i m
INTERACTIVE
MULTIMODAL
### INFORMATION
MANAGEMENT

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

(IM)2 in Bern

On April, 18th, (IM)2 was in Bern for the mid-term presentation of the 2000-2003 message from the Federal Council on the encouragement of teaching, research and technology. There were speeches from Federal Councilors R. Dreyfuss and P. Couchepin, and various other representatives. Actual examples of the current Swiss strategy were presented. SNF chose (IM)2 to demonstrate the NCCR concept. The (IM)2 booth, pictured below, attracted the full attention of Federal Councilor Ruth Dreyfuss, Head of the Federal Department of Home Affairs.

—

—
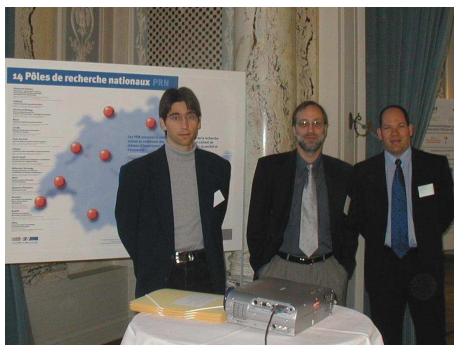
(IM)2 status update

Four months after its start on January 1st, the inauguration of (IM)2 is the first occasion to look back on the early days of the NCCR and recall the actions taken so far.

Projects

As of late April, seven of the nine Individual Projects (IP) have been accepted in their definitive forms, slightly updated from the original propositions of March 15, 2000. This process, important for the focus and cohesion of the NCCR, was not always easy, and even implied some personnel changes. The result, though, is a much tighter network of teams with well defined tasks and responsibilities who will work together toward a common goal.

The other research planning tool within (IM)2, based on the submission of White Papers, has seen 16 propositions coming from almost all (IM)2 partners. These propositions are being reviewed, and the most promising will start this summer.

Aside from these administrative issues, work has already started on the core topics of (IM)2. The full specifications of the Smart Meeting Minutes application are being agreed upon, including the various data formats that will allow smooth and efficient collaboration among the teams.

Structures

The (IM)2 management has set up structures and communication tools, which include the (IM)2 web site www.im2.ch, several mailing lists, and the monthly Newsletter. A brochure for a more general presentation of (IM)2, its members and its goals is also in preparation.

The exchange agreement with the International Computer Science Institute (ICSI) in Berkeley is effective, potential candidates will find all the details on the (IM)2 web page. Industrial relations and technology transfer will be carried out in collaboration with CIMTEC and CCSO.

Last but not least, the construction of the new IDIAP building is in progress, and should be finished by mid-June, The new facilities will offer a conference room with video-conference, a meeting room fully equipped for audio-video recordings, and office space for about 50 people.

(IM)2 Inauguration

The (IM)2-NCCR on Interactive Multimodal Information Management will be officially inaugurated on May, 4th in Martigny. The ceremony features talks by:

M. Pierre Crittin,
President of Martigny and of IDIAP Foundation Council

M. Jean-René Fournier,
State Councilor of Valais

Prof Hervé Bourlard,
Director of IDIAP and (IM)2

Prof Martin Hasler,
Member of SNF Research Committee

Prof Stefan Catsicas,
EPFL Vice-President for Research

M. Thierry Gattlen,
Director of SportAccess Kudelski SA

and musical intermezzi performed by students of the Ecole supérieure de musique Tibor Varga. The day continues with a visit to two famous attractions in the city of Martigny, the Château de la Batiaz and the Fondation Pierre Gianadda.

Events

IEEE NNSP'02, Martigny 4–6.09.02
The 2002 IEEE International Workshop on Neural Networks for Signal Processing is organised by Prof. Hervé Bourlard (IDIAP). For more details, see eivind.imm.dtu.dk/nnsp2002

DAGM'02, Zürich   16–18.09.02
This year's German Conference on Pattern Recognition is organised by Prof. Luc Van Gool (Vision Group, ETHZ). The focus is on image processing and computer vision, but other aspects related to (IM)2 will also be covered. For more details, see dagm02.vision.ee.ethz.ch.

INTERACTIVE MULTIMODAL INFORMATION MANAGEMENT

The Speech Processing Group at IDIAP

The overall goals of the IDIAP speech processing group are to research and develop robust recognition and understanding techniques for realistic speaking styles and acoustic conditions, as well as robust speaker verification and identification techniques. This includes advanced research activities, maintenance of language resources for the training and testing of recognition systems, and development of real-time prototypes. The group has been involved in speech research projects for several years and is today at the leading edge of technology. The IDIAP Speech Processing group is also involved in numerous national and European collaborative projects, as well as industrial projects.

People

Lead by Hervé Bourlard, Professor at EPFL and Director of IDIAP and of (IM)2, the group is currently composed of 4 senior researchers, 11 Phd students and regularly welcomes short term visitors. Recently, the group also appointed Dr Pierre Wellner, former Xerox and AT&T scientist and CEO of Spiderphone, Inc (www.spiderphone.com), who will start on August 1, 2002.

Part of the Speech Processing Group at IDIAP: Front: Darren Moore, Johnny Mariéthoz, Hemant Misra, Iain McCowan, Hervé Bourlard, Todd Stephenson. Back: Mohamed Faouzi Benzeghiba, Itshak Lapidot, Simon Miles Payne, Jaume Escofet Carmona, Andrew Morris, Mathew Magimai Doss, Jitendra Ajmera, Guillaume Lathoud, Shajith Ikbal.

Research Themes

● Automatic recognition of (isolated, continuous, or natural) speech based on phonetic (sub-word) modeling, using spectro-temporal profiles of speech, as well as articulatory parameters.
● Development and improvmeent of different speaker verification systems (voice print authentication) including: text independent and text dependent (text prompted) speaker verification, as well as user-customized password speaker verification. These systems exhibit state-of-the-art performance in regular international NIST (National Institute of Standards and Technology) evaluation campains.
● Development and improvement of state-of-the-art speech recognition systems based on hidden Markov models (HMM).
● Using discriminant artificial neural networks (ANN) to estimate *a posteriori* probabilities. In this regard, IDIAP is recognized as a leader in the use of hybrid HMM/ANN systems, exhibiting several advantages compared to standard HMM approaches.
● Estimation of confidence levels, i.e., attaching a confidence score to each recognized word to indicate how likely the word is correctly recognized. In this context, the problem of detecting out-of-vocabulary words is also investigated.
● Multi-stream and multi-band speech recognition: improving robustness of state-of-the-art systems based on multiple feature streams. This includes the extraction of multiple features from the same input utterance, exhibiting different properties, such as multiple temporal resolutions and/or containing some new, novel, or robust type of information.
● Multi-stream combination: Developing novel methods to combine information generated from multiple experts trained on multi-stream features to improve word recognition and increase robustness of the recognition to corrupting environmental conditions.
● Acoustic change detection and clustering, as required when dealing with large audio and multimedia databases (such as broadcast news and sport videos). In this framework, different approaches are investigated towards automatic segmentation of (multimedia) sound tracks, including, among others, changes in acoustic environments, speaker change detection, speaker identification and tracking, and speech/music discrimination.
● Speaker adaptation: Improving recognition accuracy by automatically adapting (a subset of) the parameters of the recognition system.
● Development and adaptation of efficient software for large vocabulary continuous speech recognition, on different computer platforms (UNIX and Windows).

Application Examples

One of the major applications targeted by IDIAP in speech processing is also at the heart of (IM)2. Efforts toward an **automatic meeting manager** are already taking place in the framework of the M4 European project which aims at processing multiple audio (microphone) streams for structuring, browsing and querying of an archive of automatically analysed meetings. IDIAP has recently acquired a fully featured 24 track audio recording system. Meetings with up to 8 people will be recorded with individual microphones as well as microphone arrays, and the resulting databases will be available to bootstrap the various research activities within (IM)2. This system will soon be upgraded to also accommodate video acquisition.

Other applications include:

Command and control systems, possibly used in noisy environments, e.g., to operate speech enabled phones in cars.

Speech enabled information systems, such as kiosks, desk tablets, and personal data assistants to enable users to find and display current information.

Information retrieval for audio documents, using transcriptions automatically generated by a large-vocabulary speech recogniser to build indexes that can be queried by information retrieval engines for searchable audio archives. The BBC is using such a system to index its daily archives of broadcast news.

hb