IDIAP PhD student awards

This year’s annual IDIAP Dinner took place in the majestic ballroom of the Grand Hotel des Bains in Lavey and was the opportunity to present two PhD student awards.

Mathew Magamai Doss was presented with the Research Award for 2004. Senior IDIAP staff nominated PhD students based on 5 criteria: Publications, Collaboration, Project involvement, Communication skills, and Autonomy. Mathew was presented with the award due to his successful publication rate, high level of autonomy, and a demonstrated willingness to collaborate with other staff and students. The 2004 Student Paper Award was presented to Guillaume Lathoud. Guillaume’s paper was chosen from a list of several papers nominated by IDIAP’s senior staff. All nominated papers were assessed by external reviewers and feedback was provided based on Originality, Contribution, Technical correctness, and Presentation. The paper "Unsupervised Location-based Segmentation of Multi-party Speech", published in the Proceedings of the 2004 ICASSP-NIST Meeting Recognition Workshop, May 2004, was chosen because of its excellent analysis, novel approach of a little-examined phenomenon, and promising results.

IDIAP would like to thank the following people for their help with the paper review process: Dr Tim Cootes, University of Manchester, UK; Asst Prof. Stephane Renals, University of Edinburgh, UK; Prof. Nelson Morgan, ICSI and University of California, USA; Dr Milind R. Naphade, IBM T J Watson Research Center, NY, USA; Prof. Steve Renals, University of Edinburgh, UK; Prof. Gerhard Rigoll, TUM, Germany; Asst Prof. Amos Storkey, University of Edinburgh, UK; Dr. Timothy Wark, CSIRO Australia.

CTI VoiceInPack: A success story

The CTI project VoiceInPack featuring IDIAP and ETHZ working for KOMODO Entertainment Software on low bit-rate speech coding and transmission for massively online games is now finished and was presented to the CTI officials on December 3rd, together with the first demo of the "One Online" game. The project was recognised by CTI as a large success and picked up to be featured as "success story" with broad media coverage. Work will go on in this area, however, in particular with a patent and licensing agreement between IDIAP and VoicelInfinity Inc., a new startup targeting telephone applications of the VoiceInPack technology.

The Swiss-french TV channel TSR featured the project on Friday, December 3. The online video can be seen from www.tsr.ch/tsr/index.html?siteSect=500001&bcid=0336697&vid=5382623

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.
2006–2009: A Hint of things to come

The original (March 2000) NCCR proposal covered the first phase (2002-2005) and only gave hints of what would come in subsequent years. Time has come to update the objectives and research plans for the second phase (2006-2009). While the details are yet to be discussed among the partners and formalised into a full proposal, the general sketch of things to come has already been defined and agreed upon. We give here an overview of the Individual Projects (IPs) which will implement the vision, building on the outcome of the first phase.

So far, (IM)2 has significantly contributed to the development of many new activities related to multimodal interaction and management of multimedia information systems. It will keep the same focus, directly leveraging upon these results, emphasizing the highest possible quality, cross-disciplinarity, collaboration between the (IM)2 partners, and integration in large international initiatives.

The core of the NCCR activities remain of course research, with focus on Multimodal interfaces, Integration of modalities and coordination among modalities, Meeting dynamics and human-human interaction modeling, and Content abstraction.

**IP1: Database Management and Meeting Analysis**

This IP will be responsible for defining meeting scenarios, identifying possible new (IM)2 database needs, extending infrastructure, etc. It will group all activities in database collection, database annotation and transcription, and database management, annotation standards, and distribution.

**IP2: Audio Processing**

This IP builds on the current speech processing IPSP and will work on audio structuring, audio processing, sound source localization, microphone arrays, speaker recognition, speaker segmentation, speaker verification, spoken language understanding, conversational speech recognition syntactical modeling in conversational speech recognition, metadata extraction, and dialog act modeling.

**IP3: Visual/video Processing**

This IP also builds on an existing IP, Scene Analysis, but will trim its focus down to image processing, image/object representation and modeling image/video indexing, segmentation, grouping and recognition, face and gesture tracking, gesture analysis, face and people detection and recognition document analysis, handwriting recognition.

**IP4: Multimodal Processing and Recognition**

This new IP builds on IP2 and IP3 and is the core of the multimodal activity, featuring algorithms for multi-stream/multi-channel processing, audio-visual tracking, speech recognition, speaker identification, and segmentation. The current brain-machine interface sub project, a particularly challenging application, is part of IP4, but it was also suggested to turn it into an individual project of its own.

**IP5: Multimodal Content Abstraction**

This IP groups all activities related to content abstraction of mono- and multi-modal information. Particular emphasis will be on multimodal content abstraction, information indexing and retrieval, and semantic analysis of different information sources (images, slides, complex printed documents, handwritten notes, text,...). Indexing and retrieval systems for audio and video sequences, and of multi-media (multi-stream) information will be developed.

**IP6: Human-Machine Interaction**

This IP will be in charge of all activities related to the development and evaluation of advanced (multi-modal) interfaces, including meeting browsers suited for highly multimodal and multimedia data sources, wizard-of-oz elicitation of user’s requirements and perception-action loop.

**IP7: Integrated Software and Research Demonstrators**

Last but not least, all activities related to the development of research demonstrators, which are currently spread across multiple IPs, will be coordinated in a new dedicated IP. This will help raise awareness of the growing importance of this part in Phase 2, hopefully attracting more interest from industries.

**Other aspects**

Non R&D activities will also be extended to Phase 2, including the ongoing efforts to properly anchor IDIAP in the Swiss academic framework through EPFL, the ICSI student exchange program, the MLMI series of workshops, the Fellowship for Female researchers, the integration in TheArk, etc. See the previous issues of the (IM)2 Newsletter for more information about these activities.