Welcome!
Workshop Motivation

- Machine Listening lacks a coherent community.
- 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.
- Segregation emphasises the differences between these domains ... this impedes progress on shared problems.
- One particularly challenging problem is robustness in multisource environments.
- We hope this workshop can bring communities together to share important insights.
What is a ‘Multisource Environment’?

By ‘multisource environment’ we are intending the following,

- 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).

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.
**The Challenge of Multisource Environments**

- **Multisource conditions are normal** in everyday listening environments – and yet they are often treated as a special case.
- The **human auditory system** is highly adept at dealing with multisource conditions,
  - Human ability has been much studied by the *Hearing* and *Computational Hearing* communities.
  - But there is still no deep understanding of how the human ear really works.
  - Computational models (e.g. CASA systems) remain a long way from human ability – a focus on toy problems.
- Historically, BSS and ASR communities have also focused on simple scenarios... but share a feeling that **the time has come to address real-world problems**.
- Real problems may demonstrate the need for significant re-design as simple systems no longer prove adequate.
Workshop Programme

Morning

9:00-9:10  Welcome and opening remarks
9:10-9:50  Overview of the CHiME Challenge
9:50-10:40  Oral 1: Challenge papers
10:40-11:00  Morning break (tea and coffee)
11:00-12:15  Oral 2: Challenge papers
12:15-13:45  Buffet lunch

Afternoon

13:45-15:45  Poster session
15:45-16:00  Afternoon break (tea and coffee)
16:00-16:50  Oral 3: Multisource event detection and classification
16:50-17:50  Plenary discussion: results and future evaluation
17:50  Closing
Notes for Presenters

- **Slides** - please upload your slides onto the computer during the morning break.
- **Timing** - oral presentations should be 20 minutes with 5 minutes for questions and handover.
- **Posters** - please hang your poster during the morning break.
Special Issue of Computer Speech and Language

Speech Separation and Recognition in Multisource Environments

Important Dates

- November 30, 2011: Paper submission
- March 30, 2012: First review
- May 30, 2012: Revised submission
- July 30, 2012: Second review
- August 30, 2012: Camera-ready submission

One of the chief difficulties of building device-to-microphone speech recognition systems for use in ‘everyday’ applications is that the noise background is typically ‘multisource’. A speech recognition system designed to operate in a busy home, for example, must contend with competing noise from televisions and radios, children playing, vacuum cleaners, and outdoor noises from open windows. Despite their complexity, multisource environments can be modelled as a mixture of sources that can be learnt and exploited using advanced source separation, machine learning and speech recognition techniques such as those presented at the 1st International Workshop on Machine Learning in Multisource Environments (CHIME-2011).

This special issue solicits papers describing advances in speech separation and recognition in multisource noise environments, including theoretical developments, algorithms or systems.

Examples of topics relevant to the special issue include:
- Multiple speaker localization, beamforming and source separation,
- Speech-inspired approaches to multisource processing,
- Background noise tracking and modeling,
- Noise-robust speech decoding,
- Model combination approaches to robust speech recognition,
- Datasets, toolkits and other resources for multisource speech separation and recognition.

Submission Instructions:
https://ees.elsevier.com/csl

Once logged in, click on “Submit New Manuscript” then select “Special Issue: Speech Separation and Recognition in Multisource Environments” in the “Choose Article Type” dropdown menu.

Important Dates:
November 30, 2011: Paper submission
March 30, 2012: First review
May 30, 2012: Revised submission
July 30, 2012: Second review
August 30, 2012: Camera-ready submission

Guest Editors:
Jim Baker, University of Sheffield, UK
Emmanuel Vincent, INRIA, France
Feedback is essential for the sustainability of the challenge

The Questionnaire

- You’ll find it in your packs.
- Please complete before 4.00 pm.
- No need to add name unless you wish!
- Place completed questionnaire in the box.
Acknowledgements

Financial support:

Organising Committee: Jon Barker, Dan Ellis, Phil Green, John Hershey, Walter Kellermann, Hiroshi Okuno, Emmanuel Vincent.

Technical Committee: Heidi Christensen, Reinhold Häb-Umbach, Walter Kellermann, Ning Ma, Atsushi Nakamura, Francesco Nesta, Hiroshi Okuno, Alexey Ozerov, Armin Sehr.

CHiME Challenge support: Ning Ma.

Admin support: Gillian Callaghan (Sheffield), Constanza Vannocci (PLS Educational, Italy).

Authors: 80 researchers contributing to today’s papers; Attendees: 68 delegates.