
Contents

| | |
|---|---|
| Abbreviations and Acronyms | 1 |
|---|---|

1 Introduction

E. Hänsler, G. Schmidt

| | |
|-----------------------------------|---|
| 1.1 Overview about the Book | 8 |
|-----------------------------------|---|

Part I Speech Enhancement

2 Low Delay Filter-Banks for Speech and Audio Processing

H. W. Löllmann, P. Vary

| | |
|---|----|
| 2.1 Introduction | 13 |
| 2.2 Analysis-Synthesis Filter-Banks | 15 |
| 2.2.1 General Structure | 15 |
| 2.2.2 Tree-Structured Filter-Banks | 16 |
| 2.2.3 Modulated Filter-Banks | 17 |
| 2.2.4 Frequency Warped Filter-Banks | 20 |
| 2.2.5 Low Delay Filter-Banks | 26 |
| 2.3 The Filter-Bank Equalizer | 29 |
| 2.3.1 Concept | 29 |
| 2.3.2 Prototype Filter Design | 31 |
| 2.3.3 Relation between GDFT and GDCT | 33 |
| 2.3.4 Realization for Different Filter Structures | 35 |
| 2.3.5 Polyphase Network Implementation | 37 |
| 2.3.6 The Non-Uniform Filter-Bank Equalizer | 41 |
| 2.3.7 Comparison between FBE and AS FB | 43 |
| 2.3.8 Algorithmic Complexity | 43 |
| 2.4 Further Measures for Signal Delay Reduction | 44 |
| 2.4.1 Concept | 45 |

VIII Contents

| | | |
|-------|---|----|
| 2.4.2 | Approximation by a Moving-Average Filter | 45 |
| 2.4.3 | Approximation by an Auto-Regressive Filter | 46 |
| 2.4.4 | Algorithmic Complexity | 47 |
| 2.4.5 | Warped Filter Approximation | 48 |
| 2.5 | Application to Noise Reduction | 49 |
| 2.5.1 | System Configurations | 49 |
| 2.5.2 | Instrumental Quality Measures | 50 |
| 2.5.3 | Simulation Results for the Uniform Filter-Banks | 51 |
| 2.5.4 | Simulation Results for the Warped Filter-Banks | 53 |
| 2.6 | Conclusions | 55 |
| | References | 56 |

3 A Pre-Filter for Hands-Free Car Phone Noise Reduction: Suppression of Harmonic Engine Noise Components

H. Puder

| | | |
|-------|--|----|
| 3.1 | Introduction | 63 |
| 3.2 | Analysis of the Different Car Noise Components | 64 |
| 3.2.1 | Wind Noise | 65 |
| 3.2.2 | Tire Noise | 65 |
| 3.2.3 | Engine Noise | 66 |
| 3.3 | Engine Noise Removal Based on Notch Filters | 68 |
| 3.4 | Compensation of Engine Harmonics with Adaptive Filters | 73 |
| 3.4.1 | Step-Size Control | 75 |
| 3.4.2 | Calculating the Optimal Step-Size | 78 |
| 3.4.3 | Results of the Compensation Approach | 80 |
| 3.5 | Evaluation and Comparison of the Results Obtained by the Notch Filter and the Compensation Approach | 84 |
| 3.6 | Conclusions and Summary | 85 |
| 3.6.1 | Conclusion | 85 |
| 3.6.2 | Summary | 86 |
| | References | 87 |

4 Model-Based Speech Enhancement

M. Krini, G. Schmidt

| | | |
|-------|---|-----|
| 4.1 | Introduction | 89 |
| 4.2 | Conventional Speech Enhancement Schemes | 91 |
| 4.3 | Speech Enhancement Schemes Based on Nonlinearities | 93 |
| 4.4 | Speech Enhancement Schemes Based on Speech Reconstruction | 97 |
| 4.4.1 | Feature Extraction and Control | 99 |
| 4.4.2 | Reconstruction of Speech Signals | 110 |
| 4.5 | Combining the Reconstructed and the Noise Suppressed Signal | 124 |
| 4.5.1 | Adding the Fully Reconstructed Signal | 125 |
| 4.5.2 | Adding only the Voiced Part of the Reconstructed Signal | 129 |

| | | |
|-----|---------------------------|-----|
| 4.6 | Summary and Outlook | 133 |
| | References | 133 |

5 Bandwidth Extension of Telephony Speech

B. Iser, G. Schmidt

| | | |
|-------|---|-----|
| 5.1 | Introduction | 135 |
| 5.2 | Organization of the Chapter | 137 |
| 5.3 | Basics | 138 |
| 5.3.1 | Human Speech Generation | 139 |
| 5.3.2 | Source-Filter Model | 141 |
| 5.3.3 | Parametric Representations of the Spectral Envelope | 143 |
| 5.3.4 | Distance Measures | 147 |
| 5.4 | Non-Model-Based Algorithms for Bandwidth Extension | 149 |
| 5.4.1 | Oversampling with Imaging | 149 |
| 5.4.2 | Spectral Shifting | 151 |
| 5.4.3 | Application of Non-Linear Characteristics | 153 |
| 5.5 | Model-Based Algorithms for Bandwidth Extension | 153 |
| 5.5.1 | Generation of the Excitation Signal | 155 |
| 5.5.2 | Vocal Tract Transfer Function Estimation | 159 |
| 5.6 | Evaluation of Bandwidth Extension Algorithms | 176 |
| 5.6.1 | Objective Distance Measures | 177 |
| 5.6.2 | Subjective Measures | 180 |
| 5.7 | Conclusions | 181 |
| | References | 182 |

6 Dereverberation and Residual Echo Suppression in Noisy Environments

E. A. P. Habets, S. Gannot, I. Cohen

| | | |
|-------|--|-----|
| 6.1 | Introduction | 186 |
| 6.2 | Problem Formulation | 188 |
| 6.3 | OM-LSA Estimator for Multiple Interferences | 191 |
| 6.3.1 | OM-LSA Estimator | 191 |
| 6.3.2 | A priori SIR Estimator | 193 |
| 6.4 | Dereverberation of Noisy Speech Signals | 195 |
| 6.4.1 | Short Introduction to Speech Dereverberation | 195 |
| 6.4.2 | Problem Formulation | 197 |
| 6.4.3 | Statistical Reverberation Model | 199 |
| 6.4.4 | Late Reverberant Spectral Variance Estimator | 200 |
| 6.4.5 | Summary and Discussion | 203 |
| 6.5 | Residual Echo Suppression | 203 |
| 6.5.1 | Problem Formulation | 204 |
| 6.5.2 | Late Residual Echo Spectral Variance Estimator | 206 |
| 6.5.3 | Parameter Estimation | 208 |
| 6.5.4 | Summary | 210 |

| | | |
|-------|---|-----|
| 6.6 | Joint Suppression of Reverberation, Residual Echo, and Noise . . . | 210 |
| 6.7 | Experimental Results | 212 |
| 6.7.1 | Experimental Setup | 214 |
| 6.7.2 | Joint Suppression of Reverberation and Noise | 214 |
| 6.7.3 | Suppression of Residual Echo | 216 |
| 6.7.4 | Joint Suppression of Reverberation, Residual Echo, and Noise | 221 |
| 6.8 | Summary and Outlook | 223 |
| | References | 224 |

7 Low Distortion Noise Cancellers – Revival of a Classical Technique

A. Sugiyama

| | | |
|-------|--|-----|
| 7.1 | Introduction | 229 |
| 7.2 | Distortions in Widrow’s Adaptive Noise Canceller | 230 |
| 7.2.1 | Distortion by Interference | 230 |
| 7.2.2 | Distortion by Crosstalk | 232 |
| 7.3 | Paired Filter (PF) Structure | 233 |
| 7.3.1 | Algorithm | 233 |
| 7.3.2 | Evaluations | 235 |
| 7.4 | Crosstalk Resistant ANC and Cross-Coupled Structure | 239 |
| 7.4.1 | Crosstalk Resistant ANC | 240 |
| 7.4.2 | Cross-Coupled Structure | 241 |
| 7.5 | Cross-Coupled Paired Filter (CCPF) Structure | 242 |
| 7.5.1 | Algorithm | 242 |
| 7.5.2 | Evaluations | 245 |
| 7.6 | Generalized Cross-Coupled Paired Filter (GCCPF) Structure | 247 |
| 7.6.1 | Algorithm | 250 |
| 7.6.2 | Evaluation by Recorded Signals | 251 |
| 7.7 | Demonstration in a Personal Robot | 261 |
| 7.8 | Conclusions | 261 |
| | References | 263 |

Part II Echo Cancellation

8 Nonlinear Echo Cancellation Based on Spectral Shaping

O. Hoshuyama, A. Sugiyama

| | | |
|-------|--|-----|
| 8.1 | Introduction | 267 |
| 8.2 | Frequency-Domain Model of Highly Nonlinear Residual Echo | 268 |
| 8.2.1 | Spectral Correlation Between Residual Echo and Echo Replica | 269 |

| | | |
|-------|---|-----|
| 8.2.2 | Model of Residual Echo Based on Spectral Correlation | 273 |
| 8.3 | Echo Canceller Based on the New Residual Echo Model | 274 |
| 8.3.1 | Overall Structure | 274 |
| 8.3.2 | Estimation of Near-End Speech | 275 |
| 8.3.3 | Spectral Gain Control | 276 |
| 8.4 | Evaluations | 277 |
| 8.4.1 | Objective Evaluations | 277 |
| 8.4.2 | Subjective Evaluation | 279 |
| 8.5 | DSP Implementation and Real-Time Evaluation | 280 |
| 8.6 | Conclusions | 280 |
| | References | 281 |

Part III Signal and System Quality Evaluation

9 Telephone-Speech Quality

U. Heute

| | | |
|-------|---|-----|
| 9.1 | Telephone-Speech Signals | 287 |
| 9.1.1 | Telephone Scenario | 287 |
| 9.1.2 | Telephone-Scenario Model | 287 |
| 9.2 | Speech-Signal Quality | 289 |
| 9.2.1 | Intelligibility | 289 |
| 9.2.2 | Speech-Sound Quality | 290 |
| 9.3 | Speech-Quality Assessment | 292 |
| 9.3.1 | Auditory Quality Assessment | 292 |
| 9.3.2 | Aims | 292 |
| 9.3.3 | Instrumental Quality Assessment | 293 |
| 9.4 | Compound-System Quality Prediction | 293 |
| 9.4.1 | The System-Planning Task | 293 |
| 9.4.2 | ETSI Network-Planning Model (E-Model) | 293 |
| 9.5 | Auditory Total-Quality Assessment | 294 |
| 9.5.1 | Conversation Tests | 294 |
| 9.5.2 | Listening Tests | 296 |
| 9.5.3 | LOTs with Pair Comparisons | 296 |
| 9.5.4 | Absolute-Category Rating (ACR) LOTs | 297 |
| 9.6 | Auditory Quality-Attribute Analysis | 298 |
| 9.6.1 | Quality Attributes | 298 |
| 9.6.2 | Attribute-Oriented LOTs | 298 |
| 9.6.3 | Search for Suitable Attributes | 302 |
| 9.6.4 | Integral-Quality Estimation from Attributes | 305 |
| 9.7 | Instrumental Total-Quality Measurement | 306 |
| 9.7.1 | Signal Comparisons | 306 |
| 9.7.2 | Evaluation Approaches | 306 |
| 9.7.3 | Psychoacoustically Motivated Measures | 312 |

XII Contents

| | | |
|-------|--|-----|
| 9.8 | Instrumental Attribute-Based Quality Measurements | 320 |
| 9.8.1 | Basic Ideas | 320 |
| 9.8.2 | Loudness | 322 |
| 9.8.3 | Sharpness | 323 |
| 9.8.4 | Roughness | 323 |
| 9.8.5 | Directness/Frequency Content (DFC) | 324 |
| 9.8.6 | Continuity | 326 |
| 9.8.7 | Noisiness | 329 |
| 9.8.8 | Combined Direct and Attribute-Based Total Quality Determination | 331 |
| 9.9 | Conclusions, Outlook, and Final Remarks | 331 |
| | References | 332 |

10 Evaluation of Hands-free Terminals

F. Kettler, H.-W. Gierlich

| | | |
|--------|--|-----|
| 10.1 | Introduction | 339 |
| 10.2 | Quality Assessment of Hands-free Terminals | 340 |
| 10.3 | Subjective Methods for Determining the Communicational Quality | 342 |
| 10.3.1 | General Setup and Opinion Scales Used for Subjective Performance Evaluation | 343 |
| 10.3.2 | Conversation Tests | 345 |
| 10.3.3 | Double Talk Tests | 346 |
| 10.3.4 | Talking and Listening Tests | 347 |
| 10.3.5 | Listening-only Tests (LOT) and Third Party Listening Tests | 348 |
| 10.3.6 | Experts Tests for Assessing Real Life Situations | 349 |
| 10.4 | Test Environment | 350 |
| 10.4.1 | The Acoustical Environment | 351 |
| 10.4.2 | Background Noise Simulation Techniques | 351 |
| 10.4.3 | Positioning of the Hands-Free Terminal | 352 |
| 10.4.4 | Positioning of the Artificial Head | 352 |
| 10.4.5 | Influence of the Transmission System | 354 |
| 10.5 | Test Signals and Analysis Methods | 354 |
| 10.5.1 | Speech and Perceptual Speech Quality Measures | 356 |
| 10.5.2 | Speech-like Test Signals | 356 |
| 10.5.3 | Background Noise | 360 |
| 10.5.4 | Applications | 363 |
| 10.6 | Result Representation | 365 |
| 10.6.1 | Interpretation of HFT “Quality Pies” | 366 |
| 10.6.2 | Examples | 368 |
| 10.7 | Related Aspects | 368 |
| 10.7.1 | The Lombard Effect | 368 |
| 10.7.2 | Intelligibility Outside Vehicles | 372 |

| | |
|------------------|-----|
| References | 375 |
|------------------|-----|

Part IV Multi-Channel Processing

11 Correlation-Based TDOA-Estimation for Multiple Sources in Reverberant Environments

J. Scheuing, B. Yang

| | | |
|------------|--|-----|
| 11.1 | Introduction | 381 |
| 11.2 | Analysis of TDOA Ambiguities | 383 |
| 11.2.1 | Signal Model | 383 |
| 11.2.2 | Multipath Ambiguity | 384 |
| 11.2.3 | Multiple Source Ambiguity | 384 |
| 11.2.4 | Ambiguity due to Periodic Signals | 386 |
| 11.2.5 | Principles of TDOA Disambiguation | 386 |
| 11.3 | Estimation of Direct Path TDOAs | 390 |
| 11.3.1 | Correlation and Extremum Positions | 390 |
| 11.3.2 | Raster Matching | 392 |
| 11.4 | Consistent TDOA Graphs | 397 |
| 11.4.1 | TDOA Graph | 397 |
| 11.4.2 | Strategies of Consistency Check | 398 |
| 11.4.3 | Properties of TDOA Graphs | 399 |
| 11.4.4 | Efficient Synthesis Algorithm | 402 |
| 11.4.5 | Initialization and Termination | 404 |
| 11.4.6 | Estimating the Number of Active Sources | 405 |
| 11.5 | Experimental Results | 406 |
| 11.5.1 | Localization System | 406 |
| 11.5.2 | TDOA Estimation of a Single Signal Block | 408 |
| 11.5.3 | Source Position Estimation | 412 |
| 11.5.4 | Evaluation of Continuous Measurements | 412 |
| 11.6 | Summary | 414 |
| References | | 415 |

12 Microphone Calibration for Multi-Channel Signal Processing

M. Buck, T. Haulick, H.-J. Pfeiderer

| | | |
|--------|--|-----|
| 12.1 | Introduction | 417 |
| 12.2 | Beamforming with Ideal Microphones | 418 |
| 12.2.1 | Principle of Beamforming | 418 |
| 12.2.2 | Evaluation of Beamformers | 421 |
| 12.2.3 | Statistically Optimum Beamformers | 424 |
| 12.3 | Microphone Mismatch and its Effect on Beamforming | 427 |
| 12.3.1 | Model for Non-Ideal Microphone Characteristics | 428 |
| 12.3.2 | Effect of Microphone Mismatch on Fixed Beamformers | 429 |

XIV Contents

| | | |
|------------|---|-----|
| 12.3.3 | Effect of Microphone Mismatch on Adaptive Beamformers | 430 |
| 12.3.4 | Comparison of Fixed and Adaptive Beamformers | 432 |
| 12.4 | Calibration Techniques and their Limits | |
| | for Real-World Applications | 432 |
| 12.4.1 | Calibration of Single Microphones | 432 |
| 12.4.2 | Analysis of Fixed Beamformers | 440 |
| 12.4.3 | Analysis of Adaptive Beamformers | 444 |
| 12.4.4 | Comparison of Fixed and Adaptive Beamformers | 448 |
| 12.5 | Self-Calibration Techniques | 449 |
| 12.5.1 | Basic Unit | 451 |
| 12.5.2 | Configurations for Array Processing | 452 |
| 12.5.3 | Recursive Configuration | 455 |
| 12.5.4 | Adaptation Control | 457 |
| 12.5.5 | Experimental Results | 458 |
| 12.6 | Summary | 459 |
| 12.A | Experimental Determination of the Directivity Index | 460 |
| 12.A.1 | Numerical Integration over a Spherical Surface | 460 |
| 12.A.2 | Definition of the Coordinate System | 462 |
| 12.A.3 | Determination of the Directivity | |
| | using the Normalized Cross Power Spectral | |
| | Densities of Microphone Signals | 464 |
| References | | 465 |

13 Convolutional Blind Source Separation for Noisy Mixtures

R. Aichner, H. Buchner, W. Kellermann

| | | |
|------------|--|-----|
| 13.1 | Introduction | 469 |
| 13.2 | Blind Source Separation for Acoustic Mixtures | |
| | Based on the TRINICON Framework | 473 |
| 13.2.1 | Matrix Formulation | 473 |
| 13.2.2 | Optimization Criterion and Coefficient Update | 474 |
| 13.2.3 | Approximations Leading to Special Cases | 478 |
| 13.2.4 | Estimation of the Correlation Matrices | |
| | and an Efficient Normalization Strategy | 484 |
| 13.2.5 | On Broadband and Narrowband BSS Algorithms | |
| | in the DFT Domain | 485 |
| 13.2.6 | Experimental Results for Reverberant Environments | 488 |
| 13.3 | Extensions for Blind Source Separation in Noisy Environments | 490 |
| 13.3.1 | Model for Background Noise in Realistic Environments | 491 |
| 13.3.2 | Pre-Processing for Noise-Robust Adaptation | 493 |
| 13.3.3 | Post-Processing for Suppression of Residual | |
| | Crosstalk and Background Noise | 498 |
| 13.4 | Conclusions | 518 |
| References | | 519 |

14 Binaural Speech Segregation

N. Roman, D.L. Wang

| | | |
|------|----------------------------------|-----|
| 14.1 | Introduction | 525 |
| 14.2 | T–F Masks for CASA | 528 |
| 14.3 | Anechoic Binaural Segregation | 529 |
| 14.4 | Reverberant Binaural Segregation | 533 |
| 14.5 | Evaluation | 536 |
| 14.6 | Concluding Remarks | 543 |
| 14.7 | Acknowledgments | 546 |
| | References | 546 |

**15 Spatio-Temporal Adaptive Inverse Filtering
in the Wave Domain**

S. Spors, H. Buchner, R. Rabenstein

| | | |
|--------|--|-----|
| 15.1 | Introduction | 551 |
| 15.2 | Problem Description | 553 |
| 15.2.1 | Nomenclature | 553 |
| 15.2.2 | Massive Multichannel Sound Reproduction | 553 |
| 15.2.3 | Multichannel Active Listening Room Compensation | 556 |
| 15.2.4 | Multichannel Active Noise Control | 559 |
| 15.2.5 | Unified Representation of Spatio-Temporal Adaptive Filtering Problems | 561 |
| 15.2.6 | Frequency-Domain Notation | 563 |
| 15.3 | Computation of Compensation Filters | 563 |
| 15.3.1 | Classification of Algorithms | 564 |
| 15.3.2 | Ideal Solution | 564 |
| 15.3.3 | Adaptive Solution | 565 |
| 15.3.4 | Problems of the Adaptive Solution | 567 |
| 15.4 | Eigenspace Adaptive Filtering | 568 |
| 15.4.1 | Generalized Singular Value Decomposition | 568 |
| 15.4.2 | Eigenspace Adaptive Filtering | 569 |
| 15.4.3 | Problems | 570 |
| 15.5 | Wave-Domain Adaptive Filtering | 570 |
| 15.5.1 | Concept | 571 |
| 15.5.2 | The Circular Harmonics Decomposition | 573 |
| 15.5.3 | The Circular Harmonics Expansion Using Boundary Measurements | 574 |
| 15.6 | Application of WDAF to Adaptive Inverse Filtering Problems | 576 |
| 15.6.1 | Application of WDAF to Active Listening Room Compensation | 577 |
| 15.6.2 | Application of WDAF to Active Noise Control | 578 |
| 15.7 | Conclusions | 578 |
| | References | 580 |

Part V Selected Applications

16 Virtual Hearing

K. Wiklund, S. Haykin

| | | |
|------|-----------------------------------|-----|
| 16.1 | Previous Work | 588 |
| 16.2 | VirtualHearing | 592 |
| 16.3 | Room Acoustic Model | 593 |
| 16.4 | HRTF Simulation | 597 |
| 16.5 | Neural Model | 600 |
| 16.6 | The Software and Interface | 605 |
| 16.7 | Software Testing | 609 |
| 16.8 | Future Work and Conclusions | 610 |
| | References | 613 |

**17 Dynamic Sound Control Algorithms
in Automobiles**

M. Christoph

| | | |
|--------|--|-----|
| 17.1 | Introduction | 615 |
| 17.1.1 | Introduction of Dynamic Volume Control Systems | 616 |
| 17.1.2 | Introduction of Dynamic Equalization Control Systems | 618 |
| 17.2 | Previous Systems – Description and Analysis | 619 |
| 17.2.1 | Speed Dependant Sound systems | 619 |
| 17.2.2 | Microphone Based Dynamic Volume Control Sound Systems | 621 |
| 17.2.3 | Non-Acoustic Sensor Based Sound Systems | 643 |
| 17.3 | Spectrum-Based Dynamic Equalization Control | 645 |
| 17.3.1 | Frequency Domain Adaptive Filter | 646 |
| 17.3.2 | Generalized Multidelay Adaptive Filter | 649 |
| 17.3.3 | Step-Size Control | 652 |
| 17.3.4 | Multi-Channel Systems | 653 |
| 17.3.5 | Estimating the Power Spectral Density of the Background Noise | 654 |
| 17.3.6 | Psychoacoustic Basics | 658 |
| 17.3.7 | The Psychoacoustic Masking Model According to Johnston | 661 |
| 17.4 | Conclusion and Outlook | 670 |
| 17.5 | Acknowledgement | 673 |
| | References | 674 |

**18 Towards Robust Distant-Talking Automatic Speech
Recognition in Reverberant Environments**

A. Sehr, W. Kellermann

| | | |
|------|--|-----|
| 18.1 | Introduction | 679 |
| 18.2 | The Distant-Talking ASR Scenario | 680 |

| | | |
|--------|---|-----|
| 18.3 | How to Deal with Reverberation in ASR Systems? | 683 |
| 18.4 | Effect of Reverberation in the Feature Domain..... | 691 |
| 18.5 | Signal Dereverberation and Beamforming | 695 |
| 18.6 | Robust Features | 699 |
| 18.7 | Model Training and Adaptation | 700 |
| 18.8 | Reverberation Modeling for Speech Recognition..... | 702 |
| 18.8.1 | Feature Production Model | 703 |
| 18.8.2 | Reverberation Model | 704 |
| 18.8.3 | Training of the Reverberation Model | 705 |
| 18.8.4 | Decoding | 708 |
| 18.8.5 | Inner Optimization | 713 |
| 18.8.6 | Solution of the Inner Optimization Problem in the Melspec Domain for Single Gaussian Densities | 714 |
| 18.8.7 | Simulations | 717 |
| 18.9 | Summary and Conclusions | 722 |
| | References | 723 |
| | Index | 729 |

Speech and Audio Processing in Adverse Environments

Hänsler, E.; Schmidt, G. (Eds.)

2008, XX, 736 p. 343 illus., Hardcover

ISBN: 978-3-540-70601-4