

2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2021)

**New Paltz, New York, USA
17-20 October 2021**



**IEEE Catalog Number: CFP21AUD-POD
ISBN: 978-1-6654-4871-0**

**Copyright © 2021 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:	CFP21AUD-POD
ISBN (Print-On-Demand):	978-1-6654-4871-0
ISBN (Online):	978-1-6654-4870-3
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
Web: www.proceedings.com

CURRAN ASSOCIATES INC.
proceedings
.com

TABLE OF CONTENTS

ADAPTIVE BINAURAL FILTERING FOR A MULTIPLE-TALKER LISTENING SYSTEM USING REMOTE AND ON-EAR MICROPHONES	1
<i>Ryan M. Corey, Andrew C. Singer</i>	
DPLM: A DEEP PERCEPTUAL SPATIAL-AUDIO LOCALIZATION METRIC	6
<i>Pranay Manocha, Anurag Kumar, Buye Xu, Anjali Menon, Israel D. Gebru, Vamsi K. Ithapu, Paul Calamia</i>	
ASSESSING SEGMENTAL IMPACT FOR OBJECTIVE SPEECH QUALITY EVALUATION.....	11
<i>Zhixing Liu, Yannan Wang, Gaoxiong Yi, Tao Yu, Fei Chen</i>	
IMPROVED INTELLIGIBILITY PREDICTION IN THE MODULATION DOMAIN	16
<i>Ahmed Alghamdi, Wai-Yip Chan, Daniel Fogerty, Jesper Jensen</i>	
MIMII DUE: SOUND DATASET FOR MALFUNCTIONING INDUSTRIAL MACHINE INVESTIGATION AND INSPECTION WITH DOMAIN SHIFTS DUE TO CHANGES IN OPERATIONAL AND ENVIRONMENTAL CONDITIONS.....	21
<i>Ryo Tanabe, Harsh Purohit, Kota Dohi, Takashi Endo, Yuki Nikaido, Toshiki Nakamura, Yohei Kawaguchi</i>	
IDENTIFYING ACTIONS FOR SOUND EVENT CLASSIFICATION.....	26
<i>Benjamin Elizalde, Radu Revutchi, Samarjit Das, Bhiksha Raj, Ian Lane, Laurie M. Heller</i>	
POINT CLOUD AUDIO PROCESSING	31
<i>Krishna Subramani, Paris Smaragdis</i>	
WHO CALLS THE SHOTS? RETHINKING FEW-SHOT LEARNING FOR AUDIO	36
<i>Yu Wang, Nicholas J. Bryan, Justin Salamon, Mark Cartwright, Juan Pablo Bello</i>	
SOUND EVENT DETECTION WITH ADAPTIVE FREQUENCY SELECTION	41
<i>Zhepei Wang, Jonah Casebeer, Adam Clemmitt, Efthymios Tzinis, Paris Smaragdis</i>	
SEPARATE BUT TOGETHER: UNSUPERVISED FEDERATED LEARNING FOR SPEECH ENHANCEMENT FROM NON-IID DATA	46
<i>Efthymios Tzinis, Jonah Casebeer, Zhepei Wang, Paris Smaragdis</i>	
SPARSE, EFFICIENT, AND SEMANTIC MIXTURE INVARIANT TRAINING: TAMING IN- THE-WILD UNSUPERVISED SOUND SEPARATION	51
<i>Scott Wisdom, Aren Jansen, Ron J. Weiss, Hakan Erdogan, John R. Hershey</i>	
CONVOLUTIVE PREDICTION FOR REVERBERANT SPEECH SEPARATION.....	56
<i>Zhong-Qiu Wang, Gordon Wichern, Jonathan Le Roux</i>	
WEAKLY SUPERVISED SOURCE-SPECIFIC SOUND LEVEL ESTIMATION IN NOISY SOUNDSCAPES.....	61
<i>Aurora Cramer, Mark Cartwright, Fatemeh Pishdadian, Juan Pablo Bello</i>	
A STREAMWISE GAN VOCODER FOR WIDEBAND SPEECH CODING AT VERY LOW BIT RATE	66
<i>Ahmed Mustafa, Jan Bütthe, Srikanth Korse, Kishan Gupta, Guillaume Fuchs, Nicola Pia</i>	
ADVERSARIAL AUTO-ENCODING FOR PACKET LOSS CONCEALMENT.....	71
<i>Santiago Pascual, Joan Serrà, Jordi Pons</i>	

SPATIAL CODING FOR MICROPHONE ARRAYS USING IPNLMS-BASED RTF ESTIMATION.....	76
<i>Daniel T. Jones, Dushyant Sharma, Stanislav Yu. Kruchinin, Patrick A. Naylor</i>	
EXCITATION-INHIBITION CELL ACTIVITY PATTERNS FOR BINAURAL SOURCE LOCALISATION.....	81
<i>Hsuan-Yang Wang, Philip Nelson, Christine Evers</i>	
SPEECH INTELLIGIBILITY OF MANDARIN- AND GERMAN-SPEAKING LISTENERS IN CHALLENGING CONDITIONS	86
<i>Hongmei Hu, Stephan D. Ewert</i>	
CONTROLLING THE REMIXING OF SEPARATED DIALOGUE WITH A NON-INTRUSIVE QUALITY ESTIMATE.....	91
<i>Matteo Torcoli, Jouni Paulus, Thorsten Kastner, Christian Uhle</i>	
SIDIQ: COMPUTATIONAL QUALITY ASSESSMENT OF ENHANCED SPEECH BASED ON AUDITORY FIGURE-GROUND SEGREGATION, SIMILARITY, AND DISTURBANCE	96
<i>Benjamin Stahl, Alois Sontacchi</i>	
OBJECTIVE METRICS TO EVALUATE RESIDUAL-ECHO SUPPRESSION DURING DOUBLE-TALK.....	101
<i>Amir Ivry, Israel Cohen, Baruch Berdugo</i>	
PREDICTION OF MISSING FREQUENCY RESPONSE FUNCTIONS THROUGH DEEP IMAGE PRIOR	106
<i>Raffaele Malvermi, Fabio Antonacci, Augusto Sarti, Roberto Corradi</i>	
USER-GUIDED ONE-SHOT DEEP MODEL ADAPTATION FOR MUSIC SOURCE SEPARATION	111
<i>Giorgia Cantisani, Alexey Ozerov, Slim Essid, Gaël Richard</i>	
VQPC-GAN: VARIABLE-LENGTH ADVERSARIAL AUDIO SYNTHESIS USING VECTOR-QUANTIZED CONTRASTIVE PREDICTIVE CODING.....	116
<i>Javier Nistal, Cyran Aouameur, Stefan Lattner, Gaël Richard</i>	
LEARNING MULTI-PITCH ESTIMATION FROM WEAKLY ALIGNED SCORE-AUDIO PAIRS USING A MULTI-LABEL CTC LOSS.....	121
<i>Christof Weiß, Geoffroy Peeters</i>	
DISENTANGLEMENT LEARNING FOR VARIATIONAL AUTOENCODERS APPLIED TO AUDIO-VISUAL SPEECH ENHANCEMENT	126
<i>Guillaume Carbajal, Julius Richter, Timo Gerkmann</i>	
FASTERIVA: UPDATE RULES FOR INDEPENDENT VECTOR ANALYSIS BASED ON NEGENTROPY AND THE MAJORITY-MINIMIZE PRINCIPLE	131
<i>Andreas Brendel, Walter Kellermann</i>	
LOW COMPLEXITY ONLINE CONVOLUTIONAL BEAMFORMING.....	136
<i>Sebastian Braun, Ivan Tashev</i>	
SUPERRESOLUTION PHOTOACOUSTIC TOMOGRAPHY USING RANDOM SPECKLE ILLUMINATION AND SECOND ORDER MOMENTS	141
<i>Osman Asif Malik, Venkatalakshmi Vyjayanthi Narumanchi, Stephen Becker, Todd W. Murray</i>	

CLOSING THE GAP BETWEEN TIME-DOMAIN MULTI-CHANNEL SPEECH ENHANCEMENT ON REAL AND SIMULATION CONDITIONS	146
<i>Wangyou Zhang, Jing Shi, Chenda Li, Shinji Watanabe, Yanmin Qian</i>	
MESHRR: A DATASET OF ROOM IMPULSE RESPONSES ON MESHED GRID POINTS FOR EVALUATING SOUND FIELD ANALYSIS AND SYNTHESIS METHODS.....	151
<i>Shoichi Koyama, Tomoya Nishida, Keisuke Kimura, Takumi Abe, Natsuki Ueno, Jesper Brunnström</i>	
ANALYSIS OF FREQUENCY-DEPENDENT BEHAVIOR OF ROOM REFLECTIONS USING SPHERICAL MICROPHONE MEASUREMENTS & VON MISES-FISHER CLUSTERING.....	156
<i>Amy Bastine, Thushara D. Abhayapala, Jihui Zhang</i>	
DF-CONFORMER: INTEGRATED ARCHITECTURE OF CONV-TASNET AND CONFORMER USING LINEAR COMPLEXITY SELF-ATTENTION FOR SPEECH ENHANCEMENT.....	161
<i>Yuma Koizumi, Shigeki Karita, Scott Wisdom, Hakan Erdogan, John R. Hershey, Llion Jones, Michiel Bacchiani</i>	
HIFI-GAN-2: STUDIO-QUALITY SPEECH ENHANCEMENT VIA GENERATIVE ADVERSARIAL NETWORKS CONDITIONED ON ACOUSTIC FEATURES.....	166
<i>Jiaqi Su, Zeyu Jin, Adam Finkelstein</i>	
ZERO-SHOT PERSONALIZED SPEECH ENHANCEMENT THROUGH SPEAKER- INFORMED MODEL SELECTION	171
<i>Aswin Sivaraman, Minje Kim</i>	
TEST-TIME ADAPTATION TOWARD PERSONALIZED SPEECH ENHANCEMENT: ZERO- SHOT LEARNING WITH KNOWLEDGE DISTILLATION	176
<i>Sunwoo Kim, Minje Kim</i>	
TOWARDS LARGE SCALE ECOACOUSTIC MONITORING WITH SMALL AMOUNTS OF LABELED DATA.....	181
<i>Enis Berk Çoban, Ali Raza Syed, Dara Pir, Michael I Mandel</i>	
ANOMALOUS SOUND DETECTION USING ATTENTIVE NEURAL PROCESSES	186
<i>Gordon Wichern, Ankush Chakrabarty, Zhong-Qiu Wang, Jonathan Le Roux</i>	
A MULTI-HEAD RELEVANCE WEIGHTING FRAMEWORK FOR LEARNING RAW WAVEFORM AUDIO REPRESENTATIONS	191
<i>Debotam Dutta, Purvi Agrawal, Sriram Ganapathy</i>	
CROSS-DOMAIN SEMI-SUPERVISED AUDIO EVENT CLASSIFICATION USING CONTRASTIVE REGULARIZATION	196
<i>Donmoon Lee, Kyogu Lee</i>	
POLYNOMIAL MATRIX EIGENVALUE DECOMPOSITION-BASED SOURCE SEPARATION USING INFORMED SPHERICAL MICROPHONE ARRAYS	201
<i>Vincent W. Neo, Christine Evers, Patrick A. Naylor</i>	
LOW-COMPLEXITY STEERED RESPONSE POWER MAPPING BASED ON NYQUIST- SHANNON SAMPLING	206
<i>Thomas Dietzen, Enzo De Sena, Toon Van Waterschoot</i>	
DIFFERENTIABLE TRACKING-BASED TRAINING OF DEEP LEARNING SOUND SOURCE LOCALIZERS	211
<i>Sharath Adavanne, Archontis Politis, Tuomas Virtanen</i>	

SPHERICAL HARMONIC DIAGONAL UNLOADING BEAMFORMING WITH EGO-NOISE REDUCTION FOR DOA ESTIMATION FROM AUTONOMOUS SYSTEMS	216
<i>Daniele Salvati, Carlo Drioli, Gian Luca Foresti</i>	
FILTERED NOISE SHAPING FOR TIME DOMAIN ROOM IMPULSE RESPONSE ESTIMATION FROM REVERBERANT SPEECH.....	221
<i>Christian J. Steinmetz, Vamsi Krishna Ithapu, Paul Calamia</i>	
BLIND ROOM PARAMETER ESTIMATION USING MULTIPLE MULTICHANNEL SPEECH RECORDINGS	226
<i>Prerak Srivastava, Antoine Deleforge, Emmanuel Vincent</i>	
SPHERICAL HARMONIC DECOMPOSITION OF A SOUND FIELD BASED ON MICROPHONES AROUND THE CIRCUMFERENCE OF A HUMAN HEAD	231
<i>Jens Ahrens, Hannes Helmholz, David Lou Alon, Sebastia V. Amengual Garí</i>	
AMBIENT-AWARE SOUND FIELD TRANSLATION USING OPTIMAL SPATIAL FILTERING	236
<i>Maximilian Kentgens, Peter Jax</i>	
INTERNAL TIME DELAY CALIBRATION OF RIGID SPHERICAL MICROPHONE ARRAYS FOR MULTI-PERSPECTIVE 6DOF AUDIO RECORDINGS	241
<i>Ege Erdem, Orhun Olgun, Hüseyin Hacihabiboglu</i>	
CROWDSOURCING STRONG LABELS FOR SOUND EVENT DETECTION	246
<i>Irene Martín-Morató, Manu Harju, Annamaria Mesaros</i>	
SELF-SUPERVISED LEARNING FROM AUTOMATICALLY SEPARATED SOUND SCENES	251
<i>Eduardo Fonseca, Aren Jansen, Daniel P. W. Ellis, Scott Wisdom, Marco Tagliasacchi, John R. Hershey, Manoj Plakal, Shawn Hershey, R. Channing Moore, Xavier Serra</i>	
ADAPTIVE GENERALIZED CROSS-ENTROPY LOSS FOR SOUND EVENT CLASSIFICATION WITH NOISY LABELS	256
<i>Jun Deng, Chunhui Gao, Qian Feng, Xinzhou Xu, Zhaopeng Chen</i>	
KERNEL LEARNING FOR SOUND FIELD ESTIMATION WITH L1 AND L2 REGULARIZATIONS.....	261
<i>Ryosuke Horiuchi, Shoichi Koyama, Juliano G. C. Ribeiro, Natsuki Ueno, Hiroshi Saruwatari</i>	
MAGNITUDE MODELLING OF INDIVIDUALIZED HRTFS USING DNN BASED SPHERICAL HARMONIC ANALYSIS	266
<i>Jingwei Xi, Wen Zhang, Thushara D. Abhayapala</i>	
2D LOCAL EXTERIOR SOUND FIELD REPRODUCTION USING AN ADDITION THEOREM BASED ON CIRCULAR HARMONIC EXPANSION.....	271
<i>Yi Ren, Yoichi Haneda</i>	
2D MULTIZONE SOUND FIELD SYNTHESIS WITH INTERIOR-EXTERIOR AMBISONICS	276
<i>Takuma Okamoto</i>	
MEAN-SQUARE-ERROR-BASED SECONDARY SOURCE PLACEMENT IN SOUND FIELD SYNTHESIS WITH PRIOR INFORMATION ON DESIRED FIELD	281
<i>Keisuke Kimura, Shoichi Koyama, Natsuki Ueno, Hiroshi Saruwatari</i>	
SPHERICAL ARRAY BASED DRONE NOISE MEASUREMENTS AND MODELLING FOR DRONE NOISE REDUCTION VIA PROPELLER PHASE CONTROL.....	286
<i>Hanwen Bi, Fei Ma, Thushara D. Abhayapala, Prasanga N. Samarasinghe</i>	

AUTO-DSP: LEARNING TO OPTIMIZE ACOUSTIC ECHO CANCELLERS	291
<i>Jonah Casebeer, Nicholas J. Bryan, Paris Smaragdis</i>	
FAST CONVERGENT METHOD FOR ACTIVE NOISE CONTROL OVER SPATIAL REGION WITH CAUSAL CONSTRAINT	296
<i>Naoki Murata, Yuhta Takida, Tetsu Magariyachi</i>	
ACTIVE NOISE CONTROL OVER 3D SPACE WITH REMOTE MICROPHONE TECHNIQUE IN THE WAVE DOMAIN.....	301
<i>Huiyuan Sun, Jihui Zhang, Thushara D. Abhayapala, Prasanga N. Samarasinghe</i>	
END-TO-END ZERO-SHOT VOICE CONVERSION USING A DDSP VOCODER.....	306
<i>Shahan Nercessian</i>	
ON THE ROLE OF LIP REFLECTION/TRANSMISSION IN THE RELATIONSHIP BETWEEN LPC AND WAVEGUIDE VOCAL TRACT MODELS.....	311
<i>Tamara Smyth, Devansh Zurale</i>	
HARP-NET: HYPER-AUTOENCODED RECONSTRUCTION PROPAGATION FOR SCALABLE NEURAL AUDIO CODING.....	316
<i>Darius Petermann, Seungkwon Beack, Minje Kim</i>	
PERIODIC ANALYSIS OF NONLINEAR VIRTUAL ANALOG MODELS	321
<i>François G. Germain</i>	
A POLYNOMIAL EIGENVALUE DECOMPOSITION MUSIC APPROACH FOR BROADBAND SOUND SOURCE LOCALIZATION	326
<i>Aidan O. T. Hogg, Vincent W. Neo, Stephan Weiss, Christine Evers, Patrick A. Naylor</i>	
JOINT DIRECTION AND PROXIMITY CLASSIFICATION OF OVERLAPPING SOUND EVENTS FROM BINAURAL AUDIO	331
<i>Daniel Aleksander Krause, Archontis Politis, Annamaria Mesaros</i>	
SALADNET: SELF-ATTENTIVE MULTISOURCE LOCALIZATION IN THE AMBISONICS DOMAIN	336
<i>Pierre-Amaury Grumiaux, Srdan Kitic, Prerak Srivastava, Laurent Girin, Alexandre Guérin</i>	
LOW-ORDER FILTER APPROXIMATION OF DIFFRACTION FOR VIRTUAL ACOUSTICS	341
<i>Christoph Kirsch, Stephan D. Ewert</i>	
SPATIAL SUBTRACTION OF REFLECTIONS FROM ROOM IMPULSE RESPONSES MEASURED WITH A SPHERICAL MICROPHONE ARRAY	346
<i>Thomas Deppisch, Jens Ahrens, Sebastia V. Amengual Garí, Paul Calamia</i>	
STOCHASTIC REVERBERATION MODEL WITH A FREQUENCY DEPENDENT ATTENUATION.....	351
<i>Achille Aknin, Roland Badeau</i>	
A UNIVERSAL DEEP ROOM ACOUSTICS ESTIMATOR.....	356
<i>Paula Sánchez López, Paul Callens, Milos Cernak</i>	
SPATIAL FILTER BANK IN THE SPHERICAL HARMONIC DOMAIN: RECONSTRUCTION AND APPLICATION	361
<i>Christoph Hold, Sebastian J. Schlecht, Archontis Politis, Ville Pulkki</i>	

SOUNDFIELD RECONSTRUCTION IN REVERBERANT ROOMS BASED ON
COMPRESSIVE SENSING AND IMAGE-SOURCE MODELS OF EARLY REFLECTIONS 366
Stefano Damiano, Federico Borra, Alberto Bernardini, Fabio Antonacci, Augusto Sarti

RENDERING OF SOURCE SPREAD FOR ARBITRARY PLAYBACK SETUPS BASED ON
SPATIAL COVARIANCE MATCHING 371
Leo McCormack, Archontis Politis, Ville Pulkki

Author Index