

# **2019 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2019)**

**New Paltz, New York, USA  
20 – 23 October 2019**



**IEEE Catalog Number: CFP19AUD-POD  
ISBN: 978-1-7281-1124-7**

**Copyright © 2019 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:	CFP19AUD-POD
ISBN (Print-On-Demand):	978-1-7281-1124-7
ISBN (Online):	978-1-7281-1123-0
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>SOUND EVENT DETECTION USING POINT-LABELED DATA</b> .....	1
<i>Bongjun Kim ; Bryan Pardo</i>	
<b>BATCH UNIFORMIZATION FOR MINIMIZING MAXIMUM ANOMALY SCORE OF DNN-BASED ANOMALY DETECTION IN SOUNDS</b> .....	6
<i>Yuma Koizumi ; Shoichiro Saito ; Masataka Yamaguchi ; Shin Murata ; Noboru Harada</i>	
<b>CITY CLASSIFICATION FROM MULTIPLE REAL-WORLD SOUND SCENES</b> .....	11
<i>Helen L. Bear ; Toni Heittola ; Annamaria Mesaros ; Emmanouil Benetos ; Tuomas Virtanen</i>	
<b>MODEL-AGNOSTIC APPROACHES TO HANDLING NOISY LABELS WHEN TRAINING SOUND EVENT CLASSIFIERS</b> .....	16
<i>Eduardo Fonseca ; Frederic Font ; Xavier Serra</i>	
<b>ANNOTATIONS TIME SHIFT: A KEY PARAMETER IN EVALUATING MUSICAL NOTE ONSET DETECTION ALGORITHMS</b> .....	21
<i>Mina Mounir ; Peter Karsmakers ; Toon Van Waterschoot</i>	
<b>END-TO-END MELODY NOTE TRANSCRIPTION BASED ON A BEAT-SYNCHRONOUS ATTENTION MECHANISM</b> .....	26
<i>Ryo Nishikimi ; Eita Nakamura ; Masataka Goto ; Kazuyoshi Yoshii</i>	
<b>TIME-SCALE MODIFICATION USING FUZZY EPOCH-SYNCHRONOUS OVERLAP-ADD (FESOLA)</b> .....	31
<i>Timothy Roberts ; Kuldip K. Paliwal</i>	
<b>HIGH-LEVEL CONTROL OF DRUM TRACK GENERATION USING LEARNED PATTERNS OF RHYTHMIC INTERACTION</b> .....	35
<i>Stefan Lattner ; Maarten Grachten</i>	
<b>INVESTIGATING KERNEL SHAPES AND SKIP CONNECTIONS FOR DEEP LEARNING-BASED HARMONIC-PERCUSSIVE SEPARATION</b> .....	40
<i>Carlos Lordelo ; Emmanouil Benetos ; Simon Dixon ; Sven Ahlbäck</i>	
<b>CUTTING MUSIC SOURCE SEPARATION SOME SLAKH: A DATASET TO STUDY THE IMPACT OF TRAINING DATA QUALITY AND QUANTITY</b> .....	45
<i>Ethan Manilow ; Gordon Wichern ; Prem Seetharaman ; Jonathan Le Roux</i>	
<b>ON THE BEHAVIOR OF DELAY NETWORK REVERBERATOR MODES</b> .....	50
<i>Orchisama Das ; Elliot K. Canfield-Dafilou ; Jonathan S. Abel</i>	
<b>GRAPHIC EQUALIZER DESIGN WITH SYMMETRIC BIQUAD FILTERS</b> .....	55
<i>Juho Liski ; Jussi Rämö ; Vesa Välimäki</i>	
<b>ACTIVE FEEDBACK SUPPRESSION FOR HEARING DEVICES EXPLOITING MULTIPLE LOUDSPEAKERS</b> .....	60
<i>Henning Schepker ; Simon Doclo</i>	
<b>PERCEPTUAL EVALUATION OF BINAURAL AURALIZATION OF DATA OBTAINED FROM THE SPATIAL DECOMPOSITION METHOD</b> .....	65
<i>Jens Ahrens</i>	
<b>SPARSE REPRESENTATION OF HRTFS BY EAR ALIGNMENT</b> .....	70
<i>Zamir Ben-Hur ; David Lou Alon ; Ravish Mehra ; Boaz Rafaely</i>	
<b>MORPHOLOGICAL WEIGHTING IMPROVES INDIVIDUALIZED PREDICTION OF HRTF DIRECTIVITY PATTERNS</b> .....	75
<i>Muhammad Shahnawaz ; Craig Jin ; Joan Glaunès ; Augusto Sarti ; Anthony Tew</i>	
<b>EEG-BASED DECODING OF AUDITORY ATTENTION TO A TARGET INSTRUMENT IN POLYPHONIC MUSIC</b> .....	80
<i>Giorgia Cantisani ; Slim Essid ; Gaël Richard</i>	
<b>INTRUSIVE AND NON-INTRUSIVE PERCEPTUAL SPEECH QUALITY ASSESSMENT USING A CONVOLUTIONAL NEURAL NETWORK</b> .....	85
<i>Hannes Gamper ; Chandan K. A. Reddy ; Ross Cutler ; Ivan J. Tashev ; Johannes Gehrke</i>	
<b>AN IMPROVED MEASURE OF MUSICAL NOISE BASED ON SPECTRAL KURTOSIS</b> .....	90
<i>Matteo Torcoli</i>	
<b>AN EFFICIENT MODEL FOR ESTIMATING SUBJECTIVE QUALITY OF SEPARATED AUDIO SOURCE SIGNALS</b> .....	95
<i>Thorsten Kastner ; Jürgen Herre</i>	
<b>A CLASSIFICATION-AIDED FRAMEWORK FOR NON-INTRUSIVE SPEECH QUALITY ASSESSMENT</b> .....	100
<i>Xuan Dong ; Donald S. Williamson</i>	

<b>IDENTIFICATION OF VOICE QUALITY VARIATION USING I-VECTORS</b> .....	105
<i>Chuyao Feng ; Eva Van Leer ; David V. Anderson</i>	
<b>3D LOCALIZED SOUND ZONE GENERATION WITH A PLANAR OMNI-DIRECTIONAL LOUDSPEAKER ARRAY</b> .....	110
<i>Takuma Okamoto</i>	
<b>MOTION-TOLERANT BEAMFORMING WITH DEFORMABLE MICROPHONE ARRAYS</b> .....	115
<i>Ryan M. Corey ; Andrew C. Singer</i>	
<b>AN EM METHOD FOR MULTICHANNEL TOA AND DOA ESTIMATION OF ACOUSTIC ECHOES</b> .....	120
<i>Jesper Rindom Jensen ; Usama Saqib ; Sharon Gannot</i>	
<b>SPEECH ENHANCEMENT USING POLYNOMIAL EIGENVALUE DECOMPOSITION</b> .....	125
<i>Vincent W. Neo ; Christine Evers ; Patrick A. Naylor</i>	
<b>SUB-SAMPLE TIME DELAY ESTIMATION VIA AUXILIARY-FUNCTION-BASED ITERATIVE UPDATES</b> .....	130
<i>Kouei Yamaoka ; Robin Scheibler ; Nobutaka Ono ; Yukoh Wakabayashi</i>	
<b>ACTIVE NOISE CONTROL OVER 3D SPACE WITH MULTIPLE CIRCULAR ARRAYS</b> .....	135
<i>Huiyuan Sun ; Thushara D. Abhayapala ; Prasanga N. Samarasinghe</i>	
<b>SOUND FIELD TRANSLATION METHODS FOR BINAURAL REPRODUCTION</b> .....	140
<i>Lachlan Birnie ; Thushara Abhayapala ; Prasanga Samarasinghe ; Vladimir Tourbabin</i>	
<b>FEEDBACK STRUCTURES FOR A TRANSFER FUNCTION MODEL OF A CIRCULAR VIBRATING MEMBRANE</b> .....	145
<i>Maximilian Schäfer ; Rudolf Rabenstein ; Sebastian J. Schlecht</i>	
<b>DENSE REVERBERATION WITH DELAY FEEDBACK MATRICES</b> .....	150
<i>Sebastian J. Schlecht ; Emanuël A. P. Habets</i>	
<b>PHYSICAL MODELS FOR FAST ESTIMATION OF GUITAR STRING, FRET AND PLUCKING POSITION</b> .....	155
<i>Jacob Møller Hjerrild ; Silvin Willemsen ; Mads Græsbøll Christensen</i>	
<b>JOINT SINGING PITCH ESTIMATION AND VOICE SEPARATION BASED ON A NEURAL HARMONIC STRUCTURE RENDERER</b> .....	160
<i>Tomoyasu Nakano ; Kazuyoshi Yoshii ; Yiming Wu ; Ryo Nishikimi ; Kin Wah Edward Lin ; Masataka Goto</i>	
<b>ANALYSIS OF ROBUSTNESS OF DEEP SINGLE-CHANNEL SPEECH SEPARATION USING CORPORA CONSTRUCTED FROM MULTIPLE DOMAINS</b> .....	165
<i>Matthew Maciejewski ; Gregory Sell ; Yusuke Fujita ; Leibny Paola Garcia-Perera ; Shinji Watanabe ; Sanjeev Khudanpur</i>	
<b>A STYLE TRANSFER APPROACH TO SOURCE SEPARATION</b> .....	170
<i>Shrikant Venkataramani ; Efthymios Tzinis ; Paris Smaragdis</i>	
<b>UNIVERSAL SOUND SEPARATION</b> .....	175
<i>Ilya Kavaleroov ; Scott Wisdom ; Hakan Erdogan ; Brian Patton ; Kevin Wilson ; Jonathan Le Roux ; John R. Hershey</i>	
<b>DEEP TENSOR FACTORIZATION FOR SPATIALLY-AWARE SCENE DECOMPOSITION</b> .....	180
<i>Jonah Casebeer ; Michael Colomb ; Paris Smaragdis</i>	
<b>INDEPENDENT VECTOR ANALYSIS WITH MORE MICROPHONES THAN SOURCES</b> .....	185
<i>Robin Scheibler ; Nobutaka Ono</i>	
<b>SPARSE ADAPTATION OF DISTRIBUTED BLIND SOURCE SEPARATION IN ACOUSTIC SENSOR NETWORKS</b> .....	190
<i>Michael Günther ; Haitham Afifi ; Andreas Brendel ; Holger Karl ; Walter Kellermann</i>	
<b>MULTIPLE HYPOTHESIS TRACKING FOR OVERLAPPING SPEAKER SEGMENTATION</b> .....	195
<i>Aidan O. T. Hogg ; Christine Evers ; Patrick A. Naylor</i>	
<b>DECLIPPING SPEECH USING DEEP FILTERING</b> .....	200
<i>Wolfgang Mack ; Emanuël A. P. Habets</i>	
<b>SPEECH BANDWIDTH EXTENSION WITH WAVENET</b> .....	205
<i>Archit Gupta ; Brendan Shillingford ; Yannis Assael ; Thomas C. Walters</i>	
<b>IRM WITH PHASE PARAMETERIZATION FOR SPEECH ENHANCEMENT</b> .....	209
<i>Xianyun Wang ; Changchun Bao ; Rui Cheng</i>	
<b>GENERATIVE SPEECH ENHANCEMENT BASED ON CLONED NETWORKS</b> .....	214
<i>Michael Chinen ; W. Bastiaan Kleijn ; Felicia S. C. Lim ; Jan Skoglund</i>	
<b>IMPROVEMENT OF SPEECH RESIDUALS FOR SPEECH ENHANCEMENT</b> .....	219
<i>Samy Elshamy ; Tim Fingscheidt</i>	
<b>SIMULTANEOUS DENOISING, DEREVERBERATION, AND SOURCE SEPARATION USING A UNIFIED CONVOLUTIONAL BEAMFORMER</b> .....	224
<i>Tomohiro Nakatani ; Keisuke Kinoshita ; Rintaro Ikeshita ; Hiroshi Sawada ; Shoko Araki</i>	

<b>A PERCEPTUAL WEIGHTING FILTER LOSS FOR DNN TRAINING IN SPEECH ENHANCEMENT</b> .....	229
<i>Ziyue Zhao ; Samy Elshamy ; Tim Fingscheidt</i>	
<b>SPEECH ENHANCEMENT USING END-TO-END SPEECH RECOGNITION OBJECTIVES</b> .....	234
<i>Aswin Shanmugam Subramanian ; Xiaofei Wang ; Murali Karthick Baskar ; Shinji Watanabe ; Toru Taniguchi ; Dung Tran ; Yuya Fujita</i>	
<b>SEPARATED NOISE SUPPRESSION AND SPEECH RESTORATION: LSTM-BASED SPEECH ENHANCEMENT IN TWO STAGES</b> .....	239
<i>Maximilian Strake ; Bruno Defraene ; Kristoff Fluyt ; Wouter Tirry ; Tim Fingscheidt</i>	
<b>FAST CONVERGENCE ALGORITHM FOR STATE-SPACE MODEL BASED SPEECH DEREVERBERATION BY MULTI-CHANNEL NON-NEGATIVE MATRIX FACTORIZATION</b> .....	244
<i>Masahito Togami ; Tatsuya Komatsu</i>	
<b>ATTENTION WAVE-U-NET FOR SPEECH ENHANCEMENT</b> .....	249
<i>Ritwik Giri ; Umut Isik ; Arvinth Krishnaswamy</i>	
<b>DILATED FCN: LISTENING LONGER TO HEAR BETTER</b> .....	254
<i>Shuyu Gong ; Zhewei Wang ; Tao Sun ; Yuanhang Zhang ; Charles D. Smith ; Li Xu ; Jundong Liu</i>	
<b>UNSUPERVISED ADVERSARIAL DOMAIN ADAPTATION BASED ON THE WASSERSTEIN DISTANCE FOR ACOUSTIC SCENE CLASSIFICATION</b> .....	259
<i>Konstantinos Drossos ; Paul Magron ; Tuomas Virtanen</i>	
<b>ZERO-SHOT AUDIO CLASSIFICATION BASED ON CLASS LABEL EMBEDDINGS</b> .....	264
<i>Huang Xie ; Tuomas Virtanen</i>	
<b>IDENTIFY, LOCATE AND SEPARATE: AUDIO-VISUAL OBJECT EXTRACTION IN LARGE VIDEO COLLECTIONS USING WEAK SUPERVISION</b> .....	268
<i>Sanjeel Parekh ; Alexey Ozerov ; Slim Essid ; Ngoc Q. K. Duong ; Patrick Pérez ; Gaël Richard</i>	
<b>WEAKLY INFORMED AUDIO SOURCE SEPARATION</b> .....	273
<i>Kilian Schulze-Forster ; Clément Doire ; Gaël Richard ; Roland Badeau</i>	
<b>TRICYCLE: AUDIO REPRESENTATION LEARNING FROM SENSOR NETWORK DATA USING SELF-SUPERVISION</b> .....	278
<i>Mark Cartwright ; Jason Cramer ; Justin Salamon ; Juan Pablo Bello</i>	
<b>DEEP RANKING-BASED SOUND SOURCE LOCALIZATION</b> .....	283
<i>Renana Opoehinsky ; Bracha Laufer-Goldshtein ; Sharon Gannot ; Gal Chechik</i>	
<b>INDEPENDENT LOW-RANK MATRIX ANALYSIS WITH DECORRELATION LEARNING</b> .....	288
<i>Rintaro Ikeshita ; Nobutaka Ito ; Tomohiro Nakatani ; Hiroshi Sawada</i>	
<b>GENERALIZED WEIGHTED-PREDICTION-ERROR DEREVERBERATION WITH VARYING SOURCE PRIORS FOR REVERBERANT SPEECH RECOGNITION</b> .....	293
<i>Toru Taniguchi ; Aswin Shanmugam Subramanian ; Xiaofei Wang ; Dung Tran ; Yuya Fujita ; Shinji Watanabe</i>	
<b>MULTICHANNEL SPEECH ENHANCEMENT BASED ON TIME-FREQUENCY MASKING USING SUBBAND LONG SHORT-TERM MEMORY</b> .....	298
<i>Xiaofei Li ; Radu Horaud</i>	
<b>PARAMETRIC RESYNTHESIS WITH NEURAL VOCODERS</b> .....	303
<i>Soumi Maiti ; Michael I. Mandel</i>	
<b>CONTINUAL LEARNING OF NEW SOUND CLASSES USING GENERATIVE REPLAY</b> .....	308
<i>Zhepei Wang ; Cem Subakan ; Efthymios Tzinis ; Paris Smaragdis ; Laurent Charlin</i>	
<b>TOYADMOS: A DATASET OF MINIATURE-MACHINE OPERATING SOUNDS FOR ANOMALOUS SOUND DETECTION</b> .....	313
<i>Yuma Koizumi ; Shoichiro Saito ; Hisashi Uematsu ; Noboru Harada ; Keisuke Imoto</i>	
<b>EVALUATION OF POST-PROCESSING ALGORITHMS FOR POLYPHONIC SOUND EVENT DETECTION</b> .....	318
<i>Léo Cancès ; Patrice Guyot ; Thomas Pellegrini</i>	
<b>POLYPHONIC SOUND EVENT AND SOUND ACTIVITY DETECTION: A MULTI-TASK APPROACH</b> .....	323
<i>Arjun Pankajakshan ; Helen L. Bear ; Emmanouil Benetos</i>	
<b>ACOUSTIC SCENE CLASSIFICATION USING HIGHER-ORDER AMBISONIC FEATURES</b> .....	328
<i>Marc C. Green ; Sharath Adavanne ; Damian Murphy ; Tuomas Virtanen</i>	
<b>JOINT MEASUREMENT OF LOCALIZATION AND DETECTION OF SOUND EVENTS</b> .....	333
<i>Annamaria Mesaros ; Sharath Adavanne ; Archontis Politis ; Toni Heittola ; Tuomas Virtanen</i>	
<b>JOINT ANALYSIS OF ACOUSTIC EVENTS AND SCENES BASED ON MULTITASK LEARNING</b> .....	338
<i>Noriyuki Tonami ; Keisuke Imoto ; Masahiro Niitsuma ; Ryosuke Yamanishi ; Yoichi Yamashita</i>	
<b>REGRESSION VERSUS CLASSIFICATION FOR NEURAL NETWORK BASED AUDIO SOURCE LOCALIZATION</b> .....	343
<i>Lauréline Perotin ; Alexandre Défossez ; Emmanuel Vincent ; Romain Serizel ; Alexandre Guérin</i>	

<b>SOUND SOURCE LOCALIZATION USING RELATIVE HARMONIC COEFFICIENTS IN MODAL DOMAIN</b> .....	348
<i>Yonggang Hu ; Prasanga N. Samarasinghe ; Thushara D. Abhayapala</i>	
<b>ACOUSTIC LOCALIZATION USING SPATIAL PROBABILITY IN NOISY AND REVERBERANT ENVIRONMENTS</b> .....	353
<i>Sebastian Braun ; Ivan Tashev</i>	
<b>SUPERVISED CONTRASTIVE EMBEDDINGS FOR BINAURAL SOURCE LOCALIZATION</b> .....	358
<i>Duowei Tang ; Maja Taseska ; Toon Van Waterschoot</i>	
<b>IMPROVED CHANGE PREDICTION FOR COMBINED BEAMFORMING AND ECHO CANCELLATION WITH APPLICATION TO A GENERALIZED SIDELobe CANCELER</b> .....	363
<i>Stefan Kühl ; Alexander Bohlender ; Matthias Schrammen ; Peter Jax</i>	
<b>TWO-DIMENSIONAL SOUND FIELD RECORDING WITH MULTIPLE CIRCULAR MICROPHONE ARRAYS CONSIDERING MULTIPLE SCATTERING</b> .....	368
<i>Masahiro Nakanishi ; Natsuki Ueno ; Shoichi Koyama ; Hiroshi Saruwatari</i>	
<b>RTF-STEERED BINAURAL MVDR BEAMFORMING INCORPORATING MULTIPLE EXTERNAL MICROPHONES</b> .....	373
<i>Nico Gößling ; Wiebke Middelberg ; Simon Doclo</i>	
<b>1ST-ORDER MICROPHONE ARRAY SYSTEM FOR LARGE AREA SOUND FIELD RECORDING AND RECONSTRUCTION: DISCUSSION AND PRELIMINARY RESULTS</b> .....	378
<i>Federico Borra ; Steven Krenn ; Israel Dejene Gebru ; Dejan Markovic</i>	
<b>DIRECTION OF ARRIVAL ESTIMATION IN HIGHLY REVERBERANT ENVIRONMENTS USING SOFT TIME-FREQUENCY MASK</b> .....	383
<i>Vladimir Tourbabin ; Jacob Donley ; Boaz Rafaely ; Ravish Mehra</i>	
<b>ANALYTICAL METHOD OF 2.5D EXTERIOR SOUND FIELD SYNTHESIS BY USING MULTIPOLE LOUDSPEAKER ARRAY</b> .....	388
<i>Kenta Imaizumi ; Kimitaka Tsutsumi ; Atsushi Nakadaira ; Yoichi Haneda</i>	
<b>A SPARSE BAYESIAN LEARNING BASED RIR RECONSTRUCTION METHOD FOR ACOUSTIC TOA AND DOA ESTIMATION</b> .....	393
<i>Zonglong Bai ; Jesper Rindom Jensen ; Jinwei Sun ; Mads Græsbøll Christensen</i>	
<b>Author Index</b>	