



SAPA - SCALE Conference 2012



Salon Ballroom, Hilton Portland
7-8 September 2012, Portland, OR,
USA

<http://www.sapaworkshops.org/2012>

- [Home](#)
- [People](#)
- [Dates](#)
- [Program](#)
- [Submit](#)
- [Registration](#)
- [Formatting instructions](#)
- [SAPA Workshops](#)

2012-08-22: Registration is open, and the preliminary program has been posted.

Papers are solicited for the SAPA-SCALE conference, an ISCA-supported event to be held as a satellite to [Interspeech 2012](#) in Portland, USA, September 2012. Following the successful [Workshops on Statistical and Perceptual Audition](#) (SAPA 2004, 06, 08, 10), the event this year is organized jointly with the [Speech Communication with Adaptive Learning \(SCALE\)](#) consortium. The principal objective of the conference is to bring together researchers addressing perceptually motivated speech and audio processing tasks with the tools of statistical signal processing and machine learning. The themes of the conference are:

- Statistical models for speech and audio processing motivated by human perception
- Developing the commonalities between speech recognition and synthesis to provide richer and more sophisticated models for speech
- Adaptive learning approaches to speech and audio signal processing and their incorporation into statistical models

This will be a two-day, single-track conference with an informal atmosphere structured to promote discussion. There will be keynotes from leading researchers in addition to a limited number of oral presentations chosen for breadth and provocation. All participants will be actively engaged in the effort to broaden perspectives and foster novel research directions and interesting variants on current approaches.

Papers describing new research and concepts are solicited on, but not limited to, the following topics:

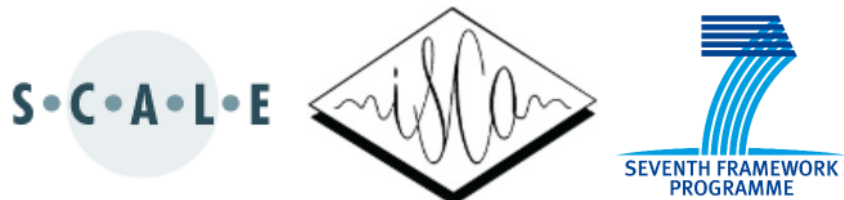
- Generalized audio and speech analysis
- Audio scene analysis and classification
- Music analysis
- Signal separation
- Automatic and Human speech recognition
- Speech synthesis
- Multi-channel analysis

In all cases, preference will be given to papers that clearly align with the themes of the conference. Manuscripts must be between 4 and 6 pages long, in standard Interspeech double-column format. Accepted papers will be published as a SAPA-SCALE conference proceedings. Papers must be received by 21 April 2012 (two weeks after the Interspeech deadline). The results of the paper review will be posted by 3 June 2012 (same as Interspeech).

Organizers:

- Paris Smaragdis, University of Illinois at Urbana-Champaign
- Bhiksha Raj, Carnegie Mellon University
- Dan Ellis, Columbia University
- Steve Renals, University of Edinburgh
- Simon King, University of Edinburgh
- Dietrich Klakow, Universitat des Saarlandes
- Herve Bourlard, IDIAP

Sponsors



The SCALE project receives EC research funding in the context of the EU's 7th framework programme.



SAPA - SCALE Conference 2012



Salon Ballroom, Hilton Portland
7-8 September 2012, Portland, OR,
USA

<http://www.sapaworkshops.org/2012>

- 🌐 [Home](#)
- 🌐 [People](#)
- 🌐 [Dates](#)
- 🌐 [Program](#)
- 🌐 [Submit](#)
- 🌐 [Registration](#)
- 🌐 [Formatting instructions](#)

- 🌐 [SAPA Workshops](#)

People

Organizers:

Bhiksha Raj

Associate Professor
Carnegie Mellon University
Pittsburgh, PA 15213
bhiksha@cs.cmu.edu

Paris Smaragdis

Assistant Professor
University of Illinois, Urbana-Champaign
Urbana, IL 61801
paris@illinois.edu

Daniel Ellis

Associate Professor
Columbia University
New York, NY 10027
dpwe@ee.columbia.edu

Herve Bourlard

Director and Professor
IDIAP
Martigny, Switzerland
bourlard@idiap.ch

Simon King

Professor and Director
University of Edinburgh Centre for Speech Technology
Research
Edinburgh, UK
Simon.King@ed.ac.uk

Dietrich Klakow

Professor
Universitat des Saarlandes
Saarbrücken, Germany

Dietrich.Klakow@lsv.uni-saarland.de

Steve Renals

Professor and Director

University of Edinburgh Institute of Language, Cognition,
and Computation

Edinburgh, UK

s.renals@ed.ac.uk

Dan Ellis <dpwe@ee.columbia.edu>

Last update: Thu Aug 25 11:35:46 AM EDT 2011





SAPA - SCALE Conference 2012



Salon Ballroom, Hilton Portland
7-8 September 2012, Portland, OR,
USA

<http://www.sapaworkshops.org/2012>

- [Home](#)
- [People](#)
- [Dates](#)
- [Program](#)
- [Submit](#)
- [Registration](#)
- [Formatting instructions](#)

- [SAPA Workshops](#)

Technical Program

The workshop will be held in Hilton Portland Salon Ballroom. This is downstairs in the Hilton Executive Tower building, which is at 545 SW Taylor (kitty-corner to the main hotel), then go down the stairs to the right of the entrance.

Click on each title to retrieve the corresponding paper, or you can download all papers in a zip file: [sapascale2012papers.zip](#) (16MB).

Polls Page

Friday September 7th

0930-0945 Welcome & introduction

0945-1045 Keynote 1:
[Human sound perception - what can we learn from it when developing audio analysis algorithms?](#)
Tuomas Virtanen (Tampere University of Technology)

1045-1105 *break*

1105-1130 [Pitch Estimation Using Mutual Information](#)
(pp. 1-4)
Majid Mirbagheri, Yanbo Xu, Shihab Shamma (University of Maryland College Park)

1130-1155 [Establishing some principles of human speech production through two-dimensional computational models](#)
(pp. 5-10)
Mauro Nicolao, Roger K. Moore (University of Sheffield)

1155-1220 [A Spectral Envelope Estimation Method Based on F0-Adaptive Multi-Frame Integration Analysis](#)
(pp. 11-16)
Tomoyasu Nakano, Masataka Goto (AIST)

1220-1315 *lunch*

1315-1340 [Cochlear Implant-like Processing of Speech Signal for Speaker Verification](#)

(pp. 17-21)

Cong-Thanh Do, Claude Barras (LIMSI-CNRS/Universite Paris-Sud)

1340-1405 Speech intelligibility enhancement for HMM-based synthetic speech in noise

(pp. 22-27)

Cassia Valentini-Botinhao, Junichi Yamagishi, Simon King (University of Edinburgh)

1405-1430 A Generalized Stein's Estimation Approach for Speech Enhancement Based on Perceptual Criteria

(pp. 28-33)

Sunder Ram Krishnan, Chandra Sekhar Seelamantula (Indian Institute of Science)

1430-1455 Non-Stationary Signal Processing and its Application in Speech Recognition

(pp. 34-39)

Zoltán Tüske, Friedhelm R. Drepper, Ralf Schlüter (RWTH Aachen University)

1455-1515 *break*

1515-1540 Joint Uncertainty Decoding with Unscented Transform for Noise Robust Subspace Gaussian Mixture Models

(pp. 40-45)

Liang Lu, Arnab Ghoshal, Steve Renals (University of Edinburgh)

1540-1605 Hierarchical Hybrid Language Models for Open Vocabulary Continuous Speech Recognition using WFST

(pp. 46-51)

M. Ali Basha Shaik, David Rybach, Stefan Hahn, Ralf Schlüter, Hermann Ney (RWTH Aachen University)

1605-1630 Template-based ASR using Posterior features and Synthetic References: comparing different TTS systems

(pp. 52-57)

Serena Soldo, Mathew Magimai.-Doss, Hervé Bourlard (Idiap Research Institute)

1630-1655 Explicit Duration Modelling in HMM-based Speech Synthesis using a Hybrid Hidden Markov Model-Multilayer Perceptron

(pp. 58-63)

Kalu U. Ogbureke, João P. Cabral, Julie Carson-Berndsen (University College Dublin)

Saturday September 8th

0945-1045 Keynote 2:
Speech processing in human auditory cortex
Nima Mesgarani (UCSF)

1045-1105 *break*

1105-1130 Language Identification using Spectro-Temporal Patch features
(pp. 110-113)

Kamal Sahni, Pranay Dinghe, Rita Singh, Bhiksha Raj (CMU)

1130-1155 Joint Detection and Localization of Multiple Speakers using a Probabilistic Steered Response Power

(pp. 68-73)

Youssef Oualil, Mathew Magimai.-Doss, Friedrich Faubel, Dietrich

Klakow (Saarland University/Idiap)

- 1155-1220 Structured Sparse Coding for Microphone Array Location Calibration
(pp. 74-79)
Afsaneh Asaei, Bhiksha Raj, Hervé Bouchard (Idiap/CMU)
- 1220-1315 *lunch*
- 1315-1340 Inharmonic Speech: A Tool for the Study of Speech Perception and Separation
(pp. 114-117)
Josh McDermott, Dan Ellis, Hideki Kawahara (NYU/Columbia/Wakayama)
- 1340-1405 Multi-Channel Speech Separation with Soft Time-Frequency Masking
(pp. 86-91)
Rahil Mahdian Toroghi, Friedrich Faubel, Dietrich Klakow (Saarland University)
- 1405-1430 Smoothing Speech Trajectories by Regularization
(pp. 92-97)
Heyun Huang, Louis ten Bosch, Bert Cranen, Lou Boves (Radboud University Nijmegen)
- 1430-1450 *break*
- 1450-1515 Data-driven Speech Representations for NMF-based Word Learning
(pp. 98-103)
Joris Driesen, Jort F. Gemmeke, Hugo Van hamme (KU Leuven)
- 1515-1540 Spectro-Temporal Features with Distribution Equalization
(pp. 104-109)
Samuel K. Ngouoko M., Britta Wrede, Martin Heckmann (Bielefeld University/Honda Research)
- 1540-1605 Log-normal matrix factorization with application to speech-music separation
(pp. 80-85)
Takuya Yoshioka, Sakaue Daichi, (NTT Communication Science Laboratories)
- 1605-1615 *Conclusion & farewell*

(not presented)

Dimensionality Reduction of Large TDOA Vectors for Speaker Diarization
(pp. 64-67)
Deepu Vijayaseenan, Fabio Valente (Universität des Saarlands/Idiap)

Dan Ellis <dpwe@ee.columbia.edu>

Last update: Wed Sep 19 09:54:45 PM EDT 2012

