

Speech Separation and Comprehension in Complex Acoustic Environments by Humans and Machines – A Workshop

November 4-7, 2004, Montreal, Québec, Canada

Organizing Committee: *Pierre Divenyi (Chair), Nat Durlach, Dan Ellis, and DeLiang Wang*

PROGRAM

Thursday Nov. 4, 17:00 – 20:00

17:00-18:00 **Welcome refreshments**

18:00-20:00 **The Speech in Speech Separation – A Primer**

Chair: Pierre Divenyi (EBIRE and VA Medical Center, Martinez, CA)

Steven Greenberg (Technical University of Denmark)

- A Multi-Tier Theoretical Framework for Understanding Spoken Language

Richard Stern (Carnegie Mellon University)

- Signal processing for sound separation and robust representation
-

Friday November 5, 8:00 – 12:30

8:00-10:00 **Brief Overview of Areas of Speech Separation: Objectives, Advantages, and Problems**

Chair: Pierre Divenyi (EBIRE and VA Martinez)

Adelbert Bronkhorst (TNO Human Factors, Soesterberg, The Netherlands)

- Human single-channel and spatial performance

Martin Cooke (Sheffield University, U.K.)

- Machine separation using human models

Te-Won Lee (University of California San Diego)

- Blind Machine Separation

Robert Zatorre (Montreal Neurological Institute, McGill University, Canada)

- Separation in the Auditory Cortex

Dan Ellis (Columbia University)

- Recognition and learning as tools of machine separation

Nathaniel Durlach (Boston University and MIT)

- Exploiting Human-Machine (HM) Collaboration to Achieve Superior Source-Separation-and-Comprehension

10:00-10:30 **Break**

10:30-11:10 **Summary of the HANSE Workshop on Speech Segregation**

Presenter: Christopher Darwin (University of Sussex, U.K.)

11:10-12:00 **Introduction of and Precipitation by Students/Postdocs**

Chair: Barbara Shinn-Cunningham (Boston University)

12:00-13:00 Lunch

13:30-16:00 Variations in Design and Performance of Sensing Arrays

Chairs: Steven Colburn (Boston University) and

Te-Won Lee (University of California San Diego)

Simon Carlile (University of Sydney, Australia)

- High Frequencies in Speech and Talker Localization and Segregation

Douglas Brungart (Wright-Patterson Air Force Research Labs)

- Symmetrical and Asymmetrical Factors Influencing Speech Segregation by Human Listeners

Andrew Oxenham (MIT)

- Role of Fundamental Frequency in Streaming/Segregation

DeLiang Wang (Ohio State University)

- Monaural and Binaural Speech Separation

Joseph Desloge (Sensimetrics Corp.)

- Directional Multimicrophone Arrays: A Spatial-filtering Approach to Source Separation

Te-Won Lee (University of California San Diego)

- ICA-based Techniques for Single Channel and Multichannel Speech Separation

James Flanagan (Rutgers University)

- Spatial Selectivity for Speech Separation

16:00-16:30 Break

16:30-18:00 What Kinds of Knowledge about Humans Are Useful for Designing Machine Systems?

Chair: Alain de Cheveigné (CNRS and IRCAM, Paris, France)

Hynek Hermansky (IDIAP, Martigny, Switzerland)

- Learning About Hearing from Speech Data

Doug Brungart (Wright-Patterson Air Force Research Labs)

- Optimizing the Performance of Multitalker Speech Displays

Rhodri Cusack (MRC-CBU, Cambridge, U.K.)

- Perceptual Organization and Attention: Relationship and Neural Architecture

19:30-22:00 Evening Poster Session

Francis Bach (University of California Berkeley, with Michael Jordan)

- Blind One-Microphone Speech Separation: A Spectral Learning Approach

Jana Eggink (Sheffield University, with Guy J. Brown)

- Using Instrument Recognition to Extract Melodies from Complex Audio

Mounya Elhilali (University of Maryland, with Shihab Shamma)

- Adaptive Cortical Model for Auditory Streaming and Monaural Speaker Separation

Frederick J. Gallun (Boston University, with Christine R. Mason and Gerald Kidd, Jr.)

- Binaural Release from Informational Masking in a Speech Recognition Task

Brian Gygi (EBIRE)

- Parsing the Blooming Buzzing Confusion

Manuel Reyes-Gomez (Columbia University, with Nebojsa Jojic and Dan Ellis)

- Deformable Spectrograms

Madhu Shashanka (Boston University, with Barbara Shinn-Cunningham and Martin Cooke)

- Effects of Reverberant Energy on Statistics of Speech

Soundarajan Srinivasan (Ohio State University, with Nicoleta Roman and DeLiang Wang)

- On Binary and Ratio Time-Frequency Masks for Robust Speech Recognition

John Hershey (Microsoft Research, with Trausti Kristjansson and Zhengyou Zhang)

- Model-Based Fusion of Bone and Air Sensors for Speech Enhancement and Robust Speech Recognition

Maria Chait (University of Maryland, with S. Greenberg, T. Arai, J. Simon, and D. Poeppel)

- Multi Resolution Analysis of Speech

Maria Chait (University of Maryland, with D. Poeppel and J. Simon)

- Binaural versus monaural processes in pitch extraction: evidence from the human auditory system

Saturday November 6, 8:00-12:00

8:00-10:00 Combating the Reverberation Problem

Chair: Barbara Shinn-Cunningham (Boston University)

Barbara Shinn-Cunningham (Boston University)

- How humans cope in natural settings

Martin Cooke (Sheffield University, U.K.)

- How speech is corrupted by reverberation

DeLiang Wang (Ohio State University)

- Effects of reverberation on pitch, onset/offset, and binaural cues

Joseph Desloge (Sensimetrics Corp.)

- Multi-microphone Source Separation in Reverberant Environments

10:00-10:30 Break

10:30-12:00 Integration of Different Machine Approaches

Chairs: DeLiang Wang (Ohio State University), Dan Ellis (Columbia University)

Lawrence Saul (University of Pennsylvania)

- Machine Learning and Auditory Scene Analysis

Les Atlas (University of Washington)

- Modulation Spectral Filtering: A New Tool for Acoustic Signal Separation

Lucas Parra (City University of New York)

- Acoustic Source Separation with Microphone Arrays

Dan Ellis (Columbia University)

- Integrating CASA information with other signal separation techniques

12:00-13:00 Lunch

Saturday November 6, 13:30-18:00

13:30-16:00 Exploiting Human-Machine (HM) Collaboration to Achieve Superior Source-Separation-and-Comprehension (SS&C) Systems

Chair: Nathaniel Durlach (Boston University and MIT)

Nat Durlach (Boston University and MIT)

- Introduction

Sumit Basu (Microsoft Research) and Joseph Desloge (Sensimetrics Corp.)

- Interfacing with the machine

Misha Pavel (Oregon Health and Science University)

- Augmented cognition: Optimal attention allocation

Betty Tuller (Florida Atlantic University)

- Adaptation and perceptual learning

Abeer Alwan (University of California Los Angeles)

- Voice differences

Barbara Shinn-Cunningham (Boston University)

- Spatial differences

16:00-16:30 Break

16:30-18:00 Neural Bases of Generalized "Scene" Analysis

Chair: Shihab Shamma (University of Maryland)

Steven McAdams (McGill University, Canada)

- Scene analysis in music

Max Riesenhuber (Georgetown University)

- Scene analysis in vision

Henning Scheich (Leibniz Institut for Neurobiology, Magdeburg, Germany)

- Scene analysis in the cortex

18:30-20:00 Workshop reception

Sunday November 7, 8:00-12:00

8:00-9:30 Evaluation of Speech Separation, Corpus Development

Chair: Dan Ellis (Columbia University)

Dan Ellis (Columbia University)

- Introduction

Alex Acero (Microsoft Research)

- The Speech Recognition Experience

Te-Won Lee (University of California San Diego)

- Evaluating Blind Source Separation Systems

Douglas Brungart (Wright-Patterson Air Force Research Labs)

- Evaluating Human Performance

Martin Cooke (Sheffield University, U.K.)

- Speech Tasks for Human/Machine Comparisons

Dan Ellis (Columbia University)

- Towards a Common Speech Separation Evaluation

9:30-10:00 *Break*

10:00-12:00 Design of Superior Systems and Research Programs Needed to Achieve Them

Chairs: Pierre Divenyi (EBIRE and VA Martinez)

Nathaniel Durlach (Boston University and MIT)

DeLiang Wang (Ohio State University)

Dan Ellis (Columbia University)

12:00-13:00 *Farewell lunch*