

Contents

Keynotes

Pathological Speech Analyses: From Classical Machine Learning to Deep Learning

Juan Rafael Orozco-Arroyave, University of Antioquia, Colombia

Statistical Signal Processing and Machine Learning for Speech Enhancement

Timo Gerkmann, Universität Hamburg, Department of Informatics, Signal Processing Research Group

Conversational AI in Production Cars

Christophe Couvreur, Cerence, Merelbeke, Belgium

Session I (Talks): Speech Recognition and Synthesis

- 01 Two-Dimensional Embeddings for Low-Resource Keyword Spotting Based on Dynamic Time Warping 9**
Kevin Wilkinghoff, Alessia Cornaggia-Urrigshardt and Fahrettin Gözköz
- 02 Multi-Head Fusion Attention for Transformer-Based End-to-End Speech Recognition 14**
Timo Lohrenz, Patrick Schwarz, Zhengyang Li and Tim Fingscheidt
- 03 Federated Learning in ASR: Not as Easy as You Think 19**
Wentao Yu, Jan Freiwald, Sören Tewes, Fabien Huennemeyer and Dorothea Kolossa

Session I (Poster): Speech Recognition and Synthesis

- 04 A Comparative Pronunciation Mapping Approach Using G2P Conversion for Anglicisms in German Speech Recognition 24**
Julia Pritzen, Michael Gref, Christoph Schmidt and Dietlind Zühlke
- 05 Bilingual I-Vector Extractor for DNN Hybrid Acoustic Model Training in German Speech Recognition Systems 29**
Yao Wang, Michael Gref, Oliver Walter and Christoph Andreas Schmidt
- 06 New Restricted Boltzmann Machines and Deep Belief Networks for Audio Classification 34**
Paul M. Baggenstoss
- 07 A Lightweight Neural TTS System for High-quality German Speech Synthesis 39**
Prachi Govalkar, Ahmed Mustafa, Nicola Pia, Judith Bauer, Metehan Yurt, Yigitcan Özer and Christian Dittmar
- 08 Automatic Speech Recognition for Dementia Screening Using ILSE-Interviews 44**
Ayimmisagul Ablimit and Tanja Schultz

Session II (Talks): Localisation, Tracking and Spatial Reproduction

- 09 On Source-Microphone Distance Estimation Using Convolutional Recurrent Neural Networks** 49
Tobias Gburrek, Joerg Schmalenstroerer and Reinhold Haeb-Umbach
- 10 On the Use of Additional Microphones in Binaural Cue Adaptation**..... 54
Sebastian Nagel and Peter Jax
- 11 Microphone Utility-based Weighting for Robust Acoustic Source Localization in Wireless Acoustic Sensor Networks** 59
Michael Günther, Andreas Brendel and Walter Kellermann

Session II (Poster): Localisation, Tracking and Spatial Reproduction

- 12 Sound Source Localisation Using Neural Networks with Circular Binary Classification** 64
Magnus Schäfer and Leonie Geyer
- 13 2D Acoustic Source Localisation Using Decentralised Deep Neural Networks on Distributed Microphone Arrays** 69
Stijn Kindt, Alexander Bohlender and Nilesh Madhu
- 14 Data-Dependent Initialization for ECM-Based Blind Geometry Estimation of a Microphone Array Using Reverberant Speech**..... 74
Klaus Brümman, Daniel Fejgin and Simon Doclo
- 15 Binaural Speaker Localization Based on Front/Back-Beamforming and Modulation-Domain Features** 79
Semih Agcaer and Rainer Martin

Session III (Talks): Speech Enhancement and Separation

- 16 Plosive Enhancement Using Phase Linearization and Smoothing** 84
Tal Peer, Klaus-Johan Ziegert and Timo Gerkmann
- 17 Speaker-conditioned Target Speaker Extraction Based on Customized LSTM Cells** 89
Ragini Sinha, Marvin Tammen, Christian Rollwage and Simon Doclo
- 18 Mixed Analog-digital Speech Communication for Underwater Applications** 94
Tim Owe Wisch and Gerhard Schmidt

Session III (Poster): Speech Enhancement and Separation

- 19 Speeding Up Permutation Invariant Training for Source Separation** 99
Thilo von Neumann, Christoph Boeddeker, Keisuke Kinoshita, Marc Delcroix and Reinhold Haeb-Umbach
- 20 Comparison of Generalized Sidelobe Canceller Structures Incorporating External Microphones for Joint Noise and Interferer Reduction**..... 104
Wiebke Middelberg and Simon Doclo

21	Feedback Cancellation for IP-based Teleconferencing Systems	109
	<i>Stefan Kühl, Carlotta Anemüller, Christiane Antweiler, Florian Heese, Patrick Vicinus and Peter Jax</i>	
22	Joint Reduction of Ego-noise and Environmental Noise with a Partially-adaptive Dictionary	114
	<i>Huajian Fang, Guillaume Carbajal, Stefan Wermter and Timo Gerkmann</i>	
23	Beam-specific System Identification	119
	<i>Mhd Modar Halimeh, Christian Hofmann and Walter Kellermann</i>	
24	Low-Complexity Multichannel Wiener Filtering Using Ambisonic Warping	124
	<i>Ashay Sathe, Adrian Herzog and Emanuel A. P. Habets</i>	
25	A Comparison and Combination of Unsupervised Blind Source Separation Techniques	129
	<i>Christoph Boeddeker, Frederik Rautenberg and Reinhold Haeb-Umbach</i>	
26	Robust and High Gain Acoustic Feedback Compensation in the Frequency Domain with a Simple Energy-decay Operator	134
	<i>Klaus Linhard, Philipp Bulling, Marco Gimm and Gerhard Schmidt</i>	
27	An Integrated Deep-Clustering Based System for Speaker Count Agnostic Speech Separation	139
	<i>Jean-Marie Lemercier, Leroy Bartel, David Ditter and Timo Gerkmann</i>	
28	Joint Multi-Channel Dereverberation and Noise Reduction Using a Unified Convolutional Beamformer with Sparse Priors	144
	<i>Henri Gode, Marvin Tammen and Simon Doclo</i>	
29	Reinforcement Learning-based Microphone Selection in Wireless Acoustic Sensor Networks Considering Network and Acoustic Utilities	149
	<i>Haitham Afifi, Michael Guenther, Andreas Brendel, Holger Karl and Walter Kellermann</i>	

Session IV (Talks): Medical Applications and Analytical Studies

30	Supervised Speech Representation Learning for Parkinson's Disease Classification	154
	<i>Parvaneh Janbakhshi and Ina Kodrasi</i>	
31	An Objective Evaluation Framework for Pathological Speech Synthesis	159
	<i>Bence Marc Halpern, Julian Fritsch, Enno Hermann, Rob van Son, Odette Scharenborg and Mathew Magimai-Doss</i>	
32	Informed Source Extraction with Application to Acoustic Echo Reduction	164
	<i>Mohamed Elminshawi, Wolfgang Mack and Emanuel A. P. Habets</i>	

Session IV (Poster): Medical Applications and Analytical Studies

33	A Cortical Model for a θ-Oscillator Segmenting Syllables	169
	<i>Harald Höge</i>	
34	Active Acoustic Equalization: Performance Bounds for Time-Invariant Systems	174
	<i>Florian Hilgemann, Johannes Fabry and Peter Jax</i>	

35	The Effect of Surprisal on Articulatory Gestures in Polish Consonant-to-Vowel Transitions: A Pilot EMA Study	179
	<i>Jacek Kudera, Lauri Tavi, Bernd Möbius, Tania Avgustinova and Dietrich Klakow</i>	

36	A Data Generation Framework for Acoustic Drone Detection Algorithms	184
	<i>Nikita Jarocky, Sebastian Urrigshardt and Frank Kurth</i>	

Session V (Talks): Quality of Speech and Speech Communication Systems

37	Predicting Conversational Quality from Simulated Conversations with Transmission Delay	189
	<i>Thilo Michael and Sebastian Möller</i>	

38	Impact of a Speaker Head Rotation on the Far-end Listening Situation	194
	<i>Tobias Hübschen, Rasool Al-Mafrachi and Gerhard Schmidt</i>	

39	Towards Non-Intrusive Prediction of Speech Recognition Thresholds in Binaural Conditions	199
	<i>David Hülsmeier, Christopher F. Hauth, Saskia Röttges, Paul Kranzusch, Jana Roßbach, Marc René Schädler, Bernd T. Meyer, Anna Warzybok and Thomas Brand</i>	

Session V (Poster): Quality of Speech and Speech Communication Systems

40	A Database for Research on Detection and Enhancement of Speech Transmitted over HF links	204
	<i>Jens Heitkaemper, Joerg Schmalenstroeer, Joerg Ullmann, Valentin Ion and Reinhold Haeb-Umbach</i>	

41	Intelligibility Prediction of Speech Reconstructed From Its Magnitude or Phase	209
	<i>Tal Peer and Timo Gerkmann</i>	

42	Acoustic Ambiance Simulation Using Orthogonal Loudspeaker Signals	214
	<i>Anton Namenas and Gerhard Schmidt</i>	

43	Assessment of Listening Effort for Various Telecommunication Scenarios	219
	<i>Jan Reimes</i>	