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

IDIAP Smart Meeting Room

Since the completion of IDIAP's new building (see previous issue), and in the scope of (IM)2 and of the M4 European project, IDIAP has equipped a meeting room with audio and video acquisition facilities for recording and processing meetings.

The smart meeting room hardware is designed to be capable of :

- recording the speech signal of each participant using both close-talking microphones and table-top microphone arrays.
- recording multiple video cameras, including wide angle and medium-range shots of the participants.

The audio acquisition system consists of 24 microphone channels that are digitised and streamed directly to the hard disk of a PC.

The video acquisition system consists of 3 video channels that are each recorded by separate MiniDV video tape recorders. This ensures that all acquired video is high-quality digital and that full PAL framerate and resolution is preserved.

To facilitate fusion of modalities in some processing tasks, the audio and video systems are synchronised using a master sync signal, and each recorded channel is accurately timestamped.

For full audio-video meeting recordings, the Smart Meeting Room can accommodate 6 meeting particpants in its current configuration. Two cameras each provide a medium-angle front-on view of 3 participants, while the 3rd tripod-mounted camera sits on the table and provides a wider angle view of the entire meeting, including the whiteboard and projector screen. Each meeting participant wears a lapel microphone, and 4 to 8 microphones are used in each of 3 table-top microphone arrays.

For audio-only recordings, the Smart Meeting Room can accommodate up to 12 meeting participants.

Full technical specifications can be found at www.idiap.ch/~mccowan/meeting/.

IDIAP receives grant from HP

In the framework of the **HP Voice Web Initiative**, a philanthropic activity to support research on the convergence of computer speech technologies and the Internet, IDIAP received a grant in the form of several HP workstations and servers. This grant will support research on the MUST project: Multimodal User Authentication for Secured Transactions, see www.idiap.ch/learning/MUST.html. The project aims in particular at improving multimodal biometric authentication support in the torch library, see www.torch.ch.

(IM)2 meets a German cousin

On July 12th, the Director of (IM)2 Prof. Hervé Bourlard and the (IM)2 Program Manager Dr Jean-Albert Ferrez, together with the Director of ICSI Prof. Nelson Morgan were invited to discover the German Research Center for Artificial Intelligence (DFKI, *Deutsche Forschungszentrum für Künstliche Intelligenz*, www.dfki.de). Lead by Prof. Wolfgang Wahlster, DFKI has offices in Saarbrücken and Kaiserslautern, and numbers approximately 200 researchers. DFKI leads many projects in areas related to (IM)2, in particular Smartkom: *Dialog-based Human-Technology Interaction by Coordinated Analysis and Generation of Multiple Modalities*, see www.smartkom.org.

From left to right: Dr Jean-Albert Ferrez, (IM)2, Prof. Hervé Bourlard, (IM)2, Prof. Wolfgang Wahlster, DFKI and Prof. Nelson Morgan, ICSI.

beowulf @ IDIAP

To fullfill the requirements of the many new researchers working on (IM)2 projects, a new 16-node Linux-based computing cluster was installed at IDIAP.

Events

Federal Authorities to visit IDIAP
As part of a three days visit in Martigny, the members of the Control Committee (*Comission de gestion - Geschäftsprüfungskommission*) of the Council of States (*Conseil des Etats - Ständerat*) will be at IDIAP on August 30th.

IEEE NNSP'02, Martigny 4–6.09.02
The 2002 IEEE International Workshop on Neural Networks for Signal Processing is organized 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 organized 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.

The conference will be followed on 19 and 20 September 2002 by a joint workshop to informally discuss the results of several European projects investigating Cognitive Vision processes and systems (COGVIS, COGVISYS, ACTIPRET, DETECT, VISATEC, LAVA, CAVIAR and VAMPIRE). For more details, see cogvis.vision.ee.ethz.ch.

(IM)2 Summer Institute   3–4.10.02
On October 3 and 4, Martigny will host the first (IM)2 internal workshop. This will be the perfect opportunity to look back on the first months of the NCCR, to identify gaps and overlaps in the various research plans, and to tighten the links among the researchers at a time when – hopefully – all positions will have been filled. Further information and registration details will be available on the local pages of the (IM)2 web site.

INTERACTIVE
MULTIMODAL
INFORMATION
MANAGEMENT

The Computer Vision and Multimedia Laboratory at UniGeneva

The Computer Vision and Multimedia Laboratory (CVML, vision.unige.ch), founded in 1989, is headed by Prof. Thierry Pun, Dr. Stéphane Marchand-Maillet and Dr. Slava Voloshynovskiy. The laboratory currently comprises about 15 researchers at various academic levels (research assistant, post-doc, assistant- and full-professor).

The laboratory carries out research in computer vision and multimedia data processing (images, videos, sounds, biological signals). Roughly 75% of CVML research is financed by external, competitive grants (Swiss and European, private and public); research results are described in over 200 refereed scientific publications and in 6 patents. 13 Ph.D. dissertations have been completed (8 more under way), as well as many B.Sc./M.Sc. projects.

Past research directions

Past areas of interest have included image segmentation and grouping, affine and photometric invariance, motion analysis and objects tracking (for e.g. highway traffic surveillance), visual attention and computational neuroscience, object recognition, learning. Amongst the completed practical projects were a pattern classifier for 2D gel electrophoresis images (part of the Melanie system), a public-domain software for image processing (LaboImage), and a machine vision system for agricultural robotics (Potato Operation).

Content-based visual indexing and retrieval (CBVIR)

Research context: strategies are being developed for the efficient indexing and retrieval of documents for multimedia archives. These developments are integrated into the CBVIR system Viper (Visual Information Processing for Enhanced Retrieval, viper.unige.ch).

Major research issues involved: use of inverted files for indexing the data, low- and high-level features extraction, inter-active segmentation, exploitation of relevance feedback, search optimization by pruning, handling of search targets that change along time, long-term learning, data annotation, benchmarking of CBVIR systems.

Main achievements:
● definition and maintenance of MRML, the Multimedia Markup Retrieval Language (www.mrml.net), which is more and more employed in the field;
● creation and maintenance of GIFT, the Gnu Image Finding Tool currently distributed under a General Public License (www.gnu.org/software/gift) and part of Linux distributions (eg SuSE since rel. 7.3);
● leadership of the international CBVIR benchmarking effort, the Benchathlon (www.benchathlon.net).

External support: Swiss NCCR IM2, Swiss NSF, EU IST projects M4 (Multimodal Meeting Manager) and Webkit (Intuitive physical interfaces to the Web). M4 concerns the structuring, browsing and querying of an archive of automatically analyzed meetings that take place in a room equipped with multimodal sensors. Webkit deals with the use of tangible user interfaces for information search and retrieval. A former EU project DVP (Digital Video Production) was related to the creation, archival and retrieval of videos.

Watermarking and stochastic signal processing

Research context: copyright protection, data authentication and tamper proofing mechanisms for images and videos are being developed, in particular using digital watermarks; these developments are integrated into the system Berkut (watermarking.unige.ch). In addition, work is underway regarding various aspects of stochastic signal processing, with applications to e.g. image denoising, compression, and segmentation.

Major research issues involved: spread-spectrum template-based watermarking, multi-resolution content-adaptive watermarking, attacks and benchmarking, authentication and tamper-proofing, steganography, analog watermarking, denoising, joint source-channel compression and coding, segmentation.

Main achievements:
● top results on international benchmarks obtained by our prototypes;

● technology transfer regarding some of these developments to the company DCT (Digital Copyright Technologies);
● creation of the public domain benchmark Checkmark (watermarking.unige.ch/Checkmark), which will eventually become part of the Certimark international benchmark.

External support: Swiss NCCR IM2, Swiss NSF, EU IST project Certimark (Certification for watermarking techniques) aiming at the creation of a public watermark benchmarking and certification tool. Former grants included the EU project JEDI-FIRE (JAVA Flexible Electronic Commerce Firewall and Data Privacy), and the Swiss Priority Research Program for Information and Communication Structures.

Multimodal interaction

Research context: with the aim to complement the visual channel, several forms of human-machine interaction paradigms are being studied: tactile (using e.g. a force-feedback mouse), auditory (by means of auditory icons projected in a virtual 3D sound space), and based on biosignals (such as EEG's). A WWW browser for visually impaired and blind users including text-to-speech and 2D image-to-3D sound conversion has been developed, and extensively tested and validated with the help of blind users. A new project is starting in the area of brain-machine interfaces.

Major research issues involved: auditory-tactile interaction, development of sonic educational games for blind pupils, sonic analysis of drawings and vectors graphics, brain-machine interaction, EEG processing.

Main achievements:
● creation of free educational software packages (games, line drawing analysis) to blind users;
● distribution of Websound, a Web sonification tool allowing audio-haptic exploration of a HTML document (websound.unige.ch);
● development of an Internet audio and text library for blind users (websound.unige.ch/library).

External support: Swiss NCCR IM2, and past grants from the Swiss ABA ("Association pour le Bien des Aveugles et Amblyopes") and from the Swiss Priority Research Program for Information and Communication Structures.

tp