CRAC

A workshop on

Consistent and reliable acoustic cues for sound analysis

September 2nd 2001
Aalborg, Denmark

Listeners are capable of sound processing tasks such as robust speech recognition in acoustic environments which defeat their automated counterparts. Part of this performance disparity may be due to different developmental conditions. Automated systems have typically attempted to ameliorate the effect of noise on representations and algorithms devised for clean conditions. Listeners, growing up outside the anechoic chamber, are immediately faced with the reality of perception in a world where competing sound sources are not simply there to be suppressed.

This workshop[1] focuses on approaches to robust sound processing which recognise the existence of such noisy realities. One theme of techniques developed in recent years has been to replace classification based on complete representations with a search for reliable acoustic cues followed by an identification process which accommodates sparse representations. The initial stage can be accomplished by techniques ranging from standard noise estimation algorithms to full-blown computational auditory scene analysis systems using pitch and location models. The subsequent identification process has been tackled using missing data and multistream processing.

Although the dominant application area is robust speech recognition, the workshop brings together research on music and non-speech audio, in addition to work on representations inspired by the auditory system. This diversity flows from the growing belief that progress in a wide range of application areas relies on confronting and solving the problem of general-purpose sound understanding faced by listeners.

CRAC Organizing Committee

Dan Ellis, Columbia University (co-chair)
Martin Cooke, Sheffield University (co-chair)

Frédéric Berthommier, Institut de la Communication Parlée, Grenoble
Andrzej Drygajlo, École Polytechnique Fédérale de Lausanne
Phil Green, Sheffield University
Andrew Morris, Institut Dalle Molle d'Intelligence Artificielle Perceptive
Hiroshi Gitchang Okuno, Kyoto University

---

1. The workshop grew out of the EU LTR Project RESPITE (28149) - Recognition of Speech by Partial Information Techniques.

Contents

Schedule

08:00	Registration
09:00	Session 1: **Perception & Auditory Models** (Phil Green)
10:30	Coffee break
11:00	Session 2: **Music & General Audio Analysis** (Dan Ellis)
12:30	Lunch
14:00	Session 3: **Missing-Data Speech Recognition** (Martin Cooke)
15:30	Coffee break
16:00	Session 4: **Approaches to Handling Noisy Speech** (Hiroshi G. Okuno)
17:30	Close

Per session:

Overview (session chair)	10 minutes
Lecture 1	15 minutes
Lecture 2	15 minutes
Posters	45 minutes