

2017 Hands-free Speech Communications and Microphone Arrays (HSCMA 2017)

**San Francisco, California, USA
1-3 March 2017**



**IEEE Catalog Number: CFP17HSM-POD
ISBN: 978-1-5090-5926-3**

**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:	CFP17HSM-POD
ISBN (Print-On-Demand):	978-1-5090-5926-3
ISBN (Online):	978-1-5090-5925-6

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

MIC: MICROPHONE ARRAY PROCESSING

MIC.1: PERFORMANCE ANALYSIS OF THE EXTENDED BINAURAL MVDR	1
BEAMFORMER WITH PARTIAL NOISE ESTIMATION IN A HOMOGENEOUS NOISE FIELD	

Nico Gößling, Daniel Marquardt, Simon Doclo, Universität Oldenburg, Germany

MIC.2: A STUDY ON SEARCH GRID POINTS FOR DATA-DRIVEN 3-D	6
BEAMSTEERING	

JeeSok Lee, Soo-Whan Chung, Yonsei University, Republic of Korea; Min-Seok Choi, NAVER Corporation, Republic of Korea; Hong-Goo Kang, Yonsei University, Republic of Korea

MIC.3: AN ITERATIVE MULTICHANNEL SUBSPACE-BASED COVARIANCE	11
SUBTRACTION METHOD FOR RELATIVE TRANSFER FUNCTION ESTIMATION	

Reza Varzandeh, Maja Taseska, Emanuel Habets, International Audio Laboratories Erlangen, Germany

MIC.4: ONLINE MEETING RECOGNITION IN NOISY ENVIRONMENTS WITH	16
TIME-FREQUENCY MASK BASED MVDR BEAMFORMING	

Shoko Araki, Nobutaka Ito, Marc Delcroix, Atsunori Ogawa, Keisuke Kinoshita, NTT Communication Science Laboratories, Japan; Takuya Higuchi, NTT Communication Science Laboratories, Japan; Takuya Yoshioka, Microsoft Research, United States; Dung Tran, Shigeki Karita, Tomohiro Nakatani, NTT Communication Science Laboratories, Japan

MIC.5: PERFORMANCE OF FIXED IN-CAR MICROPHONE ARRAY BEAMFORMER	21
UNDER VARIATIONS IN CAR NOISE	

Vladimir Tourbabin, Ilan Malka, Eli Tzirkel-Hancock, General Motors, Israel

NOISE: NOISE & REVERBERATION SUPPRESSION

NOISE.1: JOINT NOISE AND MASK AWARE TRAINING FOR DNN-BASED SPEECH	101
ENHANCEMENT WITH SUB-BAND FEATURES	

Qing Wang, Jun Du, Li-Rong Dai, University of Science and Technology of China, China; Chin-Hui Lee, Georgia Institute of Technology, United States

NOISE.2: TWO-STAGE SPEECH ENHANCEMENT WITH MANIPULATION OF THE	106
CEPSTRAL EXCITATION	

Samy Elshamy, Technische Universität Braunschweig, Germany; Nilesch Madhu, Wouter Tirry, NXP Software, Belgium; Tim Fingscheidt, Technische Universität Braunschweig, Germany

NOISE.3: MODULATION-DOMAIN SPEECH ENHANCEMENT USING A KALMAN	111
FILTER WITH A BAYESIAN UPDATE OF SPEECH AND NOISE IN THE LOG-SPECTRAL DOMAIN	

Nikolaos Dionelis, Mike Brookes, Imperial College London, United Kingdom

NOISE.4: EVD-BASED MULTI-CHANNEL DEREVERBERATION OF A MOVING	116
SPEAKER USING DIFFERENT RETF ESTIMATION METHODS	

Ina Kodrasi, Simon Doclo, University of Oldenburg, Germany

NOISE.5: ROOM EQUALIZATION BASED ON MEASUREMENTS WITH MOVING MICROPHONES	121
<i>Radoslaw Mazur, Fabrice Katzberg, Huy Phan, Alfred Mertins, University of Lübeck, Germany</i>	
NOISE.6: PERFORMANCE COMPARISON OF REAL-TIME SINGLE-CHANNEL SPEECH DEREVERBERATION ALGORITHMS	126
<i>Feifei Xiong, Fraunhofer IDMT/HSA, Germany; Bernd T. Meyer, Johns Hopkins University, United States; Benjamin Cauchi, Fraunhofer IDMT/HSA, Germany; Ante Jukic, Simon Doclo, University of Oldenburg, Germany; Stefan Goetze, Fraunhofer IDMT/HSA, Germany</i>	
NOISE.7: ESTIMATION OF THE PERCEIVED LEVEL OF REVERBERATION USING NON-INTRUSIVE SINGLE-CHANNEL VARIANCE OF DECAY RATES	131
<i>James Eaton, Hamza Javed, Patrick Naylor, Imperial College London, United Kingdom</i>	
SEP: SOURCE SEPARATION	
SEP.1: MULTIPLE-TARGET DEEP LEARNING FOR LSTM-RNN BASED SPEECH ENHANCEMENT	136
<i>Lei Sun, Jun Du, Li-Rong Dai, University of Science and Technology of China, China; Chin-Hui Lee, Georgia Institute of Technology, United States</i>	
SEP.2: DISCRIMINATIVE NON-NEGATIVE MATRIX FACTORIZATION WITH MAJORIZATION-MINIMIZATION	141
<i>Li Li, University of Tsukuba, Japan; Hirokazu Kameoka, NTT Communication Science Laboratories, Nippon Telegraph and Telephone Corporation, Japan; Shoji Makino, University of Tsukuba, Japan</i>	
SEP.3: EXTRACTION OF ACOUSTIC SOURCES FOR MULTIPLE ARRAYS BASED ON THE RAY SPACE TRANSFORM	146
<i>Federico Borra, Fabio Antonacci, Augusto Sarti, Stefano Tubaro, Politecnico di Milano, Italy</i>	
SEP.4: SOUND FIELD SEPARATION IN A MIXED ACOUSTIC ENVIRONMENT USING A SPARSE ARRAY OF HIGHER ORDER SPHERICAL MICROPHONES	151
<i>Abdullah Fahim, Prasanga N. Samarasinghe, Thushara D. Abhayapala, The Australian National University, Australia</i>	
SEP.5: ROBUST AND LOW-COMPLEXITY BLIND SOURCE SEPARATION FOR MEETING ROOMS	156
<i>W. Bastiaan Kleijn, Victoria University of Wellington, New Zealand; Felicia Lim, Google Inc., United States</i>	
SEP.6: TOWARDS REAL-TIME SOURCE COUNTING BY ESTIMATION OF COHERENT-TO-DIFFUSE RATIOS FROM AD-HOC MICROPHONE ARRAY RECORDINGS	161
<i>Shahab Pasha, Jacob Donley, Christian Ritz, University of Wollongong, Australia; Yue Xian Zou, Peking University, China</i>	
SEP.7: SINGLE SENSOR AUDIOVISUAL SPEECH SOURCE SEPARATION.....	166
<i>Pierre Narvor, Bertrand Rivet, Christian Jutten, Gipsa-Lab, France</i>	

NET: NEURAL NETWORKS FOR ROBUST SPEECH PROCESSING

NET.1: AN EXTENDED EXPERIMENTAL INVESTIGATION OF DNN UNCERTAINTY26 **PROPAGATION FOR NOISE ROBUST ASR**

Karan Nathwani, INRIA, France; Juan A. Morales Cordovilla, Universidad de Granada, Austria; Sunit Sivasankaran, Irina Illina, Emmanuel Vincent, INRIA, France

NET.2: AN IMPROVED UNCERTAINTY DECODING SCHEME WITH WEIGHTED31 **SAMPLES FOR MULTICHANNEL DNN-HMM HYBRID SYSTEMS**

Christian Huemmer, Friedrich-Alexander University Erlangen-Nuremberg, Germany; Ramón Fernández Astudillo, INESC-ID-Lisboa, Portugal; Walter Kellermann, Friedrich-Alexander University Erlangen-Nuremberg, Germany

NET.3: A UNIFIED DEEP MODELING APPROACH TO SIMULTANEOUS SPEECH36 **DEREVERBERATION AND RECOGNITION FOR THE REVERB CHALLENGE**

Bo Wu, Xidian University, China; Kehuang Li, Zhen Huang, Georgia Institute of Technology, United States; Sabato Marco Siniscalchi, University of Enna Kore, Italy; Minglei Yang, Xidian University, China; Chin-Hui Lee, Georgia Institute of Technology, United States

NET.4: DEEP BI-DIRECTIONAL LONG SHORT-TERM MEMORY BASED SPEECH41 **ENHANCEMENT FOR WIND NOISE REDUCTION**

Jinkyu Lee, Keulbit Kim, Yonsei University, Republic of Korea; Turaj Shabestary, Google Incorporated, United States; Hong-Goo Kang, Yonsei University, Republic of Korea

NET.5: EFFICIENT TARGET ACTIVITY DETECTION BASED ON RECURRENT46 **NEURAL NETWORKS**

Daniel Gerber, Stefan Meier, Walter Kellermann, Friedrich-Alexander-University Erlangen-Nuremberg, Germany

ECHO: ECHO REDUCTION, MEASUREMENT, AND RENDERING

ECHO.1: ROBUST VARIABLE-REGULARIZED RLS ALGORITHMS.....171

Camelia Elisei-Iliescu, Cristian Stanciu, Constantin Paleologu, University Politehnica of Bucharest, Romania; Jacob Benesty, INRS-EMT, University of Quebec, Canada; Cristian Anghel, Silviu Ciochina, University Politehnica of Bucharest, Romania

ECHO.2: COHERENCE-AWARE STEREOPHONIC RESIDUAL ECHO ESTIMATION.....176

Maria Luis Valero, Ilkay Yildiz, International Audio Laboratories Erlangen, Germany; Edwin Mabande, Fraunhofer Institut Integrierte Schaltungen, Germany; Emanuel Habets, International Audio Laboratories Erlangen, Germany

ECHO.3: KALMAN FILTER BASED STEREO SYSTEM IDENTIFICATION WITH AUTO-181 **AND CROSS-DECORRELATION**

Stefan Kühl, Christiane Antweiler, Tobias Hübschen, Peter Jax, RWTH Aachen, Germany

ECHO.4: CLOCK DRIFT ESTIMATION AND COMPENSATION FOR ASYNCHRONOUS186 **IMPULSE RESPONSE MEASUREMENTS**

Hannes Gamper, Microsoft Research, United States

ECHO.5: MULTIGRID RECONSTRUCTION OF SOUND FIELDS USING MOVING MICROPHONES	191
<i>Fabrice Katzberg, Radoslaw Mazur, Marco Maass, Philipp Koch, Alfred Mertins, University of Luebeck, Germany</i>	
ECHO.6: LISTENING-AREA-INFORMED SOUND FIELD REPRODUCTION WITH GAUSSIAN PRIOR BASED ON CIRCULAR HARMONIC EXPANSION	196
<i>Natsuki Ueno, Shoichi Koyama, Hiroshi Saruwatari, The University of Tokyo, Japan</i>	
ECHO.7: FAST COMPUTATION OF CUBATURE FORMULAE FOR THE SPHERE.....	201
<i>Mark Thomas, Dolby Laboratories, United States</i>	
 ROBOT: ROBOT AUDITION	
ROBOT.1: MICROPHONE ARRAY SIGNAL PROCESSING FOR ROBOT AUDITION.....	51
<i>Heinrich Loellmann, Friedrich-Alexander University Erlangen-Nuremberg, Germany; Alastair Moore, Patrick Naylor, Imperial College London, United Kingdom; Boaz Rafaely, Ben-Gurion University of the Negev, Israel; Radu Horaud, INRIA Grenoble Rhone-Alpes, France; Alexandre Mazel, Softbank Robotics Europe, France; Walter Kellermann, Friedrich-Alexander University Erlangen-Nuremberg, Germany</i>	
ROBOT.2: HRTF-BASED TWO-DIMENSIONAL ROBUST LEAST-SQUARES FREQUENCY-INVARIANT BEAMFORMER DESIGN FOR ROBOT AUDITION	56
<i>Hendrik Barfuss, Michael Buerger, Jasper Podschus, Walter Kellermann, Friedrich-Alexander University Erlangen-Nuremberg, Germany</i>	
ROBOT.3: LONG-TERM ROBOT MOTION PLANNING FOR ACTIVE SOUND SOURCE LOCALIZATION WITH MONTE CARLO TREE SEARCH	61
<i>Quan Van Nguyen, Francis Colas, Emmanuel Vincent, François Charpillet, INRIA, France</i>	
ROBOT.4: A UNIFIED FRAMEWORK FOR MULTIPLE ARRAYS ON A ROBOT AND APPLICATION TO SOUND LOCALIZATION	66
<i>Lior Madmoni, Ben-Gurion University of the Negev, Israel; Hendrik Barfuss, Friedrich-Alexander University Erlangen-Nuremberg, Germany; Boaz Rafaely, Ben-Gurion University of the Negev, Israel; Walter Kellermann, Friedrich-Alexander University Erlangen-Nuremberg, Germany</i>	
ROBOT.5: AUDIO-VISUAL TRACKING BY DENSITY APPROXIMATION IN A SEQUENTIAL BAYESIAN FILTERING FRAMEWORK	71
<i>Israel D. Gebru, INRIA, France; Christine Evers, Patrick A. Naylor, Imperial College London, United Kingdom; Radu Horaud, INRIA, France</i>	
 LOC: LOCALIZATION & TRACKING	
LOC.1: COST FUNCTION FOR SOUND SOURCE LOCALIZATION WITH ARBITRARY MICROPHONE ARRAYS	76
<i>Ivan Tashev, Microsoft Corporation, United States; Long Le, University of Illinois at Urbana-Champaign, United States; Vani Gopalakrishna, Andrew Lovitt, Microsoft Corporation, United States</i>	
LOC.2: MULTI-SOURCE ESTIMATION CONSISTENCY FOR IMPROVED MULTIPLE DIRECTION-OF-ARRIVAL ESTIMATION	81
<i>Sina Hafezi, Alastair H. Moore, Patrick A. Naylor, Imperial College London, United Kingdom</i>	

LOC.3: DOA ESTIMATION IN NOISY ENVIRONMENT WITH UNKNOWN NOISE	86
POWER USING THE EM ALGORITHM	
<i>Ofer Schwartz, Yuval Dorfan, Bar-Ilan University, Israel; Maja Taseska, Emanuel Habets, International Audio Laboratories Erlangen, Germany; Sharon Gannot, Bar-Ilan University, Israel</i>	
LOC.4: SPEAKER TRACKING IN REVERBERANT ENVIRONMENTS USING	91
MULTIPLE DIRECTIONS OF ARRIVAL	
<i>Christine Evers, Imperial College London, United Kingdom; Boaz Rafaely, Ben-Gurion University of the Negev, Israel; Patrick A. Naylor, Imperial College London, United Kingdom</i>	
LOC.5: TOWARDS ACOUSTICALLY ROBUST LOCALIZATION OF SPEAKERS IN A	96
REVERBERANT ENVIRONMENT	
<i>Boaz Rafaely, Ben-Gurion University of the Negev, Israel; Dorothea Kolossa, Ruhr-University Bochum, Germany; Yanir Maymon, Ben-Gurion University of the Negev, Israel</i>	