iM

INTERACTIVE
MULTIMODAL
INFORMATION
MANAGEMENT

IM2 Newsletter

Contents

News

2005-07-27
Norman Poh receives the AVBPA Best Student Poster Award

2005-08-03
The 34th issue of the IM2 Newsletter is online

Cover Story

VISIT TO A HIGH SCHOOL IN CRANS-MONTANA

In order to increase the awareness of children for science in general and the various aspects of multimodal interaction and multimedia information processing covered by IM2 in particular, a group of IDIAP seniors and students visited the Crans-Montana high school on June 9, 2005.

The two major objectives of this presentation were to introduce students to career paths they might otherwise not have considered, and to let girls in particular know that research in our fields warmly welcomes female players. IDIAP also projected a little movie called «Samira et la magie de la science» by the office of equality of chances at EPFL.

Children spent the afternoon participating in presentations in various fields. They had the opportunity to try and understand demonstrations on:

• Speech processing, through presentation of speaker localization by Olivier Masson
• Computer vision, through face tracking and hand movement tracking demonstrations by Agnes Just, Yann Rodrigez and Tiffany Sauquet
• Biometric authentication, presented by Mikaela Keller
• Brain computer interface for robot control, by Silvia Chiappa and Pierre Ferrez
• Information Retrieval, presented by Alessandro Vinciarelli
  Robota, a video about a robot developed at EPFL by Aude Billard and presented by Jose Millan.

A week after these presentations, children filled out a form to provide IDIAP with feedback on the event. They appreciated the afternoon and especially that researchers kindly answered to all their questions. Overall they understood the complexity of the research and some of them asked IDIAP to come back another year when they would be a little older in order to grasp all the details they had missed.

Many of those children did not consider studying sciences either because they were not aware of careers in these fields or because they had never seen the "fun" of research. Thanks to these presentations and demonstrations they are now considering getting more information about these careers. The EPFL video was a strong help to widen their interest.

The feedback on this event was extremely positive. IDIAP will again be involved in a public event with children and thanks to this first attempt it will be able to better address children's interest.

We would like to thank the Crans-Montana high school teachers and management for their support in setting up this event.

INTERACTIVE MUSEUM GUIDE

Recently, the computer vision lab of ETHZ (BIWI)[1] developed a prototype of a vision-based interactive museum guide. It has been shown to the public during the 150 years anniversary celebrations of the Federal Institute of Technology (ETH) in Zürich, Switzerland. Demonstrations took place in the Swiss National Museum Zürich and encompassed about 250 visitors in 20 guided tours of 10-15 persons each.

The guide runs on a tablet PC that features a touchscreen, a conventional webcam and a Bluetooth receiver, see Figure 1.
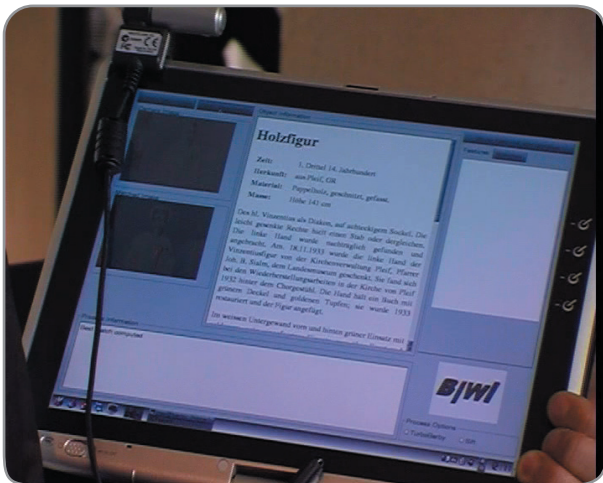

Figure 1: Tablet PC with the USB webcam fixed on the screen. The interface of the object recognition software is operated via a touchscreen.

This guide allows to recognize objects in museums based on images of the latter which are directly taken by the visitor. Furthermore, at any time, the computer can determine the visitor's location by receiving signals emitted by Bluetooth senders that are placed in the exhibition rooms of the museum. These so called BTnodes were developed by the Computer Engineering and Networks Laboratory (TIK) at ETHZ[2].

The location information obtained from the BTnodes is used to reduce the number of objects that the system can be expected to be in its field of view. Hence, the recognition accuracy is increased and the search time reduced. Moreover, this information can be used to indicate the user's current location in the museum. This device can be extended to display location dependent information, such as the closest emergency exits, the restrooms or the direction to the next coffee shop. Note that our museum guide does not impose a predefined visiting order.

During operation, our hand-held device allows the visitor to simply take a picture of an object of interest from any position and is provided, almost immediately, with a detailed description about it, see Figure 2. The visitor can also browse to some more specific information on the Intranet/Internet or to related objects that are currently exposed in the museum (e.g. made by the same artist). Furthermore, the visitor has the option to have the object description read out by the computer via a text-to-speech synthesis engine. This enhances the comfort by allowing the visitor to look at the exhibits instead of viewing the screen all the time.
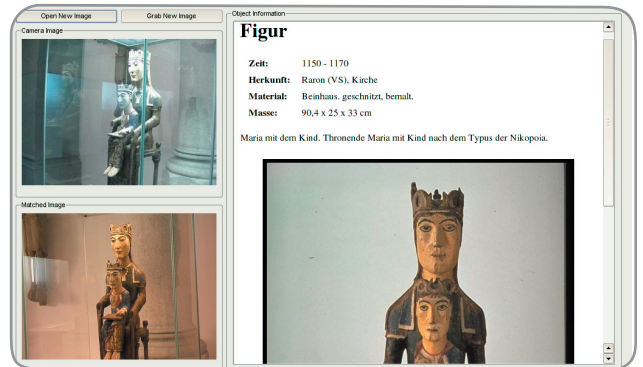

Figure 2: Interface of the object recognition application. On the upper left is shown the camera input image. On the lower left, the matched reference image is displayed. On the right hand side can be seen the browser window displaying the description of the object that has been associated to the matched reference image
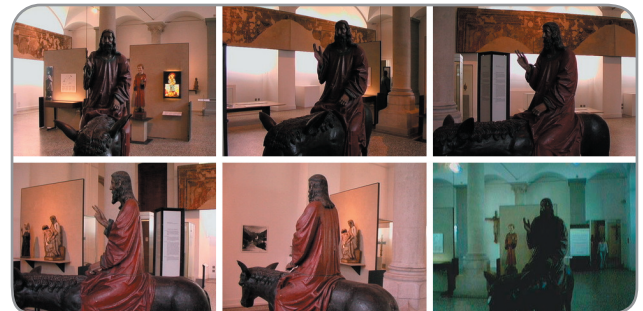

Figure 3: Sample of reference images and input image (lower right image) of an object in the museum. Note the important difference in appearance between the reference images and the input image.

Our interactive museum guide has been tested on 22 museum objects such as wooden statues, paintings, metal and stone items as well as coins and objects enclosed in glass cabinets, which produce interfering reflections. The input images were taken from substantially different viewpoints. The system uses a gray-scale image based algorithm that is invariant to changes in scale and in-plane rotation. Preparing the system to recognize additional objects is simple: a small set of reference images has to be added to the database, see Figure 3.

1 *http://www.vision.ee.ethz.ch*

2 *http://www.tik.ee.ethz.ch*

INTERACTIVE
MULTIMODAL
INFORMATION
MANAGEMENT

MARKET RESEARCH & BUSINESS ANALYSES
Commercial Opportunities for Meeting Archive Browser and Information Retrieval System

In order to have a better understanding of the commercial environment surrounding the research activities going on in IM2 and related projects, IDIAP has asked well known business consultant Christine Perey to conduct a preliminary market study specifically on two major components: the Meeting Browser and the Information Retrieval System. We quote here excerpts of her report. For further details regarding this study, please contact Dr Frank Crittin, Industrial Relations Manager, at *Frank.Crittin@idiap.ch*.

There are many commercial products designed to enhance the experience of and communications between people participating in business meetings. As part of the research process, this project compiled a list of over 100 companies which are either dedicated to improving productivity and information transfer during meetings or view meetings as one of the applications for their solutions. Products range from appliances incorporated into the furniture in conference rooms to lightweight and yet versatile software applications. Service providers are offering everything from fully-automated, self-serve, hosted "virtual rooms" to moderated or facilitated meeting management.

Once the meeting market was canvassed, another objective of this research project was to predict where and how two of the AMI and IM2 technologies would be most valued when applied in real world scenarios, and to identify the most likely partners to take the research products to their next stage of development. Virtually every company identified in this study and all of their customers and users could at some point in the future become interested in the technologies spawning from research underway in the AMI and IM2 projects. But, we conclude that, in general, the AMI and IM2 technologies studied for this report are still at least two to three years ahead of their target market need. Vendors and customers of meeting platforms report that the users of today's computer-assisted meeting tools rarely record their simple telephone meetings, even less frequently video or multi-modal sessions and even those who do are not encountering the challenges AMI and IM2 technologies are designed to address (obstacles or barriers to browsing and searching in one or more meeting archives, or barriers to identifying content introduced during business meetings).

Nevertheless, many experts agree that there will come a time when as a general rule business meetings will be technology-assisted and recordings will become commonplace. It is in the best interest of the IM2 and AMI projects and their partners to continue monitoring those who will detect changes in meeting patterns and technology use. At the same time, AMI and IM2 project participants will engage in numerous programs and activities to accelerate the pace at which technology-assisted meeting participation (or viewing) and the use of analytical tools will be discovered and their value understood.

Until the mainstream business technology purchaser detects a need for AMI and IM2-developed meeting related technologies and enhanced meeting participant or meeting browser services, this study revealed

several areas where multimedia content repositories are reaching a critical size. Vendors of multimedia content publishing and management platforms report that their users could apply AMI and IM2 project technologies to solve emerging needs in the near future. We pursued this line of exploration. The anecdotal evidence gathered for this study indicates that the most popular application for multimedia content, the objective for which multimedia assets are accumulating and the situations in which improved searching and navigation could be beneficial, are closely tied to learning.

As the barriers to the production of compelling and interactive computer-delivered content fall, multimedia programs for employee or student education are accumulating on network servers in enterprises and institutes of learning. Some exceptional customers have thousands of "learning units", each including video/audio, presentation materials and text files. The size of some repositories has reached the point where the problems of searching, navigating and browsing these repositories and archives is felt by those creating content and managing content on behalf of user communities.

Another focal point for accumulation of simpler yet valuable archives of digital audio and synchronized presentations is in the business of hosting and delivering network-based earnings conferences on behalf of publicly held companies. We believe that as the leaders of companies become increasingly comfortable with the recording processes and results, they will consider the many other applications for the same technology, leading to the possible recording of executive briefings and committee meetings.

In general, until the complete integration of capture technologies becomes more popular, we believe that the opportunities identified through this study can be explored by way of discussions with multimedia content publishing platform developers, content and courseware publishers and providers of knowledge and asset management systems.

©2005 Perey Consulting

Timeline: Content Accumulation

IM2 Newsletter
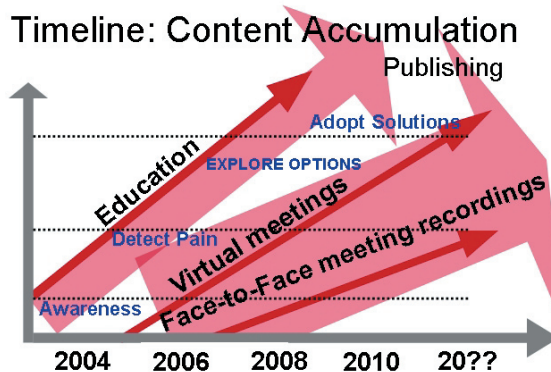
Upcoming Events

IM2 2005 Summer Institute

This year's Summer Institute will take place in Lausanne (EPFL) during the week of November 14-18, together with the SNSF Review Panel Site Visit and the Scientific and Industrial Advisory Board Meeting. The detailed program will be announced later.

Partner News

Norman Poh receives the AVBPA Best Student Poster Award

Norman Poh, a PhD student at the IDIAP Research Institute working in the framework of the «Access Control and Protection (ACP)» project of the National Center of Competence in Research on Interactive Multimodal Information Management (NCCR IM2) lead by IDIAP, has been awarded the Best Student Poster Award in recognition of his work on «Biometric Fusion» during the 5th International Conference on Audio Visual Biometric Person Authentication which took place on July 20-22 in Rye Brook, NY. *(http://biometrics.cse.msu.edu/avbpa2005.html)*

The conference is considered the most important in the field of biometric authentication and has a rejection rate of 40%. In the future, it will be called International Conference on Biometrics (ICB).

The paper supporting Norman Poh's poster can be found as:
N. Poh and S. Bengio, A Novel Approach to Combining Client-Dependent and Confidence Information in Multimodal Biometric, in the 5th Int'l. Conf. Audio- and Video-Based Biometric Person Authentication (AVBPA), LNCS 3546, pages 1120-1129, 2005.

IM2 Talks and Posters at MLMI'05

Following the very successful first edition in Martigny, the second Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms (MLMI) took place in Edinburgh, UK, on July 11-13, 2005.

We list here the papers and posters presented by IM2 members at MLMI'05 in Edinbrugh. See http://groups.inf.ed.ac.uk/mlmi05/ for the full online proceedings and recordings of the workshop.

Oral Presentations

Can Chimeric Persons Be Used in Multimodal Biometric Authentication Experiments?
Norman Poh and Samy Bengio (IDIAP)

Detection and Resolution of References to Meeting Documents
Popescu-Belis A., Lalanne D. (Uni GE + Uni FR)

Hierarchical Multi-Stream Posterior Based Speech Recognition System
H. Ketabdar, H. Bourlard and S. Bengio (IDIAP)

The Development of the AMI System for the Transcription of Speech in Meetings
Thomas Hain, Lukas Burget, John Dines, Iain McCowan, Martin Karafiat, Mike Lincoln, Darren Moore, Giulia Garau, Vincent Wan, Roeland Ordelman, and Steve Renals (IDIAP + AMI Partners)

Poster Presentations

Using natural language processing to improve German LVCSR accuracy
R. Beutler, T. Kaufmann, H. Romsdorfer, and B. Pfister (ETHZ)

Answering natural language queries on spoken dialogs in meeting discussions
Ghorbel, H. Ailomaa, M, Rajman M. (EPFL)

Semantic Segmentation of Video Collections using Boosted Random Fields
B. Janvier, E. Bruno, S. Marchand-Maillet, and T. Pun (Uni GE)

Characterisng optimal structures in multimedia collections for enhanced exploration
S. Marchand-Maillet and E. Bruno (Uni GE)

Spectral plane investigation for probabilistic features for ASR
F. Grezl (IDIAP)

Continuous Microphone Array Speech Recognition on Wall Street Journal Corpus
M. Hari Krishna (IDIAP)

Browsing multimedia archives through implicit and explicit cross-modal links
Rigamonti M., Lalanne D., Evequoz F., Ingold R. (Uni FR)

Two-Handed Gesture Recognition
A. Just and S. Marcel (IDIAP)

Hierarchical approach for spotting keywords
M. Lehtonen (IDIAP)

Rapid Multimodal Dialogue Design: Application in a Multimodal Meeting Retrieval and Browsing System
Melichar, M., Lisowska, A., Armstrong, S., Rajman, M. Rapid (EPFL + Uni GE)

Towards Automatic Video Conference Proceedings
J.-M. Odobez (IDIAP)

Document interactive navigatgion in multimodal databases
Y. Rytsar, S. Voloshynovskiy, O.Koval, F. Deguillaume, S.Startchik and T. Pun (Uni GE)

Toward Joint Segmentation and Classification of Dialog Acts in Multiparty Meetings
M. Zimmermann (ICSI)

Quasi Text-Independent Speaker Verification with Neural Networks
M. Gerber and B. Pfister (ETHZ)