The (IM)2 Newsletter
First part of an overview of the (IM)2-related projects in which (IM)2 members are involved.

The DIVA research group at the University of Fribourg, lead by Prof. Rolf Ingold, is dealing with multimedia engineering.

Issue N° 07
October 2002
Editor: Jean-Albert Ferrez
www.im2.ch/nl nl@im2.ch

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

This issue of the (IM)2 Newsletter is the first in a series dedicated to the presentation of (IM)2-related projects in which (IM)2 members are involved. We start with a selection of European projects in the Information Society Technologies (IST) programme of the 5th Framework, see www.cordis.lu/ist. More IST projects, as well as other projects (DARPA, CTI/KTI,...) will be covered in future issues.

M4

MultiModal Meeting Manager, IST-2001-34485, 2002-2005, involving IDIAP, EPFL/ITS and UniGe. M4 concerns the structuring, browsing and querying of an archive of automatically analyzed meetings that take place in a room equipped with multimodal sensors. It is important to note here that this project, initiated just before the official start of (IM)2, fits perfectly with the application scenario defined in (IM)2, and will further build upon the Smart Meeting Room available at IDIAP. More information at www.dcs.shef.ac.uk/spandh/projects/m4.

CIMWOS

Combined IMage and WOrd Spotting, IST-1999-12203, 2001-2003, involving IDIAP and ETHZ (Prof. Luc Van Gool). CIMWOS investigates the automation of image and speech based annotation of video material. Besides fruitful contacts with other research groups, the project also created a formal link between the Swiss partners and Canal+, which gives useful insights into the needs of one of the top players in the European media landscape. More information at www.xanthi.ilsp.gr/cimwos.

BANCA

Biometric Access control for Networked and e-Commerce Applications, IST-1999-11159, 2000-2003, involving IDIAP and EPFL/ITS. The project aims at developing and implementing a secured system for user identification in the context of Internet applications, like teleworking, web-banking and others. The new security system will rely on combining classical authentication protocols (like PIN codes) with multimodal verification schemes based on speech and image. Three main research directions are pursued: (i) the development of scalable and robust multimodal verification algorithms; (ii) the development of scalable classifier combination techniques (fusion); and (iii) the design and implementation of an overall secure architecture including security protocols adapted to biometrics. More information at falbala.ibermatica.com/banca.

WEBKIT

Intuitive physical interfaces to the Web, IST-2001-34171, 2002-2004, involving UniGe. Webkit deals with the use of Tangible User Interfaces (TUI) for searching and retrieving information. TUIs give physical form to information, using small objects both as representations of information, and as controls to manipulate the underlying platform. The end users are children from primary and secondary schools. WebKit attempts to break away from the standard interface of the mouse and keyboard and should make retrieval and interaction an intuitive and fun process. More information at www.projectwebkit.com.

CERTIMARK

CERTIfication for waterMARKing techniques, IST-1999-10987, 2000 - 2002, involving EPFL/ITS and UniGe. The aim of CERTIMARK, based on the benchmark reference, is to make watermarking algorithms labeled with an international certification. The Certimark project consists of 15 academic and industrial partners. Certimark leads to the creation of a non-profitable organization CAST (www.cast-forum.de) that will provide the final certification of the watermarking technologies. More information at vision.unige.ch/certimark.

LAVA

Learning for Adaptable Visual Assistants, IST-2001-34405, 2002-2005, involving IDIAP. The objective of the project is to create fundamental enabling technologies for cognitive vision systems and to understand the system- and user-level aspects of their applications. Technologically, the objectives are the robust and efficient categorisation and interpretation of large numbers of objects, scenes and events, in real settings, and automatic online acquisition of knowledge of categories, for convenient construction of applications. We aim to exploit this technology in integrated systems that employ vision for information retrieval in a mobile setting, and systems that derive symbolic representations from video. More information at www.l-a-v-a.org.

HOARSE

Hearing Organization And Recognition of Speech in Europe, RTN2-2001-00140, 2002-2006, involving IDIAP. Hoarse is a Training and Mobility in Research (TMR) network focusing on multi-disciplinary training in speech science and speech technology, and aiming at exploiting human hearing properties towards improving speech recognition technology. More information at www.hoarsenet.org.

Events

(IM)2 Summer Institute 3–4.10.02

On October 3 and 4, Martigny will host the first (IM)2 internal workshop. Full details are available on the local pages of the (IM)2 web site.

NCCR Site Visit 17–18.10.02

On October 17 and 18, IDIAP will host the first Site Visit of the (IM)2 NCCR. The Review Panel will focus this year on the setting up of the NCCR, based on the first annual progress report as well as presentations and discussions which will take place during the site visit.
The review panel has been appointed by the SNSF. Members are Dr Phil Janson (SNSF, Chairman), Prof. Marco Baggiolini (SNSF), Dr Giordano Bruno Beretta (Hewlett Packard Laboratories, Palo Alto, USA), Prof. Shih-Fu Chang (Columbia University, New York, USA), Prof. Beat Hirsbrunner (SNSF), Prof. Mari Ostendorf (University of Washington, Seattle, USA), Prof. Steve Renals (University of Sheffield, UK), Prof. Gerhard Rigoll (Technische Universität München, D), Dr Andrew William Senior (IBM T.J. Watson Research Center, Hawthorne, USA).

The (IM)2 Newsletter is a publication of the NCCR on Interactive Multimodal Information Management, hosted by the Dalle Molle Institute for Perceptual Artificial Intelligence, Martigny, Switzerland. The National Centers of Competence in Research are managed by the Swiss National Science Foundation on behalf of the Federal Authorities.

SCHWEIZERISCHER NATIONALFONDS
FONDS NATIONAL SUISSE
SWISS NATIONAL SCIENCE FOUNDATION

The (IM)2 Newsletter backpage
Every month the (IM)2 Newsletter backpage presents one research group involved in the (IM)2 network: people, projects, research themes, infrastructure, etc.

Issue N° 07
October 2002
Editor: Jean-Albert Ferrez
www.im2.ch/nl nl@im2.ch

INTERACTIVE MULTIMODAL INFORMATION MANAGEMENT

The DIVA research group at Fribourg University

The Department of Informatics at University of Fribourg holds a research group called DIVA (for Document, Image and Voice Analysis) that is leaded by Prof. Rolf Ingold and Dr Dijana Petrovska-Delacretaz. The team currently has 10 members at various academic levels (professor, senior-assistant, post-doc, assistant, Ph.D. student).

The members of the DIVA research group.

The DIVA research group is dealing in general with multimedia engineering, with a focus on image and signal processing, speech recognition and document analysis. More than half of the research activity is financed by external grants, from public institutions as well as from industrial partners. The group deploys its activities in well established international networks. Prof. Ingold is a member of the editorial board of several international journals as well as a member of the board of directors of the French association GRCE (Groupe de Recherche sur la Communication Ecrite).

UNIVERSITAS FRIBURGENSIS

Past research themes and results

During the last decade the research activity of the group, previously called GIRAF (for Groupe de recherche sur l'Image et la Reconnaissance Automatique de Formes), was concentrated on image analysis in three main research area: 1) scene analysis, 2) content based image retrieval and 3) document analysis.

Scene Analysis: This research theme has dealt with computer vision in a non-calibrated environment. Invariant properties of projective geometry were used to perform matching of planar surfaces of a 3D scene observed from arbitrary viewpoints.

Content Based Image Retrieval: The goal of this project was to retrieve images from a heterogeneous database according to a user-drawn sketch. The technique relied on a sophisticated distance measure between the query (sketch) and the images contained in the database.

Document Image Analysis: The most significant research activity of the group has been about structure analysis of composite printed documents. The final goal was to extract the logical structure according to a given document type description. To achieve accurate results several low level processing were necessary. Thus, the group has also developed novel methods for layout analysis and for font recognition, a topic that has been neglected by the scientific community, and in which the University of Fribourg has become an uncontestable leader.

Current research projects

Since summer 2001, the group has extended its research activities towards multimedia, especially by including signal and voice analysis. The following projects are currently running.

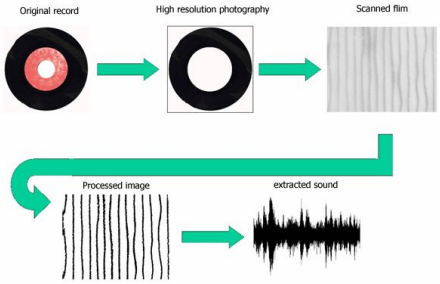
Interactive Document Reverse Engineering: This project consists of a general revision of the document analysis strategy. The idea is to perform document structure analysis in an assisted interactive environment, which is able to improve with use, by learning document models incrementally. An XML-based platform called XMIllum has been developed to achieve this goal and a framework called Edelweiss has been designed to make the whole system available on the Internet.

DocMining: The goal of this project, which is done in collaboration with industrial partners is to extend the XMIllum platform with the capability of handling graphics and technical drawings.

Faslav: The main concern of this project is to study and develop new methods for fully automated spoken language acquisition, understanding, and speaker verification by machines. We investigate by minimizing the currently inevitable usage of annotated corpora, by using methods that do not require phonetic nor orthographic transcription of the speech data.

Biomet: The goal of this project, done in collaboration with the Groupement des Ecoles de Telecom, France, is to collect a multimodal person authentication database, and to study the possible fusion of five different modalities in order to achieve better performance, and/or acceptability of the person authentication step.

Visual Audio: This project, which is done in collaboration with HES-SO Fribourg and the Swiss National Sound Archives aims at restoring sound from old records. The principle consists of analyzing the grove shape from high resolution photographs of these old disks.



The idea behind the Visual Audio project.

Smart Meeting Minutes: The goal of this project is to develop a fully operational demonstrator that integrates the various technologies developed by all (IM)2 partners. It deals with recording multiple sound and video tracks of a meeting in order to produce multimedia abstracts that can later be retrieved by content using multimodal interactions.

More information about the DIVA research group is available at diuf.unifr.ch/diva

ri

The (IM)2 Newsletter is a publication of the NCCR on Interactive Multimodal Information Management, hosted by the Dalle Molle Institute for Perceptual Artificial Intelligence, Martigny, Switzerland. The National Centers of Competence in Research are managed by the Swiss National Science Foundation on behalf of the Federal Authorities.

FNSNF
Schweizerischer Nationalfonds
Fonds national suisse
Swiss National Science Foundation