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

More European projects

In this issue of the (IM)2 Newsletter we follow up on our survey of (IM)2-related European projects in which (IM)2 members are involved. A first series of projects was covered in issue 07 (October 2002). Other projects (DARPA, CTI/KTI,...) will be covered in a future issue.

FGnet

Face and Gesture Recognition Working Group, IST-2000-26434, 2001-2004, involving IDIAP. FGnet is the European working group on face and gesture recognition funded by the IST program. The objectives of FGnet are to encourage development of technology for face and gesture recognition. The network goals are to assist development of face and gesture recognition technology, to create a set of foresight reports defining development roadmaps and future use scenarios for the technology in the medium (5-7 years) and long (10-20 years) term, and to specify, develop and supply resources (eg image sets) supporting these scenarios. More information at www-prima.inrialpes.fr/FGnet.

HASSIP

Harmonic Analysis and Statistics for Signal and Image Processing, involving EPFL-ITS. In the HASSIP project, the emphasis is on developing a new framework for jointly exploring deterministic image non-linear approximation techniques (i.e. unveiling the structure of images) and the associated stochastic models. This will allow a clean modelling of image features in the settings of denoising, compression, communications. More information at www.cmi.univ-mrs.fr/HASSIP.

ISLE

International Standards for Language Engineering, involving UniGE-ISSCO. The ISLE Project is a joint initiative supported by the European Community and the National Science Foundation of the USA. Following the EAGLES Project, ISLE aims at developing standards for natural language engineering, in three main directions: computational lexicons, multimodal human-machine interaction, and evaluation. More information at www.ilc.pi.cnr.it/EAGLES96/isle.

COST-275

COST-275 Action, Biometrics-Based Recognition of People over the Internet, involving IDIAP and EPFL. The main objective of this action is to investigate effective methods for the recognition of people over the Internet, based primarily on voice and facial characteristics, in order to facilitate, protect and promote financial and other services over this growing telecommunication medium. As implied above, the scope of the work also includes investigation into the usefulness of other types of biometrics and their integration into the proposed technology. More information at www.fub.it/cost275.

COST-278

COST-278 Action, Spoken Language Interaction in Telecommunication, involving IDIAP, ETHZ and Eurecom. The main objective of this Action is to create knowledge in several problem areas of spoken language interaction in telecommunications in order to achieve communicative interfaces that provide a natural human-computer interaction through more cognitive, intuitive and robust interfaces, whether monolingual, multilingual or multimodal. More information at www.cost278.org.

ASSAVID

Automatic Segmentation and Semantic Annotation of Sports Video, IST-1999-13082, 2000-2002, involving IDIAP. In the ASSAVID project an exploration is being continued into improvement of techniques for extraction of features from audio visual media, and the depth of annotation that may be achieved by fusing the multiple modes and features extracted. Assavid follows up on the THISL project (Thematic Indexing of Spoken Language) with BBC as the main industrial partner. More information at www.bpe-rnd.co.uk/assavid.

Visit of Léonard Favre

Mr Léonard Favre, Director of Economic Affairs of the Canton du Valais, was at IDIAP for a presentation of the (IM)2 NCCR last November. He showed a strong interest for the various technologies under development and confirmed that CIMTEC was the partner of choice for the technology transfer activities of the NCCR. Several business opportunities are already being evaluated and will hopefully start during the year.

Events

(IM)2 Scientific and Industrial Advisory Boards   13–14.2.03
The first joint meeting of the (IM)2 Scientific and Industrial Advisory Boards will take place in Martigny on February 13 and 14. The two day meeting will start with a general presentation of the NCCR and its objectives, and then focus on three topics: input modalities, multimodal processing, and applications. The members of the Boards are:
Scientific Advisory Board: Prof. Maurice Bellanger, C.N.A.M., France; Prof. Edward Delp, Purdue University, USA; Prof. Renato De Mori, University of Avignon, France; Prof. Bernd Girod, Stanford University, USA; Prof. Fred Jelinek, Johns Hopkins University, USA; Prof. Goesta Granlund, Linkoping University, Sweden; Prof. Thomas Huang, University of Illinois, USA; Prof. Joseph Kittler, University of Surrey, UK.
Industrial Advisory Board: Dr Aldo Bussien, VP of Engineering, Logitech SA, CH; Prof. Gordon Edge, The Generics Group plc, Cambridge, UK; Dr Kari-Pekka Estola, Nokia Research Center, Finland; Dr Pier Carlo Faloti, Consultant, ex CEO of DEC and Oracle Europe, CH; Dr Claude Galand, Director, AT&T Business Services, USA; Dr Denis Gonseth, Director, HPI Holding SA, CH; Dr Bob Liang, Director Media Lab, Intel Research, USA; Dr Tom Malzbender, HP Laboratories, Palo Alto, USA; Dr Christophe Meier, Director, CCSO, CH; Dr Arun Netravali, President, Lucent, USA; Dr Pierre-Yves Saintoyant, Microsoft Research, Cambridge, UK.

INTERACTIVE
MULTIMODAL
INFORMATION
MANAGEMENT

The Database Laboratory at EPFL

Since 1988, Prof. Stefano Spaccapietra heads the Database Laboratory (LBD), whose staff averages at 15-20 members. Research at LBD aims at enriching the functionalities supported by current data management software towards more user-friendly services. A powerful conceptual data modeling approach provides a sound and formal basis for the specification of system-independent services relying on user's perception and concepts. Results achieved so far, in cooperation with Prof. Christine Parent at UNIL-Inforge, include:

• A conceptual data model, MADS, which extends traditional data modeling to harmoniously include concepts for modeling of spatial, temporal, and spatio-temporal data. MADS offers unmatched expressive power, in particular for applications using geographical data, ranging from environmental management to geomarketing.
• An extension of MADS to support multiple representation of real world phenomena (EEC IST MurMur project, 2000-2002). This allows database design to adhere to the diversity of perceptions and requirements from the user community sharing the same database.
• Operational CASE tools that provide users with visual conceptual interfaces, based on MADS, for the definition of database schemas (schema editor) and for query formulation (query editor) on MADS spatio-temporal databases. The tools map conceptual specifications into existing DBMS (Oracle) or GIS (MapInfo, ArcView, Geotask) specifications.
• A database integration methodology supporting alignment or integration of heterogeneous sources without requiring prior translation into a common data model.

LBD also hosts a research group lead by Dr Pearl Pu on human-computer interaction. Research from this group is in the area of personalized search of electronic catalogs for configurable products and services, focusing on modeling and elicitation of user preferences, and human agent interaction models. Advanced fish-eye visualization techniques are developed to improve human-computer communication.

Current research projects

MADS and MADS-based CASE tools
Work on the MADS conceptual data model continues towards adding concepts for description and management of imprecise information. Enrichment of the MADS CASE tools is planned to support multi-level modeling allowing management of very large schemas, to improve visual querying of temporal data and to manage information uncertainty.

Schema evolution and multimedia query languages Within (IM)2 our goal is to design and implement services to support interactive multimodal information management techniques. The targeted services include interfaces and supporting facilities to store and retrieve video and audio data, and the corresponding descriptive metadata (namely, various annotations to describe in particular relevant segmentations of the audio and video files and, as far as possible, their semantic content). Supporting easy and fast metadata evolution along paths that are not necessarily foreseeable in advance is a major challenge for this project, calling for investigation into innovative techniques for schema evolution. In parallel, we investigate concepts and techniques for content-based querying of multimedia databases, focusing on image and video databases. An image content data model and an associated query language have been defined.

Spatial Semantic Services for Mobile Users Research within the Mobile Information and Communication Systems NCCR targets the development of location-based spatial information services by designing and implementing semantic extraction services that dynamically adjust to the information sources available at a given moment in time on the data servers currently accessible by a mobile user. This calls for designing facilities to address ontological issues, including reconciliation, addition and removal of ontologies, semantic extraction from structured as well as unstructured sources, personalization and matching of preformatted queries with currently available data sources.

Ontology Engineering and Federated Services This project aims at developing techniques for building global federated services supporting semantic interoperability across heterogeneous sources including spatio-temporal information. Applications in the area of environmental data management are foreseen. Ontology engineering is a kernel component of the approach.

Semantic Fisheye View for Multimedia Information Retrieval This project develops a novel approach, semantic fisheye view (FEV), for multimedia information access and retrieval. Semantic fisheye view is a visual interface to a large information store. It organizes users' information needs by a set of task-based queries represented by semantic structures. Fisheye view algorithms use graphical techniques to dynamically emphasize retrieved data based on relevance to a query and user interaction.

LBD is routinely involved in international activities, holding responsibilities in working groups, conferences and journals. For further details, please refer to lbd.epfl.ch.

ss