The 2nd International Workshop on Machine Listening in Multisource Environments

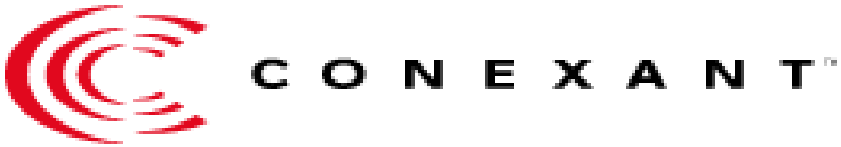
Emmanuel Vincent[1], Jon Barker[2], Shinji Watanabe[3], Jonathan Le Roux[3], Francesco Nesta[4] and Marco Matassoni[5]

[1]Inria Nancy – Grand Est, France
[2]Department of Computer Science, University of Sheffield, UK
[3]Mitsubishi Electric Research Labs, Boston, MA, USA
[4]Conexant Systems, Newport Beach, CA, USA
[5]FBK-Irst, Trento, Italy

Vancouver, June 1, 2013

Acknowledgements

Financial support



Industrial Board

Masami Akamine
Carlos Avendano
Li Deng
Erik McDermott
Gautham Mysore
Atsushi Nakamura
Peder A. Olsen
Trausti Thormundsson
Daniel Willett

Admin support: Édith Blin, Gillian Callaghan

Acknowledgements

Scientific Committee

Alberto Abad
Masami Akamine
Shoko Araki
Carlos Avendano
Heidi Christensen
Dan Ellis
Valentin Emiya
Ramon Fernandez Astudillo
Cédric Févotte
Federico Flego
Phil Green
Reinhold Häb-Umbach
Mark Hasegawa-Johnson
John R. Hershey
Denis Jouvet
Walter Kellermann
Mathieu Lagrange
Erik McDermott
Gautham Mysore
Atsushi Nakamura
Tomohiro Nakatani
Peder A. Olsen
Nobutaka Ono
Alexey Ozerov
Bhiksha Raj
Steven J. Rennie
Hiroshi Sawada
Armin Sehr
Paris Smaragdis
Trausti Thormundsson
Daniel Willett

Workshop motivation

- Machine Listening researchers often identify themselves by specific application domains, for example,
  - speech recognition people,
  - music transcription and analysis people,
  - acoustic event detection people,
  - source separation people.
- This emphasises the differences between these domains instead of promoting shared solutions to shared problems.
- One particularly challenging problem is robustness in multisource environments.
- We hope this workshop can bring communities together.

What is a multisource environment?

- By multisource environment we intend the following conditions:
  - Environments containing multiple sources of sound.
  - The sound sources are typically individually localised in space.
  - The activity level of the sources is changing over time.
  - The sound sources may be static or moving.
  - There may be some prior expectations, but many critical parameters are unknown (e.g. number of sources).
- These conditions are normal in everyday listening environments – and yet they are often treated as a special case.
- Multisource conditions lead to challenging tasks, e.g.,
  - Recognising distant microphone speech in everyday settings.
  - Transcribing a string quartet from a live recording.
  - Detecting a specific bird call in a woodland recording.
  - Enhancing a target speaker while suppressing multisource noise background.

Workshop programme

9:00	Welcome
9:10	Keynote 1: Steven J. Rennie (IBM T.J. Watson Research Center)
10:00	Break
10:20	Overview of the 2nd CHiME Challenge
10:50	Oral session 1: challenge papers
12:10	Lunch
13:40	Poster session
15:40	Break
16:00	Oral session 2: challenge-related paper
16:20	Keynote 2: Daniel P.W. Ellis (Columbia University)
17:10	Plenary discussion
17:50	Closing

Notes for presenters

Slides

Please upload your slides onto the computer
during the morning break.

Timing

Talks should be 15 minutes
with 5 minutes for questions and handover.

Posters

Please hang your poster during the morning break.

Plenary discussion

Feedback is essential for the sustainability of CHiME, e.g.:

- What went right or wrong?
- Why so few non-challenge papers?
- Why few participants in Track 2?
- How to improve the instructions and the tools?
- How to improve the reproducibility of the results?
- How to gain more outcomes?
- Which data and task come next? Who is interested in organising?

Think about it. Discuss it today. Email us feedback.

Other discussion topics and announcements are welcome.

Best Paper Award selected by the Industrial Board

Best Paper Award selected by the Industrial Board

Jürgen Geiger, Felix Weninger, Antti Hurmalainen, Jort Gemmeke, Martin Wöllmer, Björn Schuller, Gerhard Rigoll and Tuomas Virtanen

The TUM+TUT+KUL approach to the 2nd CHiME Challenge: Multi-stream ASR exploiting BLSTM networks and sparse NMF