

The (IM)2 Newsletter
HEVs delivers SMA, Seniors meet, DARPA funds, MDM hosts a workshop on annotation,...

The Speech Group at ICSI
lead by Prof. Nelson Morgan, interacts with (IM)2 in a number of ways

Issue N° 15
June 2003
Editor: Jean-Albert Ferrez
www.im2.ch/nl nl@im2.ch

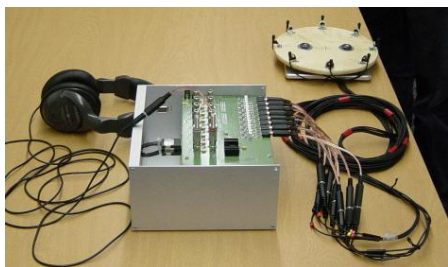


The (IM)2 Newsletter

Every month the (IM)2 Newsletter brings you the latest and hottest scientific and administrative news about the (IM)2 NCCR and related topics

Delivery of the first SMA

On May 19, IDIAP and HEVs celebrated the delivery of the first Small Microphone Array (SMA) prototype. This joint development took place in the framework of the (IM)2.RTMAP white paper (Real Time Microphone Array Processing). The objective of the SMA is to obtain an integrated device in which some of the microphone array preprocessing can be performed in real time before the output is sent to a standard PC (through a FireWire interface) for further processing.



Inside view of the SMA hooked-up to an 8-way circular microphone array.

The SMA features a TigerSHARC DSP board, a codec board, eight audio inputs and eight audio outputs, RS232, USB and FireWire interfaces. The infotronics group at HEVs has developed custom hardware to integrate all these elements as well as all necessary APIs and interfaces controlling the SMA from a PC.



The SMA, the DSP and the codec boards (on the table), and the joint HEVs/IDIAP team who worked on it: from left to right: Olivier Masson, Christian Briguët, Charles Praplan, Gilbert Maître, Joseph Moerschell, Darren Moore, Iain McCowan (missing: Jean-Pierre Gherig).

While this first prototype is fully functional and can be produced in small series, work will now continue in the direction of adding more capabilities, as well as in using the SMA as the basis for further research. In particular, the beam-forming

techniques used to improve the quality of the recording, to locate and track speakers, and to perform speaker segmentation, are now being ported to the DSP.

(IM)2 seniors meeting

A meeting of seniors involved in (IM)2 took place on Thursday May 15 at IDIAP. Most of the day was dedicated to informally discussing the potential White Paper project proposals. This gave a chance for the participants to better identify strengths and weaknesses in their ideas, lead to better awareness of other team's objectives, and increased the potential for fruitful collaboration. The result will obviously materialize in better proposals, even if participation in this event will not have an influence on the actual selection process.

Thanks to (IM)2, IDIAP raises more funding from the US through DARPA



The exchange program set up in the framework of (IM)2, which tightens the links between (IM)2 partners and US institutions, mainly through ICSI/Berkeley, produces its first results. Not only do several (IM)2 researchers (PhD students or post-docs) benefit from the opportunity, the program also fosters a significant increase of collaboration between IDIAP and US research programs. In late 2002, the US Defense Advanced Research Projects Agency (DARPA) granted IDIAP a small research project in robust speech recognition. As of July 1, 2003, and acknowledging the quality of the collaboration, DARPA significantly increased this funding from approx. 60 K\$/year to 270 K\$/year to work in the framework of the EARS project (Effective, Affordable, Reusable, Speech-to-text systems, see www.darpa.mil/iao/EARS.htm). This is definitely a major success for (IM)2, and we want to thank all our collaborators for their involvement in this effort.

Advisory Boards Review

The (IM)2 Scientific and Industrial Advisory Boards met for the first time on February 13-14 in Martigny. Based on the presentations and discussions, they issued a report and recommendations available from the (IM)2 local web site.

(IM)2.MDM workshop: "Which annotation for which application?"

On May 7, ISSCO hosted a full-day workshop on annotations. (IM)2.MDM invited contributions from six other IPs, along with presentations from its own participants. The goal was to encourage communication between IPs on a specific topic: which phenomena do we target in the meeting data, and what is their relevance to the targeted applications? The various talks listed below, as well as the discussions, showed that the present IPs were focusing on complementary phenomena in the data, with an encouraging convergence at the level of applications.

The workshop grouped about 20 participants. The slides of the talks can be accessed on the (IM)2.MDM website, in the partners' area at www.issco.unige.ch/projects/im2/mdm/. The following talks were given:

- Beyond scripted meetings: next round of meeting recordings (Daniel Gatica-Perez, IDIAP, (IM)2.SP & (IM)2.SA)
- Annotation tools and formats (Pierre Wellner, IDIAP, (IM)2.IP)
- Database and interfaces for multimodal dialogues (Fabrice Jouanot & Oleksandr Drutskyy, EPFL, (IM)2.DS)
- Shallow dialogue processing (Andrei Popescu-Belis & Alexander Clark, UniGe, (IM)2.MDM)
- Meeting modelling for enhanced browsing (Stéphane Marchand-Maillet, UniGe, (IM)2.IIR)
- Document annotation and the Smart Meeting Minutes application (Denis Lalanne, UniFr, (IM)2.DI)
- Annotation of argumentation structure in meetings (Vincenzo Pallotta & Hatem Ghorbel, EPFL, (IM)2.MDM)
- A multimodal query interface on dialogue databases (Martin Rajman, EPFL, (IM)2.MDM)

Events

2003 Summer Institute 6-8.10.03

The second edition of the (IM)2 Summer Institute will take place on October 6-7-8 in Martigny. The full announcement, including call for talks and social program will follow, but you can already set those days aside.

The Speech Group at the International Computer Science Institute (ICSI)

The International Computer Science Institute is one of the few fully independent, nonprofit basic research institutes in the United States, and is closely affiliated with (though not part of) the University of California in Berkeley, California. ICSI has a strong international program, including visitor exchanges with Germany, Finland, Spain, and Switzerland. There are 5 major scientific departments at ICSI: Networking (focused on the basic architecture of the Internet); Algorithms (the design, analysis and application of algorithms to problems such as bioinformatics and communication in distributed systems like the Web); the Berkeley Center for the Information Society (an interdisciplinary effort to study the impact of information technology on societies, at an international level); Artificial Intelligence (computational models and techniques based on natural intelligence); and Speech. The Speech group conducts research in the areas of algorithms, architectures, and systems for speech and audio signal processing and pattern recognition, with a particular emphasis on spoken language processing.

The Speech group at ICSI interacts with (IM)2 in a number of ways. First, while the visitor exchange was originally a broad program covering all areas of computer science studied at ICSI, the current program is focused on (IM)2 concerns - visitors to ICSI work on research topics that are of specific interest and ultimate benefit to the NCCR, such as speaker segmentation or automatic sound unit learning. Secondly, through the (IM)2 White Paper mechanism, ICSI researchers participate directly in topics of

mutual interest, such as the current effort on tools, data, and studies related to spoken language from meetings. Thirdly, ICSI and IDIAP collaborate on current and proposed European projects, many of which are likely to provide leverage to the Swiss funding in related research areas. Finally, through the mechanism of US funding sources, ICSI is sponsoring research in Switzerland that complements large efforts in the US, particularly in the area of conversational speech recognition.

In the current (IM)2-specific research, ICSI is completing the final preparations for delivery of a corpus that includes 75 meetings collected at ICSI (averaging almost an hour in length, for a total of over 70 hours). These meetings are simultaneously recorded on both close-talking and far-field microphones, and are fully transcribed at the word level. ICSI researchers also work on the manual labeling of dialog acts (e.g., statements vs questions vs back-channels) and the automatic detection of such events. There is an effort to define and automatically locate "hot spots" or, in some sense, "interesting" periods in the meeting, which could be very useful for future work on summarization. Finally, there is an ongoing study of the applicability of high performance microarchitectures for speech recognition in meetings, such as might be implemented in future high-end PDAs.

The Speech Group is composed of over 20 researchers, currently consisting of 12 staff, 6 graduate students, and 4 visiting researchers (see accompanying picture). For 15 years it has had a strong focus on the recognition of speech under "natural" con-

ditions. This has sometimes meant a study of speech recognition under noisy or reverberant conditions, but recently it also has had an emphasis on speech from natural conversations. Over the last 5 years both interests have led to our current corpus and research in the area of spoken language in meetings, which has been a key component in our collaboration with (IM)2 and IDIAP. We have been chosen to lead a DARPA-funded collaboration with IDIAP and four US labs (SRI, University of Washington, Columbia, and OGI) as part of the program called Effective Affordable Reusable Speech-to-text, or "EARS". Our mission in this collaboration is to experiment with approaches that depart significantly from the mainstream in order to help achieve the goals of the EARS program. These goals include the extremely challenging problem of reducing word error rates on conversational speech to under 10% in the next few years. We also participate strongly in other aspects of the EARS project, including work on segmentation of speech streams into parts associated with each speaker (in collaboration with IDIAP), the development of a Mandarin Chinese recognizer, and the annotation and subsequent classification of so-called "metadata" in conversational speech, such as indicators for various kinds of disfluencies. We also work (in the context of the German research program "SmartKom" on a range of subsystems for spoken dialog systems, for instance, to retrieve traveler information. In addition, we continue our long-term interest in developing algorithms to improve performance under difficult acoustic conditions, often in close collaboration with IDIAP.

Ultimately, one of the closest ties between ICSI and (IM)2 is through their Directors. Hervé Bourlard was one of the earliest visitors to ICSI (arriving before its official inauguration in 1988), and afterwards began joint research with ICSI's current Director, Nelson Morgan. Together they have had a long and fruitful collaboration that is now substantially expanded through the great efforts of the other personnel at IDIAP and ICSI. Professor Bourlard is also a Trustee of the ICSI governing Board, which gives him an additional window into ICSI activities.

nmm

