2014 14th International Workshop on Acoustic Signal Enhancement

(IWAENC 2014)

Antibes – Juan les Pins, France
8-11 September 2014

IEEE Catalog Number: CFP14WAE-POD
ISBN: 978-1-4799-6809-1

Registration

Welcome reception

Registration

I1: Welcome Address

K1: Keynote 1

Next Generation 3 GPP Speech Coding: Enhanced Voice Services (EVS)

The EVS (Enhanced Voice Services) project in 3GPP aims at improving user experience through enhancement of the key element of telephony - speech quality in speech coding. EVS is the next generation speech coding standard in 3GPP after the successful coders used in mobile telephony. The EVS coder is especially designed for packet-switched networks, such as voice-over-LTE as a key target application. Besides ensuring enhanced quality for VoLTE, EVS addresses all networks, including mobile VoIP with QoS, best effort VoIP, and CS. EVS provides improved user experience through super-wideband bandwidth at low bit rates (at 32 kHz sampling), improved robustness through significantly better error resilience, improved music performance, wide bit-rate range and all bandwidths (narrowband, wideband, super-wideband, full band) for maximum flexibility, and improved capacity by introduction of a variable-bit-rate mode. The standardization process includes a qualification phase to determine the most promising technologies, a selection phase which actually includes a candidate developed jointly by all proponent candidates, and a characterization testing to obtain more detailed information about the performance of the coder. The presentation will address the goals of the project, the standardization process, the speech quality, and the algorithmic elements of the coder.

K2: Keynote 2

Design and Implementation of Small Microphone Arrays for Acoustic Speech Signal

Microphone array is a generic expression used to refer to a sound system that has multiple microphones. These microphones can be distributed into an arbitrary network (often called a microphone sensor network) or arranged into a particular geometry (then called an organized array). In most of the cases, however, when we say microphones arrays, we mean organized arrays in which the sensors' positions relative to a reference point are known to the subsequent processors. This kind of arrays can be used to solve many important problems such as source localization/tracking, noise reduction/speech enhancement, source separation, dereverberation, spatial sound recording, etc. Consequently, the design of such microphone arrays and the associated processing algorithms has attracted a significant amount of research and engineering interest over the last three decades. In this talk, we will address the problems and challenges of microphone array processing. Based on how they respond to the sound field, the microphone arrays are categorized into two classes: additive arrays and differential ones. The former refers to those arrays with large inter-element spacing (from a couple of centimeters to a couple of decimeters) and optimal beamforming in broadside directions while the latter refers to those microphone arrays that are responsive to the spatial derivatives of an acoustic pressure field. We will present a brief overview of the basic principles underlying the two classes of arrays. We will then focus on discussing the design of differential microphone arrays (DMAs), which have many advantages over additive arrays in practical applications, particularly with small devices. We will elaborate, using examples, on how to design DMA beamforming that is good for processing broadband signals like speech. We will also discuss how to deal with the white noise amplification problem that used to be a big problem preventing DMA from practical usage.

Coffee Break

D1: Demonstration 1

Real-time Microphone Selection in Noisy Reverberant Environments for Teleconferencing Systems

P1: Poster Session 1

P1.1 ***HMM-Based Artificial Bandwidth Extension Supported by Neural Networks***
Patrick Bauer (Technische Universität Braunschweig & Institute for Communications Technology, Germany); Johannes Abel and Tim Fingscheidt (Technische Universität Braunschweig, Germany)
pp. 1-5

P1.2 ***Unbiased Coherent-to-Diffuse Ratio Estimation for Dereverberation***
Andreas Schwarz (Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany); Walter Kellermann (University Erlangen-Nuremberg, Germany)
pp. 6-10

P1.3 ***A Posteriori Speech Presence Probability Estimation Based on Averaged Observations and a Super-Gaussian Speech Model***
Balazs Fodor (Technische Universität Braunschweig & Institute for Communications Technology, Germany); Timo Gerkmann (University of Oldenburg, Germany)
pp. 11-15

P1.4 ***An Adaptive Microphone Array Topology For Target Signal Extraction With Humanoid Robots***
Hendrik Barfuss (University of Erlangen-Nuremberg, Germany); Walter Kellermann (University Erlangen-Nuremberg, Germany)
pp. 16-20

P1.5 ***On the Statistics and the Detection of Multichannel Common Zeros***
Gerald Enzner and Philipp Thüne (Ruhr-University Bochum, Germany)
pp. 21-25

P1.6 ***The ABCIT Research Platform***
Kamil Adiloğlu (HörTech gGmbH, Germany); Tobias Herzke (HörTech gGmbH and Cluster of Excellence Hearing4all, Germany); Volker Hohmann (Universität Oldenburg and Cluster of Exellence Hearing4all, Germany); Matthieu Recugnat, Martin Besnard, Teng Huang and Bradford Backus (Oticon Medical, France)
pp. 26-30

P1.7 ***LPC-based speech dereverberation using Kalman-EM algorithm***
Boaz Schwartz and Sharon Gannot (Bar-Ilan University, Israel); Emanuël Habets (International Audio Laboratories Erlangen, Germany)
pp. 30-34

P1.8 ***Post-filter design for speech enhancement in various noisy environments***
Kenta Niwa (NTT Media Intelligence Laboratories, Japan); Yusuke Hioka (University of Auckland, New Zealand); Kazunori Kobayashi (NTT Media Intelligence Laboratories, Japan)
pp. 35-39

P1.9 ***The Single- and Multichannel Audio Recordings Database (SMARD)***
Jesper Kjær Nielsen (Aalborg University & Bang & Olufsen, Denmark); Jesper Rindom Jensen, Søren Holdt Jensen and Mads Græsbøll Christensen (Aalborg University, Denmark)
pp. 40-44

P1.10 ***Estimation of the Common Part of Acoustic Feedback Paths in Hearing Aids using Iterative Quadratic Programming***
Henning Schepker and Simon Doclo (University of Oldenburg, Germany)
pp. 45-49

P1.11 ***Speech dereverberation with convolutive transfer function approximation using MAP and variational deconvolution approaches***
Ante Jukić (University of Oldenburg, Germany); Toon van Waterschoot (KU Leuven, Belgium); Timo Gerkmann and Simon Doclo (University of Oldenburg, Germany)
pp. 50-54

Lunch Break

K3: Keynote 3

Exploiting Structure: New Models and Algorithms for Source Separation, Optimization and Deep Learning

A cornerstone of machine learning, signal processing, and numerical optimization models and algorithms is the identification and exploitation of structure. Structural assumptions can make ill-posed problems like under-determined source separation well defined, structural realities such as complex source models and explaining away effects can make such problems computationally intractable, and subtle hidden structure, once uncovered, can make the same problem tractable once again. In this talk, I'll begin by describing approximate inference algorithms for source separation that exploit the inherent structure of the problem to efficiently search over literally trillions of states and realize super-human multi-talker speech separation and recognition. I'll then talk about the integral role of structure regularization when defining and learning state-of-the-art neural networks for automatic speech recognition, and discuss neural network architectures that make it feasible to train neural networks with millions of neurons. Finally, I'll spend some time characterizing the surprisingly broad class of functions for which Hessian-vector products can be efficiently computed, and describe how such methods have recently been utilized to efficiently solve high-dimensional structure-learning and compression problems.

Speech and sensing at NXP Software

Coffee Break

D2: Demonstration 2

The ABCIT Research Platform

P2: Poster Session 2

P2.1 ***An improved non-intrusive intelligibility metric for noisy and reverberant speech***
João Felipe Santos (INRS-EMT, Canada); Mohammed Senoussaoui (INRS, Canada); Tiago Falk (INRS-EMT, Canada)
pp. 55-59

P2.2 ***Wave-domain canceling of residual echo with subspace tracking***
Satoru Emura and Hitoshi Ohmuro (NTT, Japan)
pp. 60-64

P2.3 ***A robust howling detection algorithm based on a statistical approach***
Joachim Flocon-Cholet (Orange Labs & Université Rennes 1, France); Julien Faure and Alexandre Guérin (Orange Labs, France); Pascal Scalart (University of Rennes, France)
pp. 65-69

Group Photo

O1: Selected Papers

O1.2 ***Identification of Surface Acoustic Impedances in a Reverberant Room Using the FDTD Method***
Niccolò Antonello (KU Leuven & FP7-PEOPLE Marie Curie Initial Training Network DREAMS, Belgium); Toon van Waterschoot and Marc Moonen (KU Leuven, Belgium); Patrick A Naylor (Imperial College London, United Kingdom)
pp. 114-118

O1.3 ***Statistical Modelling of Multichannel Blind System Identification Errors***
Felicia Lim and Patrick A Naylor (Imperial College London, United Kingdom)
pp. 119-123

O1.4 ***Multichannel dereverberation for hearing aids with interaural coherence preservation***
Sebastian Braun and Matteo Torcoli (International Audio Laboratories Erlangen, Germany); Daniel Marquardt (University of Oldenburg, Germany); Emanuël Habets (International Audio Laboratories Erlangen, Germany); Simon Doclo (University of Oldenburg, Germany)
pp. 124-128

Coffee Break

D3: Demonstration 3

Intelligent Microphone for Stress-free Communication

P3: Poster Session 3

P3.1 ***Spatial Perception of Virtual X-Y Recordings***
Konrad Kowalczyk (Fraunhofer Institute for Integrated Circuits IIS, Germany); Alexandra Craciun (International Audio Laboratories Erlangen, Germany); Christian Dachmann (Fraunhofer IIS, Germany); Emanuël Habets (International Audio Laboratories Erlangen, Germany)
pp. 129-133

P3.2 ***Characterisation and modelling of non-linear loudspeakers***
Leela Gudupudi (EURECOM, France); Christophe Beaugeant (Intel, France); Nicholas Evans (EURECOM, France)
pp. 134-138

P3.3 ***Joint Dereverberation and Noise Reduction Based on Acoustic Multichannel Equalization***
Ina Kodrasi and Simon Doclo (University of Oldenburg, Germany)
pp. 139-143

P3.4 ***A Wind-Noise Suppressor Based on Wind-Onset Detection and Spectral Gain Modification***
Masanori Kato and Akihiko K. Sugiyama (NEC Corporation, Japan)
pp. 144-148

P3.5 ***Generalized amplitude interpolation by beta-divergence for virtual microphone array***
Hiroki Katahira (University of Tsukuba, Japan); Nobutaka Ono (National Institute of Informatics, Japan); Shigeki Miyabe, Takeshi Yamada and Shoji Makino (University of Tsukuba, Japan)
pp. 149-153

P3.6 ***Generalization of supervised learning for binary mask estimation***
Tobias May (Technical University of Denmark, Denmark); Timo Gerkmann (University of Oldenburg, Germany)
pp. 154-158

P3.7 ***A Computationally Constrained Optimization Framework for Implementation and Tuning of Speech Enhancement Systems***
Daniele Giacobello (Beats Electronics, USA); Jason Wung (Beats Electronics, LLC & Georgia Institute of Technology, USA); Ramin Pichevar (Beats Electronics LLC, USA); Joshua Atkins (Apple, USA)
pp. 159-163

P3.8 ***Spectral tilt modelling with extrapolated GMMs for intelligibility enhancement of narrowband telephone speech***
Emma Jokinen (International Audio Laboratories Erlangen, Germany); Ulpu Remes, Marko Takanen and Kalle Palomäki (Aalto University, Finland); Mikko Kurimo (Aalto University, Denmark); Paavo Alku (Aalto University, Finland)
pp. 164-168

P3.9 ***Near-field Source Extraction Using Speech Presence Probabilities for Ad hoc Microphone Arrays***
Maja Taseska (International Audio Laboratories Erlangen, Germany); Shmulik Markovich-Golan (Bar-Ilan University, Israel); Emanuël Habets (International Audio Laboratories Erlangen, Germany); Sharon Gannot (Bar-Ilan University, Israel)
pp. 169-173

P3.10 ***Multichannel Adaptive Filtering in Compressive Domains***
Karim Helwani (Huawei European Research Center, Germany); Herbert Buchner (University of Cambridge, United Kingdom)
pp. 174-177

P3.11 ***Numerical Formulae for TOA-based Microphone and Source Localization***
Trung-Kien Le and Nobutaka Ono (National Institute of Informatics, Japan)
pp. 178-182

Lunch Break

Social Event

Visit to the island and Abbaye de Lérins

Lerins- known in ancient times as Lérina - is an archipelago composed of two islands: St Marguerit and St Honorat located in the bay of Cannes, about 15 minutes by boat from Juan-les-Pins.
The smallest and most aside - St. Honorat - is a lush garden surrounded by the turquoise waters of the Mediterranean Sea. For over 16 centuries St. Honorat has hosted a monastery and is today the home of the Cistercian Congregation of the Immaculate Conception. This community - made of 20 monks - is living modestly and peacefully in this oasis of silence.
The Lérins Abbey host regular "retreats" in its monastic hospitality for people seeking serenity. The vineyard offers wine production in an unspoiled environment over the crystal clear Mediterranean waters.

Banquet

The banquet will be held after the visit to the island and Abbaye de Lérins at the seaside restaurant Bijou Plage. Overlooking the island of Lérins, Cap d'Antibes and the Palm Beach Point (Cannes), Bijou Plage offers you an exceptional and relaxing view.

K4: Keynote 4

Dereverberation for Professional Audio and Consumer Electronics

The concept of Illusonic's dereverberation algorithms involves a number of microphone elements, capturing samples of the sound field at various points of the recording space. A reverberation estimate is derived and used to suppress the reverberation in frequency subbands. Professional audio requires only slight dereverberation, but high signal quality. In this case, a second microphone element is added to a conventional shot-gun (interference tube) microphone design and dereverberation with beamforming is applied. Whereas in consumer electronics, the use of one, two, or three consumer grade omni-directional microphone elements is aimed at suppressing reverberation and noise more aggressively. In this case, the algorithms have to cope with inter-microphone mismatch, high microphone/electronic noise levels, and limited frequency responses.

D4: Demonstration 4

Real-Time Speech Source Separation with the Independent Vector Analysis Algorithm on the TI TMS320C6713

P4: Poster Session 4

Coffee Break

P5: Poster Session 5

pp. 248-252

P5.4 ***A new structure for acoustic echo cancellation in double-talk scenario using auxiliary filter***
Mahfoud Hamidia (University of Science and Technology Bab Ezzouar & USTHB, Algeria); Abderrahmane Amrouche (USTHB, Algeria)
pp. 253-257

P5.5 ***Multiple source localisation in the spherical harmonic domain***
Christine Evers, Alastair H. Moore and Patrick A Naylor (Imperial College London, United Kingdom)
pp. 258-262

P5.6 ***Acoustic modeling based on Early-to-Late Reverberation Ratio for robust ASR***
Marco Matassoni, Alessio Brutti and Piergiorgio Svaizer (Fondazione Bruno Kessler, Italy)
pp. 263-267

P5.7 ***Relaxed disjointness based clustering for joint blind source separation and dereverberation***
Ito Nobutaka (NTT, Japan); Shoko Araki and Takuya Yoshioka (NTT Communication Science Laboratories, Japan); Tomohiro Nakatani (NTT Corporation, Japan)
pp. 268-272

P5.8 ***Reverberant Audio Source Separation using Partially Pre-trained Nonnegative Matrix Factorization***
Mahmoud Fakhry (University of Trento & Fondazione Bruno Kessler - irst, Italy); Piergiorgio Svaizer (Fondazione Bruno Kessler, Italy); Maurizio Omologo (Fondazione Bruno Kessler - irst, Italy)
pp. 273-277

P5.9 ***Investigation of Self-Masking Effects for the Evaluation of In-Car Communication Systems***
Anne Theiß (Christian-Albrecht-Universität zu Kiel, Germany); Gerhard Schmidt (CAU Kiel, Germany)
pp. 278-282

P5.10 ***Single Channel Noise Reduction based on an Auditory Filterbank***
Steffen Kortlang (University of Oldenburg, Germany); Stephan Ewert (Medizinische Physik, Universit¨at Oldenburg, Germany); Timo Gerkmann (University of Oldenburg, Germany)
pp. 283-287

P5.11 ***Optimal Binaural LCMV Beamformers for Combined Noise Reduction and Binaural Cue Preservation***
Daniel Marquardt (University of Oldenburg, Germany); Elior Hadad and Sharon Gannot (Bar-Ilan University, Israel); Simon Doclo (University of Oldenburg, Germany)
pp. 288-292

Lunch Break

K5: Keynote 5

Melody Extraction from Polyphonic Music Signals

P6: Poster Session 6

Kollmeier (Medizinische Physik, Universit¨at Oldenburg, Germany); Stefan Goetze (Fraunhofer IDMT, Germany)
pp. 332-336

P6.10 ***Time-Frequency Constraints for Phase Estimation in Single-Channel Speech Enhancement***
Pejman Mowlaee (Graz University of Technology (TU Graz) & Signal Processing and Speech Communication Laboratory, Austria); Rahim Saeidi (Aalto University, Finland)
pp. 337-341

P6.11 ***2.5D Sound Field Reproduction in Higher Order Ambisonics***
Wen Zhang and Thushara D. Abhayapala (Australian National University, Australia)
pp. 342-346

I2: Closing Address / Best Student Paper Award