

1. Record Nr.	UNINA9910830736303321
Autore	Pulkki Ville
Titolo	Parametric time-frequency domain spatial audio / / edited by Ville Pulkki, Symeon Delikaris-Manias, Archontis Politis, Aalto University, Aalto, Finland
Pubbl/distr/stampa	Hoboken, New Jersey, USA : , : Wiley, , 2018 [Piscataway, New Jersey] : , : IEEE Xplore, , [2017]
ISBN	1-119-25261-X 1-119-25258-X 1-119-25263-6
Edizione	[First edition.]
Descrizione fisica	1 online resource (1 volume) : illustrations
Classificazione	TEC008000
Disciplina	621.382/2
Soggetti	Surround-sound systems - Mathematical models Time-domain analysis Signal processing
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references at the end of each chapters and index.
Nota di contenuto	Contents -- List of Contributors xiii -- Preface xv -- About the Companion Website xix -- Part I Analysis and Synthesis of Spatial Sound 1 -- Time–Frequency Processing: Methods and Tools 3 /Juha Vilkkamo and Tom Ba¨ckström -- 1.1 Introduction 3 -- 1.2 Time–Frequency Processing 4 -- 1.2.1 Basic Structure 4 -- 1.2.2 Uniform Filter Banks 5 -- 1.2.3 Prototype Filters and Modulation 6 -- 1.2.4 A Robust Complex-Modulated Filter Bank, and Comparison with STFT 8 -- 1.2.5 Overlap-Add and Windowing 12 -- 1.2.6 Example Implementation of a Robust Filter Bank in Matlab 13 -- 1.2.7 Cascaded Filters 15 -- 1.3 Processing of Spatial Audio 16 -- 1.3.1 Stochastic Estimates 17 -- 1.3.2 Decorrelation 18 -- 1.3.3 Optimal and Generalized Solution for Spatial Sound Processing Using Covariance Matrices 19 -- References 23 -- 2 Spatial Decomposition by Spherical Array Processing 25 /David Lou Alon and Boaz Rafaely -- 2.1 Introduction 25 -- 2.2 Sound Field Measurement by a Spherical Array 26 -- 2.3 Array Processing and Plane-Wave Decomposition 26 -- 2.4

Sensitivity to Noise and Standard Regularization Methods	29
Optimal Noise-Robust Design	32
Measure	32
Simulation Study	35
Performance Limit	37
Aliasing Cancellation	39
PWD Simulation Study	40
Example	42
Minimizing Broadband PWD Error	42
Summary	44
Acknowledgment	45
References	46
3 Sound Field Analysis Using Sparse Recovery	49
Introduction	49
The Plane-Wave Decomposition Problem	50
Sparse Plane-Wave Decomposition	51
The Iteratively Reweighted Least-Squares Algorithm	51
Bayesian Approach to Plane-Wave Decomposition	53
Calculating the IRLS Noise-Power Regularization Parameter	55
Estimation of the Relative Noise Power	56
Numerical Simulations	58
Experiment: Echoic Sound Scene Analysis	59
Conclusions	65
Appendix	65
References	66
Part II	
Reproduction of Spatial Sound	69
Overview of Time–Frequency Domain Parametric Spatial Audio Techniques	71
Introduction	71
Parametric Processing Overview	73
Analysis Principles	74
Synthesis Principles	75
Spatial Audio Coding and Up-Mixing	76
Spatial Sound Recording and Reproduction	78
Auralization of Measured Room Acoustics and Spatial Rendering of Room Impulse Responses	81
References	82
5 First-Order Directional Audio Coding (DirAC)	89
Introduction	89
Representing Spatial Sound with First-Order B-Format Signals	89
Some Notes on the Evolution of the Technique	92
DirAC with Ideal B-Format Signals	94
Analysis of Directional Parameters with Real Microphone Setups	97
DOA Analysis with Open 2D Microphone Arrays	97
DOA Analysis with 2D Arrays with a Rigid Baffle	99
DOA Analysis in Underdetermined Cases	101
DOA Analysis: Further Methods	102
Effect of Spatial Aliasing and Microphone Noise on the Analysis of Diffuseness	103
First-Order DirAC with Monophonic Audio Transmission	105
First-Order DirAC with Multichannel Audio Transmission	106
Stream-Based Virtual Microphone Rendering	106
Evaluation of Virtual Microphone DirAC	109
Discussion of Virtual Microphone DirAC	111
Optimized DirAC Synthesis	111
DirAC-Based Reproduction of Spaced-Array Recordings	114
DirAC Synthesis for Headphones and for Hearing Aids	117
Reproduction of B-Format Signals	117
DirAC in Hearing Aids	118
Optimizing the Time–Frequency Resolution of DirAC for Critical Signals	119
Example Implementation	120
Executing DirAC and Plotting Parameter History	122
DirAC Initialization	125
DirAC Runtime	131
Simplistic Binaural Synthesis of Loudspeaker Listening	136
Summary	137
References	138
6 Higher-Order Directional Audio Coding	141
Introduction	141
Sound Field Model	144
Energetic Analysis and Estimation of Parameters	145
Analysis of Intensity and Diffuseness in the Spherical Harmonic Domain	146
Higher-Order Energetic Analysis	147
Sector Profiles	149
Synthesis of Target Setup Signals	151

-- 6.4.1 Loudspeaker Rendering	152
-- 6.4.2 Binaural Rendering	155
-- 6.5 Subjective Evaluation	157
-- 6.6 Conclusions	157
-- References	158
7 Multi-Channel Sound Acquisition Using a Multi-Wave Sound Field Model	161
/Oliver Thiergart and Emanuele Habets	
-- 7.1 Introduction	161
-- 7.2 Parametric Sound Acquisition and Processing	163
-- 7.2.1 Problem Formulation	163
-- 7.2.2 Principal Estimation of the Target Signal	166
-- 7.3 Multi-Wave Sound Field and Signal Model	167
-- 7.3.1 Direct Sound Model	168
-- 7.3.2 Diffuse Sound Model	169
-- 7.3.3 Noise Model	169
-- 7.4 Direct and Diffuse Signal Estimation	170
-- 7.4.1 Estimation of the Direct Signal $Y_s(k, n)$	170
-- 7.4.2 Estimation of the Diffuse Signal $Y_d(k, n)$	176
-- 7.5 Parameter Estimation	179
-- 7.5.1 Estimation of the Number of Sources	179
-- 7.5.2 Direction of Arrival Estimation	181
-- 7.5.3 Microphone Input PSD Matrix	181
-- 7.5.4 Noise PSD Estimation	182
-- 7.5.5 Diffuse Sound PSD Estimation	182
-- 7.5.6 Signal PSD Estimation in Multi-Wave Scenarios	185
-- 7.6 Application to Spatial Sound Reproduction	186
-- 7.6.1 State of the Art	186
-- 7.6.2 Spatial Sound Reproduction Based on Informed Spatial Filtering	187
-- 7.7 Summary	194
-- References	195
8 Adaptive Mixing of Excessively Directive and Robust Beamformers for Reproduction of Spatial Sound	201
/Symeon Delikaris-Manias and Juha Vilkkamo	
8.1 Introduction	201
-- 8.2 Notation and Signal Model	202
-- 8.3 Overview of the Method	203
-- 8.4 Loudspeaker-Based Spatial Sound Reproduction	204
-- 8.4.1 Estimation of the Target Covariance Matrix C_y	204
-- 8.4.2 Estimation of the Synthesis Beamforming Signals W_s	206
-- 8.4.4 Processing the Synthesis Signals (W_{sx}) to Obtain the Target Covariance Matrix C_y	206
-- Spatial Energy Distribution	207
-- 8.4.5 Listening Tests	208
-- 8.5 Binaural-Based Spatial Sound Reproduction	209
-- 8.5.1 Estimation of the Analysis and Synthesis Beamforming Weight Matrices	210
-- 8.5.2 Diffuse-Field Equalization of HRTFs	210
-- 8.5.3 Adaptive Mixing and Decorrelation	211
-- 8.5.4 Subjective Evaluation	211
-- 8.6 Conclusions	212
-- References	212
9 Source Separation and Reconstruction of Spatial Audio Using Spectrogram Factorization	215
/Joonas Nikunen and Tuomas Virtanen	
-- 9.1 Introduction	215
-- 9.2 Spectrogram Factorization	217
-- 9.2.1 Mixtures of Sounds	217
-- 9.2.2 Magnitude Spectrogram Models	218
-- 9.2.3 Complex-Valued Spectrogram Models	221
-- 9.2.4 Source Separation by Time-Frequency Filtering	225
-- 9.3 Array Signal Processing and Spectrogram Factorization	226
-- 9.3.1 Spaced Microphone Arrays	226
-- 9.3.2 Model for Spatial Covariance Based on Direction of Arrival	227
-- 9.3.3 Complex-Valued NMF with the Spatial Covariance Model	229
-- 9.4 Applications of Spectrogram Factorization in Spatial Audio	231
-- 9.4.1 Parameterization of Surround Sound: Upmixing by Time-Frequency Filtering	231
-- 9.4.2 Source Separation Using a Compact Microphone Array	233
-- 9.4.3 Reconstruction of Binaural Sound Through Source Separation	238
-- 9.5 Discussion	243
-- 9.6 Matlab Example	243
-- References	247
Part III Signal-Dependent Spatial Filtering	251
10 Time-Frequency Domain Spatial Audio Enhancement	253
/Symeon Delikaris-Manias and Pasi Perttinen	
-- 10.1 Introduction	253
-- 10.2 Signal-Independent Enhancement	254
-- 10.3 Signal-Dependent Enhancement	255
-- 10.3.1 Adaptive Beamformers	255
-- 10.3.2 Post-Filters	257
-- 10.3.3 Post-Filter Types	257
-- 10.3.4 Estimating Post-Filters with Machine Learning	259
-- 10.3.5 Post-Filter Design Based on Spatial Parameters	259
-- References	261
11 Cross-Spectrum-Based Post-Filter Utilizing Noisy and Robust Beamformers	265
/Symeon Delikaris-Manias and Ville Pulkki	
-- 11.1 Introduction	265
-- 11.2 Notation and Signal Model	267
-- 11.2.1	

Virtual Microphone Design Utilizing Pressure Microphones	268	-- 11.3
Estimation of the Cross-Spectrum-Based Post-Filter	269	-- 11.3.1
Post-Filter Estimation Utilizing Two Static Beamformers	270	-- 11.3.2
Post-Filter Estimation Utilizing a Static and an Adaptive Beamformer	272	-- 11.3.3
Smoothing Techniques	277	-- 11.4
Implementation Examples	279	-- 11.4.1
Ideal Conditions	279	-- 11.4.2
Prototype Microphone Arrays	281	-- 11.5
Conclusions and Further Remarks	283	-- 11.6
Source Code	284	-- References
287	-- 12	Microphone-Array-Based Speech Enhancement Using Neural Networks /Pasi Pertila
291	-- 12.1	Introduction
291	-- 12.2	Time–Frequency Masks for Speech Enhancement Using Supervised Learning
293	-- 12.2.1	Beamforming with Post-Filtering
293	-- 12.2.2	Overview of Mask Prediction
294	-- 12.2.3	Features for Mask Learning
295	-- 12.2.4	Target Mask Design
297	-- 12.3	Artificial Neural Networks
298	-- 12.3.1	Learning the Weights
299	-- 12.3.2	Generalization
301	-- 12.3.3	Deep Neural Networks
305	-- 12.4	Mask Learning: A Simulated Example
305	-- 12.4.1	Feature Extraction
306	-- 12.4.2	Target Mask Design
306	-- 12.4.3	Neural Network Training
307	-- 12.4.4	Results
308	-- 12.5	Mask Learning: A Real-World Example
310	-- 12.5.1	Brief Description of the Third CHiME Challenge Data
310	-- 12.5.2	Data Processing and Beamforming
312	-- 12.5.3	Description of Network Structure, Features, and Targets
312	-- 12.5.4	Mask Prediction Results and Discussion
314	-- 12.5.5	Speech Enhancement Results
316	-- 12.6	Conclusions
318	-- 12.7	Source Code
318	-- 12.7.1	Matlab Code for Neural-Network-Based Sawtooth Denoising Example
318	-- 12.7.2	Matlab Code for Phase Feature Extraction
321	-- References	324
-- Part IV	Applications	327
-- 13	Upmixing and Beamforming in Professional Audio	329
-- 13.1	Introduction	329
-- 13.2	Stereo-to-Multichannel Upmix Processor	329
-- 13.2.1	Product Description	329
-- 13.2.2	Considerations for Professional Audio and Broadcast	331
-- 13.2.3	Signal Processing	332
-- 13.3	Digitally Enhanced Shotgun Microphone	336
-- 13.3.1	Product Description	336
-- 13.3.2	Concept	336
-- 13.3.3	Signal Processing	336
-- 13.3.4	Evaluations and Measurements	339
-- 13.4	Surround Microphone System Based on Two Microphone Elements	341
-- 13.4.1	Product Description	341
-- 13.4.2	Concept	344
-- 13.5	Summary	345
-- References	345	-- 14
Spatial Sound Scene Synthesis and Manipulation for Virtual Reality and Audio Effects	347	/Ville Pulkki, Archontis Politis, Tapani Pihlajama¨ki, and Mikko-Ville Laitinen
347	-- 14.1	Introduction
347	-- 14.2	Parametric Sound Scene Synthesis for Virtual Reality
348	-- 14.2.1	Overall Structure
348	-- 14.2.2	Synthesis of Virtual Sources
350	-- 14.2.3	Synthesis of Room Reverberation
352	-- 14.2.4	Augmentation of Virtual Reality with Real Spatial Recordings
352	-- 14.2.5	Higher-Order Processing
353	-- 14.2.6	Loudspeaker-Signal Bus
354	-- 14.3	Spatial Manipulation of Sound Scenes
355	-- 14.3.1	Parametric Directional Transformations
356	-- 14.3.2	Sweet-Spot Translation and Zooming
356	-- 14.3.3	Spatial Filtering
356	-- 14.3.4	Spatial Modulation
357	-- 14.3.5	Diffuse Field Level Control
358	-- 14.3.6	Ambience Extraction
359	-- 14.3.7	Spatialization of Monophonic Signals
360	-- 14.4	Summary
360	-- References	361
-- 15	Parametric Spatial Audio Techniques in Teleconferencing and Remote Presence	363
-- 15.1	Introduction and Motivation	363
-- 15.2	Background	365
-- 15.3	Immersive Audio Communication System (ImmACS)	366
-- 15.3.1	Encoder	366
-- 15.3.2	Decoder	373
-- 15.4	Capture and Reproduction of Crowded Acoustic Environments	376
15.4.1	Sound Source Positioning Based on VBAP	376
15.4.2	Non-	

Sommario/riassunto

"A comprehensive guide that addresses the theory and practice of spatial audio This book provides readers with the principles and best practices in spatial audio signal processing. It describes how sound fields and their perceptual attributes are captured and analyzed within the time-frequency domain, how essential representation parameters are coded, and how such signals are efficiently reproduced for practical applications. The book is split into four parts starting with an overview of the fundamentals. It then goes on to explain the reproduction of spatial sound before offering an examination of signal-dependent spatial filtering. The book finishes with coverage of both current and future applications and the direction that spatial audio research is heading in. Parametric Time-frequency Domain Spatial Audio focuses on applications in entertainment audio, including music, home cinema, and gaming—covering the capturing and reproduction of spatial sound as well as its generation, transduction, representation, transmission, and perception. This book will teach readers the tools needed for such processing, and provides an overview to existing research. It also shows recent up-to-date projects and commercial applications built on top of the systems. Provides an in-depth presentation of the principles, past developments, state-of-the-art methods, and future research directions of spatial audio technologies Includes contributions from leading researchers in the field Offers MATLAB codes with selected chapters An advanced book aimed at readers who are capable of digesting mathematical expressions about digital signal processing and sound field analysis, Parametric Time-frequency Domain Spatial Audio is best suited for researchers in academia and in the audio industry"--
