

# **2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2017)**

**New Paltz, New York, USA  
15 – 18 October 2017**



**IEEE Catalog Number: CFP17AUD-POD  
ISBN: 978-1-5386-1633-8**

**Copyright © 2017 by the Institute of Electrical and Electronics Engineers, Inc.  
All Rights Reserved**

*Copyright and Reprint Permissions:* Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

***\*\*\* This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

|                         |                   |
|-------------------------|-------------------|
| IEEE Catalog Number:    | CFP17AUD-POD      |
| ISBN (Print-On-Demand): | 978-1-5386-1633-8 |
| ISBN (Online):          | 978-1-5386-1632-1 |
| ISSN:                   | 1931-1168         |

**Additional Copies of This Publication Are Available From:**

Curran Associates, Inc  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: (845) 758-0400  
Fax: (845) 758-2633  
E-mail: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

CURRAN ASSOCIATES INC.  
**proceedings**  
.com

# TABLE OF CONTENTS

|                                                                                                                              |            |
|------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>METRIC LEARNING BASED DATA AUGMENTATION FOR ENVIRONMENTAL SOUND CLASSIFICATION.....</b>                                   | <b>1</b>   |
| <i>Rui Lu ; Zhiyao Duan ; Changshui Zhang</i>                                                                                |            |
| <b>TRANSFER LEARNING OF WEAKLY LABELLED AUDIO .....</b>                                                                      | <b>6</b>   |
| <i>Aleksandr Diment ; Tuomas Virtanen</i>                                                                                    |            |
| <b>SOUND EVENT DETECTION IN SYNTHETIC AUDIO: ANALYSIS OF THE DCASE 2016 TASK RESULTS.....</b>                                | <b>11</b>  |
| <i>Grégoire Lafay ; Emmanouil Benetos ; Mathieu Lagrange</i>                                                                 |            |
| <b>LEARNING VOCAL MODE CLASSIFIERS FROM HETEROGENEOUS DATA SOURCES .....</b>                                                 | <b>16</b>  |
| <i>Zhao Shuyang ; Toni Heittola ; Tuomas Virtanen</i>                                                                        |            |
| <b>MULTI-SCALE MULTI-BAND DENSENETS FOR AUDIO SOURCE SEPARATION .....</b>                                                    | <b>21</b>  |
| <i>Naoya Takahashi ; Yuki Mitsufuji</i>                                                                                      |            |
| <b>UNDERDETERMINED METHODS FOR MULTICHANNEL AUDIO ENHANCEMENT WITH PARTIAL PRESERVATION OF BACKGROUND SOURCES.....</b>       | <b>26</b>  |
| <i>Ryan M. Corey ; Andrew C. Singer</i>                                                                                      |            |
| <b>A CONVEX OPTIMIZATION APPROACH FOR TIME-FREQUENCY MASK ESTIMATION .....</b>                                               | <b>31</b>  |
| <i>Feng Bao ; Waleed H. Abdulla</i>                                                                                          |            |
| <b>MUSIC/VOICE SEPARATION USING THE 2D FOURIER TRANSFORM.....</b>                                                            | <b>36</b>  |
| <i>Prem Seetharaman ; Fatemeh Pishdadian ; Bryan Pardo</i>                                                                   |            |
| <b>EXPLOITING THE INTERMITTENCY OF SPEECH FOR JOINT SEPARATION AND DIARIZATION.....</b>                                      | <b>41</b>  |
| <i>Dionysos Kounades-Bastian ; Laurent Girin ; Xavier Alameda-Pineda ; Radu Horaud ; Sharon Gannot</i>                       |            |
| <b>A NOVEL TARGET SPEAKER DEPENDENT POSTFILTERING APPROACH FOR MULTICHANNEL SPEECH ENHANCEMENT .....</b>                     | <b>46</b>  |
| <i>Ritwik Giri ; Karim Helwani ; Tao Zhang</i>                                                                               |            |
| <b>EXPLAINING THE PARAMETERIZED WIENER FILTER WITH ALPHA-STABLE PROCESSES .....</b>                                          | <b>51</b>  |
| <i>Mathieu Fontaine ; Antoine Liutkus ; Laurent Girin ; Roland Badeau</i>                                                    |            |
| <b>AN EM ALGORITHM FOR AUDIO SOURCE SEPARATION BASED ON THE CONVOLUTIVE TRANSFER FUNCTION.....</b>                           | <b>56</b>  |
| <i>Xiaofei Li ; Laurent Girin ; Radu Horaud</i>                                                                              |            |
| <b>GUIDING AUDIO SOURCE SEPARATION BY VIDEO OBJECT INFORMATION .....</b>                                                     | <b>61</b>  |
| <i>Sanjeel Parekh ; Slim Essid ; Alexey Ozerov ; Ngoc Q. K. Duong ; Patrick Pérez ; Gaël Richard</i>                         |            |
| <b>LOW-LATENCY APPROXIMATION OF BIDIRECTIONAL RECURRENT NETWORKS FOR SPEECH DENOISING .....</b>                              | <b>66</b>  |
| <i>Gordon Wichern ; Alexey Lukin</i>                                                                                         |            |
| <b>LOW LATENCY SOUND SOURCE SEPARATION USING CONVOLUTIONAL RECURRENT NEURAL NETWORKS.....</b>                                | <b>71</b>  |
| <i>Gaurav Naithani ; Tom Barker ; Giambattista Parascandolo ; Lars Bramsløw ; Niels Henrik Pontoppidan ; Tuomas Virtanen</i> |            |
| <b>PSD ESTIMATION OF MULTIPLE SOUND SOURCES IN A REVERBERANT ROOM USING A SPHERICAL MICROPHONE ARRAY .....</b>               | <b>76</b>  |
| <i>Abdullah Fahim ; Prasanga N. Samarasinghe ; Thushara D. Abhayapala</i>                                                    |            |
| <b>JOINT WIDEBAND SOURCE LOCALIZATION AND ACQUISITION BASED ON A GRID-SHIFT APPROACH .....</b>                               | <b>81</b>  |
| <i>Christos Tzagkarakis ; W. Bastiaan Kleijn ; Jan Skoglund</i>                                                              |            |
| <b>EXPERIMENTAL STUDY OF ROBUST BEAMFORMING TECHNIQUES FOR ACOUSTIC APPLICATIONS .....</b>                                   | <b>86</b>  |
| <i>Yingke Zhao ; Jesper Rindom Jensen ; Mads Græsbøll Christensen ; Simon Doclo ; Jingdong Chen</i>                          |            |
| <b>MODULATION SPECTRUM BASED BEAMFORMING FOR SPEECH ENHANCEMENT .....</b>                                                    | <b>91</b>  |
| <i>Sam Karimian-Azari ; Tiago H. Falk</i>                                                                                    |            |
| <b>A CONFORMAL, HELMET-MOUNTED MICROPHONE ARRAY FOR AUDITORY SITUATIONAL AWARENESS AND HEARING PROTECTION .....</b>          | <b>96</b>  |
| <i>Paul Calamia ; Shakti Davis ; Christopher Smalt ; Christine Weston</i>                                                    |            |
| <b>MULTI-CHANNEL LATE REVERBERATION POWER SPECTRAL DENSITY ESTIMATION BASED ON NUCLEAR NORM MINIMIZATION .....</b>           | <b>101</b> |
| <i>Ina Kodrasi ; Simon Doclo</i>                                                                                             |            |

|                                                                                                                                                                      |            |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>DESIGN OF ROBUST TWO-DIMENSIONAL POLYNOMIAL BEAMFORMERS AS A CONVEX OPTIMIZATION PROBLEM WITH APPLICATION TO ROBOT AUDITION.....</b>                              | <b>106</b> |
| <i>Hendrik Barfuss ; Markus Bachmann ; Michael Buerger ; Martin Schneider ; Walter Kellermann</i>                                                                    |            |
| <b>MULTIZONE SOUND REPRODUCTION IN REVERBERANT ENVIRONMENTS USING AN ITERATIVE LEAST-SQUARES FILTER DESIGN METHOD WITH A SPATIOTEMPORAL WEIGHTING FUNCTION .....</b> | <b>111</b> |
| <i>Michael Buerger ; Christian Hofmann ; Cornelius Frankenbach ; Walter Kellermann</i>                                                                               |            |
| <b>AMPLITUDE ENGINEERING FOR BEAMFORMERS WITH SELF-BENDING DIRECTIVITY BASED ON CONVEX OPTIMIZATION .....</b>                                                        | <b>116</b> |
| <i>Jens Ahrens</i>                                                                                                                                                   |            |
| <b>DIRECTIONAL SOURCE MODELING IN WAVE-BASED ROOM ACOUSTICS SIMULATION .....</b>                                                                                     | <b>121</b> |
| <i>Stefan Bilbao ; Brian Hamilton</i>                                                                                                                                |            |
| <b>FREQUENCY DOMAIN SINGULAR VALUE DECOMPOSITION FOR EFFICIENT SPATIAL AUDIO CODING .....</b>                                                                        | <b>126</b> |
| <i>Sina Zamani ; Tejaswi Nanjundaswamy ; Kenneth Rose</i>                                                                                                            |            |
| <b>BLIND MICROPHONE GEOMETRY CALIBRATION USING ONE REVERBERANT SPEECH EVENT .....</b>                                                                                | <b>131</b> |
| <i>Ofer Schwartz ; Axel Plinge ; Emanuel A. P. Habets ; Sharon Gannot</i>                                                                                            |            |
| <b>BROADBAND DOA ESTIMATION USING CONVOLUTIONAL NEURAL NETWORKS TRAINED WITH NOISE SIGNALS.....</b>                                                                  | <b>136</b> |
| <i>Soumitro Chakrabarty ; Emanuel A. P. Habets</i>                                                                                                                   |            |
| <b>ANTIDERIVATIVE ANTIALIASING, LAGRANGE INTERPOLATION AND SPECTRAL FLATNESS .....</b>                                                                               | <b>141</b> |
| <i>Stefan Bilbao ; Fabian Esqueda ; Vesa Valimäki</i>                                                                                                                |            |
| <b>AN AUGMENTED LAGRANGIAN METHOD FOR PIANO TRANSCRIPTION USING EQUAL LOUDNESS THRESHOLDING AND LSTM-BASED DECODING .....</b>                                        | <b>146</b> |
| <i>Sebastian Ewert ; Mark B. Sandler</i>                                                                                                                             |            |
| <b>TOWARDS END-TO-END POLYPHONIC MUSIC TRANSCRIPTION: TRANSFORMING MUSIC AUDIO DIRECTLY TO A SCORE.....</b>                                                          | <b>151</b> |
| <i>Ralf Gunter Correa Carvalho ; Paris Smaragdis</i>                                                                                                                 |            |
| <b>A NOTE ON THE IMPLEMENTATION OF AUDIO PROCESSING BY SHORT-TERM FOURIER TRANSFORM .....</b>                                                                        | <b>156</b> |
| <i>James A. Moorer</i>                                                                                                                                               |            |
| <b>COLOURATION IN 2.5D LOCAL WAVE FIELD SYNTHESIS USING SPATIAL BANDWIDTH-LIMITATION .....</b>                                                                       | <b>160</b> |
| <i>Fiete Winter ; Christoph Hold ; Hagen Wierstorf ; Alexander Raake ; Sascha Spors</i>                                                                              |            |
| <b>DIFFERENTIAL MICROPHONE ARRAYS FOR THE UNDERWATER ACOUSTIC CHANNEL .....</b>                                                                                      | <b>165</b> |
| <i>Rotem Mulayoff ; Yaakov Buchris ; Israel Cohen</i>                                                                                                                |            |
| <b>LOCALIZATION OF ACOUSTIC SOURCES IN THE RAY SPACE FOR DISTRIBUTED MICROPHONE SENSORS.....</b>                                                                     | <b>170</b> |
| <i>Federico Borra ; Fabio Antonacci ; Augusto Sarti ; Stefano Tubaro</i>                                                                                             |            |
| <b>A PENALIZED INEQUALITY-CONSTRAINED MINIMUM VARIANCE BEAMFORMER WITH APPLICATIONS IN HEARING AIDS .....</b>                                                        | <b>175</b> |
| <i>Wenqiang Pu ; Jinjun Xiao ; Tao Zhang ; Zhi-Quan Luo</i>                                                                                                          |            |
| <b>ANGULAR SPECTRUM DECOMPOSITION-BASED 2.5D HIGHER-ORDER SPHERICAL HARMONIC SOUND FIELD SYNTHESIS WITH A LINEAR LOUDSPEAKER ARRAY .....</b>                         | <b>180</b> |
| <i>Takuma Okamoto</i>                                                                                                                                                |            |
| <b>EXTENDED SOUND FIELD RECORDING USING POSITION INFORMATION OF DIRECTIONAL SOUND SOURCES .....</b>                                                                  | <b>185</b> |
| <i>Keigo Wakayama ; Jorge Trevino ; Hideaki Takada ; Shuichi Sakamoto ; Yoiti Suzuki</i>                                                                             |            |
| <b>ASYMMETRIC BEAMPATTERNS WITH CIRCULAR DIFFERENTIAL MICROPHONE ARRAYS .....</b>                                                                                    | <b>190</b> |
| <i>Yaakov Buchris ; Israel Cohen ; Jacob Benesty</i>                                                                                                                 |            |
| <b>ROBUST PHASE REPLICATION METHOD FOR SPATIAL ALIASING PROBLEM IN MULTIPLE SOUND SOURCES LOCALIZATION .....</b>                                                     | <b>195</b> |
| <i>Kainan Chen ; Jürgen Geiger ; Wenyu Jin ; Mohammad Taghizadeh ; Walter Kellermann</i>                                                                             |            |
| <b>SPEECH ENHANCEMENT USING EXTREME LEARNING MACHINES .....</b>                                                                                                      | <b>200</b> |
| <i>Babafemi O. Odelowo ; David V Anderson</i>                                                                                                                        |            |
| <b>CONTINUOUS MEASUREMENT OF SPATIAL ROOM IMPULSE RESPONSES USING A NON-UNIFORMLY MOVING MICROPHONE .....</b>                                                        | <b>205</b> |
| <i>Nara Hahn ; Sascha Spors</i>                                                                                                                                      |            |

|                                                                                                                                                     |            |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| <b>INCOHERENT IDEMPOTENT AMBISONICS RENDERING .....</b>                                                                                             | <b>209</b> |
| <i>W. Bastiaan Kleijn ; Andrew Allen ; Jan Skoglund ; Felicia Lim</i>                                                                               |            |
| <b>COMPARISON OF REVERBERATION MODELS FOR SPARSE SOUND FIELD DECOMPOSITION .....</b>                                                                | <b>214</b> |
| <i>Shoichi Koyama ; Laurent Daudet</i>                                                                                                              |            |
| <b>A DNN REGRESSION APPROACH TO SPEECH ENHANCEMENT BY ARTIFICIAL BANDWIDTH EXTENSION .....</b>                                                      | <b>219</b> |
| <i>Johannes Abel ; Tim Fingscheidt</i>                                                                                                              |            |
| <b>COMPARING MODELED AND MEASUREMENT-BASED SPHERICAL HARMONIC ENCODING FILTERS FOR SPHERICAL MICROPHONE ARRAYS .....</b>                            | <b>224</b> |
| <i>Archontis Politis ; Hannes Gamper</i>                                                                                                            |            |
| <b>MULTI-MICROPHONE ACOUSTIC ECHO CANCELLATION USING RELATIVE ECHO TRANSFER FUNCTIONS .....</b>                                                     | <b>229</b> |
| <i>Maria Luis Valero ; Emanuel A. P. Habets</i>                                                                                                     |            |
| <b>NOISE POWER SPECTRAL DENSITY ESTIMATION FOR BINAURAL NOISE REDUCTION EXPLOITING DIRECTION OF ARRIVAL ESTIMATES .....</b>                         | <b>234</b> |
| <i>Daniel Marquardt ; Simon Doclo</i>                                                                                                               |            |
| <b>RAY SPACE ANALYSIS WITH SPARSE RECOVERY.....</b>                                                                                                 | <b>239</b> |
| <i>Craig Jin ; F. Antonacci ; A. Sarti</i>                                                                                                          |            |
| <b>DISTRIBUTED LCMV BEAMFORMING: CONSIDERATIONS OF SPATIAL TOPOLOGY AND LOCAL PREPROCESSING .....</b>                                               | <b>244</b> |
| <i>Dovid Y. Levin ; Shmulik Markovich-Golan ; Sharon Gannot</i>                                                                                     |            |
| <b>PERFORMANCE ANALYSIS OF A PLANAR MICROPHONE ARRAY FOR THREE DIMENSIONAL SOUNDFIELD ANALYSIS .....</b>                                            | <b>249</b> |
| <i>Prasanga N. Samarasinghe ; Hanchi Chen ; Abdullah Fahim ; Thushara D. Abhayapala</i>                                                             |            |
| <b>DEEP RECURRENT NMF FOR SPEECH SEPARATION BY UNFOLDING ITERATIVE THRESHOLDING .....</b>                                                           | <b>254</b> |
| <i>Scott Wisdom ; Thomas Powers ; James Pitton ; Les Atlas</i>                                                                                      |            |
| <b>LÉVY NMF FOR ROBUST NONNEGATIVE SOURCE SEPARATION .....</b>                                                                                      | <b>259</b> |
| <i>Paul Magron ; Roland Badeau ; Antoine Liutkus</i>                                                                                                |            |
| <b>SEPARATING TIME-FREQUENCY SOURCES FROM TIME-DOMAIN CONVOLUTIVE MIXTURES USING NON-NEGATIVE MATRIX FACTORIZATION.....</b>                         | <b>264</b> |
| <i>Simon Leglaive ; Roland Badeau ; Gael Richard</i>                                                                                                |            |
| <b>CONSISTENT ANISOTROPIC WIENER FILTERING FOR AUDIO SOURCE SEPARATION .....</b>                                                                    | <b>269</b> |
| <i>Paul Magron ; Jonathan Le Roux ; Tuomas Virtanen</i>                                                                                             |            |
| <b>PREDICTING ALGORITHM EFFICACY FOR ADAPTIVE MULTI-CUE SOURCE SEPARATION.....</b>                                                                  | <b>274</b> |
| <i>Ethan Manilow ; Prem Seetharaman ; Fatemeh Pishdadian ; Bryan Pardo</i>                                                                          |            |
| <b>THE SELECTION OF SPECTRAL MAGNITUDE EXPONENTS FOR SEPARATING TWO SOURCES IS DOMINATED BY PHASE DISTRIBUTION NOT MAGNITUDE DISTRIBUTION .....</b> | <b>279</b> |
| <i>Stephen Voran</i>                                                                                                                                |            |
| <b>LOW-COMPLEXITY KALMAN FILTER FOR MULTI-CHANNEL LINEAR-PREDICTION-BASED BLIND SPEECH DEREVERBERATION .....</b>                                    | <b>284</b> |
| <i>Thomas Dietzen ; Simon Doclo ; Ann Spriet ; Wouter Tirry ; Marc Moonen ; Toon van Waterschoot</i>                                                |            |
| <b>DYNAMIC RANGE COMPRESSION FOR NOISY MIXTURES USING SOURCE SEPARATION AND BEAMFORMING .....</b>                                                   | <b>289</b> |
| <i>Ryan M. Corey ; Andrew C. Singer</i>                                                                                                             |            |
| <b>AMPLITUDE AND PHASE DEREVERBERATION OF HARMONIC SIGNALS .....</b>                                                                                | <b>294</b> |
| <i>Arthur Belhomme ; Roland Badeau ; Yves Grenier ; Eric Humbert</i>                                                                                |            |
| <b>AUDIO SOFT DECLIPPING BASED ON WEIGHTED L<sub>1</sub>-NORM .....</b>                                                                             | <b>299</b> |
| <i>Flavio R. Avila ; Luiz W. P. Biscainho</i>                                                                                                       |            |
| <b>IMINET: CONVOLUTIONAL SEMI-SIAMESE NETWORKS FOR SOUND SEARCH BY VOCAL IMITATION .....</b>                                                        | <b>304</b> |
| <i>Yichi Zhang ; Zhiyao Duan</i>                                                                                                                    |            |
| <b>LEVERAGING REPETITION TO DO AUDIO IMPUTATION .....</b>                                                                                           | <b>309</b> |
| <i>Ethan Manilow ; Bryan Pardo</i>                                                                                                                  |            |
| <b>A KALMAN-BASED FUNDAMENTAL FREQUENCY ESTIMATION ALGORITHM .....</b>                                                                              | <b>314</b> |
| <i>Liming Shi ; Jesper K. Nielsen ; Jesper R. Jensen ; Max A. Little ; Mads G. Christensen</i>                                                      |            |
| <b>ASSESSMENT OF HUMAN AND MACHINE PERFORMANCE IN ACOUSTIC SCENE CLASSIFICATION: DCASE 2016 CASE STUDY .....</b>                                    | <b>319</b> |
| <i>Annamaria Mesaros ; Toni Heittola ; Tuomas Virtanen</i>                                                                                          |            |

|                                                                                                                             |            |
|-----------------------------------------------------------------------------------------------------------------------------|------------|
| <b>SPEECH CODING WITH TRANSFORM DOMAIN PREDICTION .....</b>                                                                 | <b>324</b> |
| <i>Lars Villemoes ; Janusz Klejsa ; Per Hedelin</i>                                                                         |            |
| <b>VOICE CONVERSION BASED ON A MIXTURE DENSITY NETWORK.....</b>                                                             | <b>329</b> |
| <i>Mohsen Ahangar ; Mostafa Ghorbandoost ; Sudhendu Sharma ; Mark J. T. Smith</i>                                           |            |
| <b>SOURCE RENDERING ON DYNAMIC AUDIO DISPLAYS .....</b>                                                                     | <b>334</b> |
| <i>Michael C. Heilemann ; David Anderson ; Mark F. Bocko</i>                                                                |            |
| <b>ZERO-DELAY LARGE SIGNAL CONVOLUTION USING MULTIPLE PROCESSOR<br/>ARCHITECTURES .....</b>                                 | <b>339</b> |
| <i>Nicholas Jillings ; Joshua D. Reiss ; Ryan Stables</i>                                                                   |            |
| <b>SCAPER: A LIBRARY FOR SOUNDSCAPE SYNTHESIS AND AUGMENTATION .....</b>                                                    | <b>344</b> |
| <i>Justin Salamon ; Duncan MacConnell ; Mark Cartwright ; Peter Li ; Juan Pablo Bello</i>                                   |            |
| <b>TRANSIENT-TO-NOISE RATIO RESTORATION OF CODED APPLAUSE-LIKE SIGNALS .....</b>                                            | <b>349</b> |
| <i>Alexander Adami ; Adrian Herzog ; Sascha Disch ; Jürgen Herre</i>                                                        |            |
| <b>DIAGONAL RNNs IN SYMBOLIC MUSIC MODELING .....</b>                                                                       | <b>354</b> |
| <i>Y. Cem Subakan ; Paris Smaragdis</i>                                                                                     |            |
| <b>DEEP RECURRENT MIXTURE OF EXPERTS FOR SPEECH ENHANCEMENT .....</b>                                                       | <b>359</b> |
| <i>Shlomo E. Chazan ; Jacob Goldberger ; Sharon Gannot</i>                                                                  |            |
| <b>FAST RECONSTRUCTION OF SPARSE RELATIVE IMPULSE RESPONSES VIA SECOND-<br/>ORDER CONE PROGRAMMING .....</b>                | <b>364</b> |
| <i>Pavel Rajmic ; Zbyněk Koldovský ; Marie Daňková</i>                                                                      |            |
| <b>QRD BASED MVDR BEAMFORMING FOR FAST TRACKING OF SPEECH AND NOISE<br/>DYNAMICS .....</b>                                  | <b>369</b> |
| <i>Anna Barnov ; Vered Bar Bracha ; Shmulik Markovich-Golan</i>                                                             |            |
| <b>AUTOMATED AUDIO CAPTIONING WITH RECURRENT NEURAL NETWORKS.....</b>                                                       | <b>374</b> |
| <i>Konstantinos Drossos ; Sharath Adavanne ; Tuomas Virtanen</i>                                                            |            |
| <b>ENHANCEMENT OF AMBISONIC BINAURAL REPRODUCTION USING DIRECTIONAL<br/>AUDIO CODING WITH OPTIMAL ADAPTIVE MIXING .....</b> | <b>379</b> |
| <i>Archontis Politis ; Leo McCormack ; Ville Pulkki</i>                                                                     |            |
| <b>OPTIMIZING DIFFERENTIATED DISCRETIZATION FOR AUDIO CIRCUITS BEYOND<br/>DRIVING POINT TRANSFER FUNCTIONS .....</b>        | <b>384</b> |
| <i>François G. Germain ; Kurt James Werner</i>                                                                              |            |
| <b>Author Index</b>                                                                                                         |            |