 786
39.3 Phone Recognition Followed by Language Modeling (PRLM) ...... 788
39.4 Vector-Space Characterization (VSC) ......................... 789
39.5 Spoken Language Verification ................................ 790
39.6 Discriminative Classifier Design .............................. 791
39.7 Summary ................................................................ 793
References .................................................................. 793

40 Spoken Language Characterization
M. P. Harper, M. Maxwell ................................................ 797
40.1 Language versus Dialect ............................................ 798
40.2 Spoken Language Collections ................................... 800
40.3 Spoken Language Characteristics ................................ 800
40.4 Human Language Identification .................................. 804
40.5 Text as a Source of Information on Spoken Languages ....... 806
40.6 Summary .............................................................. 807
References .................................................................. 807

41 Automatic Language Recognition Via Spectral and Token Based Approaches
D. A. Reynolds, W. M. Campbell, W. Shen, E. Singer................. 811
41.1 Automatic Language Recognition ............................... 811
41.2 Spectral Based Methods ............................................ 812
41.3 Token-Based Methods ............................................. 815
41.4 System Fusion .......................................................... 818
41.5 Performance Assessment ............................................ 820
41.6 Summary .............................................................. 823
References .................................................................. 823

42 Vector-Based Spoken Language Classification
H. Li, B. Ma, C.-H. Lee ................................................. 825
42.1 Vector Space Characterization ..................................... 826
42.2 Unit Selection and Modeling ....................................... 827
42.3 Front-End: Voice Tokenization and Spoken Document Vectorization 830
42.4 Back-End: Vector-Based Classifier Design .................... 831
42.5 Language Classification Experiments and Discussion .......... 835
42.6 Summary .............................................................. 838
References .................................................................. 839
Part H Speech Enhancement

43 Fundamentals of Noise Reduction
J. Chen, J. Benesty, Y. Huang, E. J. Diethorn ....................................................... 843
43.1 Noise .................................................................................................................. 843
43.2 Signal Model and Problem Formulation ............................................................ 845
43.3 Evaluation of Noise Reduction ........................................................................... 846
43.4 Noise Reduction via Filtering Techniques .......................................................... 847
43.5 Noise Reduction via Spectral Restoration .......................................................... 857
43.6 Speech-Model-Based Noise Reduction ............................................................... 863
43.7 Summary .............................................................................................................. 868
References .................................................................................................................. 869

44 Spectral Enhancement Methods
I. Cohen, S. Gannot ...................................................................................................... 873
44.1 Spectral Enhancement .......................................................................................... 874
44.2 Problem Formulation .......................................................................................... 875
44.3 Statistical Models ............................................................................................... 876
44.4 Signal Estimation ............................................................................................... 879
44.5 Signal Presence Probability Estimation .............................................................. 881
44.6 A Priori SNR Estimation ...................................................................................... 882
44.7 Noise Spectrum Estimation ............................................................................... 888
44.8 Summary of a Spectral Enhancement Algorithm ............................................... 891
44.9 Selection of Spectral Enhancement Algorithms ............................................... 896
44.10 Conclusions ....................................................................................................... 898
References .................................................................................................................. 899

45 Adaptive Echo Cancelation for Voice Signals
M. M. Sondhi .............................................................................................................. 903
45.1 Network Echoes ............................................................................................... 904
45.2 Single-Channel Acoustic Echo Cancelation ...................................................... 915
45.3 Multichannel Acoustic Echo Cancelation ......................................................... 921
45.4 Summary ............................................................................................................ 925
References .................................................................................................................. 926

46 Dereverberation
Y. Huang, J. Benesty, J. Chen ..................................................................................... 929
46.1 Background and Overview ................................................................................... 929
46.2 Signal Model and Problem Formulation ............................................................. 931
46.3 Source Model-Based Speech Dereverberation ..................................................... 932
46.4 Separation of Speech and Reverberation via Homomorphic Transformation .... 936
46.5 Channel Inversion and Equalization .................................................................. 937
46.6 Summary ............................................................................................................. 941
References .................................................................................................................. 942
47 Adaptive Beamforming and Postfiltering
S. Gannot, I. Cohen .......................................................... 945
47.1 Problem Formulation .................................................. 947
47.2 Adaptive Beamforming ............................................. 948
47.3 Fixed Beamformer and Blocking Matrix ....................... 953
47.4 Identification of the Acoustical Transfer Function .......... 955
47.5 Robustness and Distortion Weighting ......................... 960
47.6 Multichannel Postfiltering ......................................... 962
47.7 Performance Analysis .............................................. 967
47.8 Experimental Results ............................................. 972
47.9 Summary .............................................................. 972
47.A Appendix: Derivation of the Expected Noise Reduction for a Coherent Noise Field ........................................ 973
47.B Appendix: Equivalence Between Maximum SNR and LCMV Beamformers ............................................... 974
References ........................................................................ 975

48 Feedback Control in Hearing Aids
A. Spriet, S. Doclo, M. Moonen, J. Wouters ....................... 979
48.1 Problem Statement .................................................... 980
48.2 Standard Adaptive Feedback Canceller ....................... 982
48.3 Feedback Cancellation Based on Prior Knowledge of the Acoustic Feedback Path ........................................ 986
48.4 Feedback Cancellation Based on Closed–Loop System Identification ......................................................... 990
48.5 Comparison ............................................................ 995
48.6 Conclusions ............................................................ 997
References ........................................................................ 997

49 Active Noise Control
S. M. Kuo, D. R. Morgan .................................................... 1001
49.1 Broadband Feedforward Active Noise Control .............. 1002
49.2 Narrowband Feedforward Active Noise Control ........... 1006
49.3 Feedback Active Noise Control .................................. 1010
49.4 Multichannel ANC ..................................................... 1011
49.5 Summary .............................................................. 1015
References ........................................................................ 1015

Part I Multichannel Speech Processing

50 Microphone Arrays
G. W. Elko, J. Meyer .......................................................... 1021
50.1 Microphone Array Beamforming ................................. 1021
50.2 Constant–Beamwidth Microphone Array System .......... 1029
50.3 Constrained Optimization of the Directional Gain ........ 1030
50.4 Differential Microphone Arrays .................................. 1031
50.5 Eigenbeamforming Arrays ......................................... 1034
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>50.6</td>
<td>Adaptive Array Systems</td>
<td>1037</td>
</tr>
<tr>
<td>50.7</td>
<td>Conclusions</td>
<td>1040</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>1040</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>51</th>
<th>Time Delay Estimation and Source Localization</th>
<th>1043</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Y. Huang, J. Benesty, J. Chen</td>
<td></td>
</tr>
<tr>
<td>51.1</td>
<td>Technology Taxonomy</td>
<td>1043</td>
</tr>
<tr>
<td>51.2</td>
<td>Time Delay Estimation</td>
<td>1044</td>
</tr>
<tr>
<td>51.3</td>
<td>Source Localization</td>
<td>1054</td>
</tr>
<tr>
<td>51.4</td>
<td>Summary</td>
<td>1061</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>1062</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>52</th>
<th>Convolutive Blind Source Separation Methods</th>
<th>1065</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M. S. Pedersen, J. Larsen, U. Kjems, L. C. Parra</td>
<td></td>
</tr>
<tr>
<td>52.1</td>
<td>The Mixing Model</td>
<td>1066</td>
</tr>
<tr>
<td>52.2</td>
<td>The Separation Model</td>
<td>1068</td>
</tr>
<tr>
<td>52.3</td>
<td>Identification</td>
<td>1071</td>
</tr>
<tr>
<td>52.4</td>
<td>Separation Principle</td>
<td>1071</td>
</tr>
<tr>
<td>52.5</td>
<td>Time Versus Frequency Domain</td>
<td>1076</td>
</tr>
<tr>
<td>52.6</td>
<td>The Permutation Ambiguity</td>
<td>1078</td>
</tr>
<tr>
<td>52.7</td>
<td>Results</td>
<td>1084</td>
</tr>
<tr>
<td>52.8</td>
<td>Conclusion</td>
<td>1084</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>1084</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>53</th>
<th>Sound Field Reproduction</th>
<th>1095</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R. Rabenstein, S. Spors</td>
<td></td>
</tr>
<tr>
<td>53.1</td>
<td>Sound Field Synthesis</td>
<td>1095</td>
</tr>
<tr>
<td>53.2</td>
<td>Mathematical Representation of Sound Fields</td>
<td>1096</td>
</tr>
<tr>
<td>53.3</td>
<td>Stereophony</td>
<td>1100</td>
</tr>
<tr>
<td>53.4</td>
<td>Vector-Based Amplitude Panning</td>
<td>1103</td>
</tr>
<tr>
<td>53.5</td>
<td>Ambisonics</td>
<td>1104</td>
</tr>
<tr>
<td>53.6</td>
<td>Wave Field Synthesis</td>
<td>1109</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>1113</td>
</tr>
</tbody>
</table>

Acknowledgements ........................................................................ 1115
About the Authors ..................................................................... 1117
Detailed Contents .................................................................... 1133
Subject Index ........................................................................... 1161