2011 Joint Workshop on Hands-free Speech Communication and Microphone Arrays

(HSCMA 2011)

Edinburgh, United Kingdom
30 May – 1 June 2011

IEEE Catalog Number: CFP11HSM-PRT
ISBN: 978-1-4577-0997-5

Program

Monday, May 30

W: Welcome and Opening Remarks

K1: Keynote 1

A Versatile Speech Front-End for Telecommunication and Speech Recognition - The Last One Mile: Implementation Issues for a Better Product
Akihiko K. Sugiyama (NEC Corporation, Japan)

MB1: Break

T1: Talks 1 - Distant Speech Recognition

Channel Selection based on Multichannel Cross-Correlation Coefficients for Distant Speech Recognition
Kenichi Kumatani (Disney Research, Pittsburgh, USA); John McDonough (Carnegie Mellon University, USA); Jill Lehman (Carnegie Mellon University, USA); Bhiksha Raj (Carnegie Mellon University, USA)
pp. 1-6

Discriminative approach to dynamic variance adaptation for noisy speech recognition
Marc Delcroix (NTT Communication Science Laboratories, Japan); Shinji Watanabe (NTT Communication Science Laboratories, Japan); Tomohiro Nakatani (NTT Corporation, Japan); Atsushi Nakamura (NTT Communication Science Laboratories, Japan)
pp. 7-12

Extension of the REMOS Concept to Frequency-Filtering-Based Features for Reverberation-Robust Speech Recognition
Roland Maas (University of Erlangen-Nuremberg, Germany); Martin Wolf (Universitat Politècnica de Catalunya, Spain); Armin Sehr (University of Erlangen-Nuremberg, Germany); Climent Nadeu (UPC, Spain); Walter Kellermann (University Erlangen-Nuremberg, Germany)
pp. 13-18

Theoretical analysis of parametric blind spatial subtraction array and its application to speech recognition performance prediction
Ryoichi Miyazaki (Nara Institute of Science and Technology, Japan); Hiroshi Saruwatari (Graduate School of Information Science, Nara Institute of Science and Technology, Japan); Ryo Wakisaka (Nara Institute of Science and Technology, Japan); Kiyohiro Shikano (Graduate School of Information Science,

ML: Lunch

P1: Posters 1

Speech Measurements Using a Laser Doppler Vibrometer Sensor: Application to Speech Enhancement
Yekutiel Avargel (AudioZoom Ltd, Israel); Israel Cohen (Technion, Israel)
pp. 109-114

Non-linear Spectro-temporal Modulations for Reverberant Speech Recognition
Marco Matassoni (Fondazione Bruno Kessler, Italy); Hari Krishna Maganti (Foundazione Bruno Kessler - Center for Information Technology -IRST, Italy); Maurizio Omologo (Fondazione Bruno Kessler - irst, Italy)
pp. 115-120

WR: Welcome Reception

Tuesday, May 31

K2: Keynote 2

Model-based approaches to handling additive noise in reverberant environments
Mark Gales (University of Cambridge, United Kingdom); Y Wang (University of Cambridge, United Kingdom)
pp. 121-126

TB1: Break

T2: Talks 2 - Microphone arrays

Towards Acoustic Self-Localization of Ad Hoc Smartphone Arrays
Marius Hennecke (TU Dortmund University, Germany); Gernot Fink (TU Dortmund University, Germany)
pp. 127-132

A Wave-Domain Model for Acoustic MIMO Systems with Reduced Complexity
Martin Schneider (University for Erlangen-Nuremberg, Germany); Walter Kellermann (University Erlangen-Nuremberg, Germany)
pp. 133-138

Closed-Form Self-Localization of Asynchronous Microphone Arrays
Pasi Pertilä (Tampere University of Technology, Finland); Mikael Mieskolainen (Tampere University of Technology, Finland); Matti S Hämäläinen (Nokia Research Center, Finland)
pp. 139-144

Dereverberation Performance of Rigid and Open Spherical Microphone Arrays: Theory & Simulation

Daniel P Jarrett (Imperial College London, United Kingdom); Emanuel Habets (International Audio Laboratories Erlangen, Germany); Mark Thomas (Imperial College London, United Kingdom); Nikolay D Gaubitch (Imperial College London, United Kingdom); Patrick A Naylor (Imperial College London, United Kingdom)
pp. 145-150

TL: Lunch

D1: Demo session

Low-latency meeting recognition and understanding using distant microphones
Shoko Araki (NTT Communication Science Laboratories, Japan); Takaaki Hori (NTT Communication Science Laboratories, Japan); Takuya Yoshioka (NTT Corporation, Japan); Masakiyo Fujimoto (NTT Communication Science Laboratories, Japan); Shinji Watanabe (NTT Communication Science Laboratories, Japan); Takanobu Oba (NTT Communication Science Laboratories, Japan); Atsunori Ogawa (NTT Corporation, Japan); Kazuhiro Otsuka (NTT, Japan); Dan Mikami (NTT, Japan); Marc Delcroix (NTT Communication Science Laboratories, Japan); Keisuke Kinoshita (NTT Communication Science Laboratories, Japan); Tomohiro Nakatani (NTT Corporation, Japan); Atsushi Nakamura (NTT Communication Science Laboratories, Japan); Junji Yamato (NTT, Japan)
pp. 151-152

A Speech-Based Conversation System for Accessing Agriculture Commodity Prices in Indian Languages
Gautam Varma Mantena (IIIT Hyderabad, India); Rajendran S. (IIIT Hyderabad, India); Rambabu B. (IIIT Hyderabad, India); Suryakanth Gangashetty (IIIT Hyderabad, India); Yegnanarayana B. (International Institute of Information Technology, Hyderabad, India); Kishore Prahallad (IIIT Hyderabad, India)
pp. 153-154

Positioning System for Mobile Terminals Using a Microphone Array Network as an Intuitive Interface
Shimpei Soda (Kobe University, Japan); Koji Kugata (Kobe University, Japan); Tomoya Takagi (Kobe University, Japan); Hiroki Noguchi (Kobe University, Japan); Shintaro Izumi (Kobe University, Japan); Masahiko Yoshimoto (Kobe University, Japan); Hiroshi Kawaguchi (Kobe University, Japan)
pp. 155-156

IllumiSense: Context Sensitive Illumination
Marius Hennecke (TU Dortmund University, Germany); Christian Kleine-Cosack (TU Dortmund University, Germany); Gernot Fink (TU Dortmund University, Germany)
pp. 157-158

A Practical Beamformer-Postfilter System for Adaptive Speech Enhancement in Non-Stationary Noise Environments
Tobias Wolff (Nuance Communications, Germany); Markus Buck (Nuance Communications, Germany)

P2: Posters 2

Approach

Emanuel Habets (International Audio Laboratories Erlangen, Germany); Jacob Benesty (INRS-EMT, University of Quebec, Canada)

pp. 191-195

Multi-Style Training of HMMs with Stereo Data for Reverberation-Robust Speech Recognition

Armin Sehr (University of Erlangen-Nuremberg, Germany); Christian Hofmann (University of Erlangen-Nuremberg, Germany); Roland Maas (University of Erlangen-Nuremberg, Germany); Walter Kellermann (University Erlangen-Nuremberg, Germany)

pp. 196-200

An Analysis of Nonstationary Variance Estimates in the Maximum Negentropy Beamformer

Barbara Rauch (Saarland University, Germany); Friedrich Faubel (Saarland University, Germany); Dietrich Klakow (Saarland University, Germany)

pp. 201-206

Incorporating localisation cues in a fragment decoding framework for distant binaural speech recognition

Ning Ma (University of Sheffield, United Kingdom); Jon Barker (University of Sheffield, United Kingdom); Heidi Christensen (University of Sheffield, United Kingdom); Phil Green (University of Sheffield, United Kingdom)

pp. 207-212

CD: Conference Dinner (The Caves)

Wednesday, June 1

K3: Keynote 3

Environmental sound recognition and classification

Daniel P W Ellis (Columbia University, USA)

WB: Break

T3: Talks 3 - Microphone arrays and applications

Statistical Method to Identify Key Anthropometric Parameters in HRTF Individualization

Mengqiu Zhang (Australian National University, Australia); Rodney Andrew Kennedy (Australian National University, Australia); Thushara D. Abhayapala (Australian National University, Australia); Wen Zhang (Process Science and Engineering, CSIRO, Australia)

pp. 213-218

A microphone array system integrating beamforming, feature enhancement, and spectral mask-based noise estimation

Takuya Yoshioka (NTT Corporation, Japan); Tomohiro Nakatani (NTT Corporation, Japan)

pp. 219-224

Single and Multichannel Enhancement of Distant Speech using characteristics of Speech Production

Yegnanarayana B. (International Institute of Information Technology, Hyderabad, India); Guruprasad S (IIIT Hyderabad, India); S. R. Mahadeva Prasanna (Indian Institute of Technology Guwahati, India); Suryakanth Gangashetty (IIIT Hyderabad, India)

pp. 225-230

CT: Closing Talk

HSCMA: Hands-free Sound Capture and Microphone Array in Kinect

Ivan Tashev (Microsoft Research, USA)

CR: Closing remarks

WL: Lunch