iM

INTERACTIVE
MULTIMODAL
INFORMATION
MANAGEMENT

IM2 Newsletter

Content

News

MLMI'07

The official website of the 4th Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithm (MLMI'07) is now available:
http://www.mlmi07.org

IM2 RSS feed

IM2 website offers a RSS feed including the content of IM2 news and press. If you want to subscribe, use this link http://www.im2.ch/newsandpress/RSS.

How IM2 shapes the future
A NEW PUBLICATION TO MARK THE END OF THE FIRST PHASE OF IM2

There will be a time...

In the courtyard of the Collège de l'Abbaye de Saint-Maurice, one day in 2006... "I had a strange dream last night," says Vincent, 15 years of age, to his friends Stéphanie and Samuel. "I was living in a world more advanced than ours. A world where machines were better adapted to us. We could communicate with them by speaking and gesturing… And even simply by telepathy!"
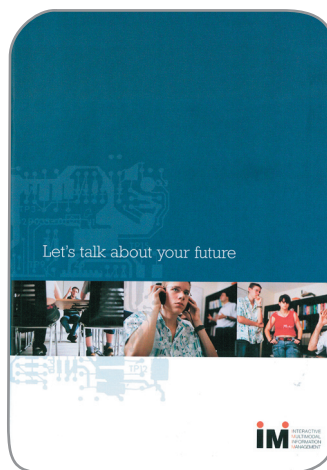
"You really watch too many movies," comments Stéphanie. Vincent is not so sure, "Perhaps it's not science fiction," he says. "I heard about some researchers who are working right near here, in Martigny…" And Samuel suggests, "Let's go meet them".

It was thus that this past summer Stéphanie, Vincent and Samuel paid a visit to some scientists of the IM2. On the premises of the IDIAP Research Institute in Martigny, they questioned researchers from all over the world who have a common goal: creating new human-machine interactions.

If you would like to share in this adventure, we invite you to read our new publication "Let's talk about your future". By following the three adolescents through their meetings with the IM2 scientists, you will discover research that will transform your daily life in the near future.

This publication marks the end of the first phase of IM2 (2002-2005). It provides an interim report on this program and outlines what will be explored in the next few years.

Please feel free to contact us, if you want to receive a copy. We hope that you will enjoy reading the brochure and feel confident that you will be pleased to learn about our activities.

Cover Story

Let's talk about your future

IDIAP

FNSNF
SCHWEIZERISCHER NATIONALFONDS
FONDS NATIONAL SUISSE
SWISS NATIONAL SCIENCE FOUNDATION

The Mobile Phone as an Interaction Device for the Meeting Room

Recently, the Computer Vision Lab of ETHZ (BIWI) [1] has developed a new algorithm that allows to detect and describe local features in images. We coined it SURF [2], which stands for Speeded Up Robust Features. As the name suggests, the algorithm is much faster than existing methods and yet is robust to lighting variations and image transformations such as changes in scale, rotation and translation.

In order to further empower the users of the IM2 Smart Meeting Room, we have implemented a scenario involving a standard mobile phone equipped with a internal digital camera and a remote object server. Hereby, the user can take a photo of a scene which contains an object known to the remote object server. The SURF algorithm running on the mobile phone extracts points of interest in the input image, transforms them into a robust representation and finally sends the resulting image description via MMS (Multi Media Service) to the object server, where the description is compared to the descriptions of all objects of interest contained in a database. Once the best matching object is found, a given text associated to this object is sent back to the mobile phone via SMS (Short Message Service) and made visible to the user.

A mockup of this technology operating in the meeting room is shown in Figure 1. A participant can take a photo of a slide that is presented. Objects contained in the image can help to determine the topic of the presentation, and a user could retrieve previous presentations with similar content. An alternative application is a bookmarking service: An attendee takes pictures of slides (or parts of them) which are of interest to

Figure 1: A mockup illustration, showing the envisioned scenario.

him. The recognition process described above is executed, and the collected slides are for instance transferred to the user's email inbox. In this way, after the meeting, the user has a collection of only those slides that were of particular interest to him.

The detection of different objects in some meeting room slides is illustrated in Figure 2. As can be seen, the approach is able to find the same objects in different slides. Even a part of a given object is often sufficient to correctly retrieve the object and therefore presentation of interest.
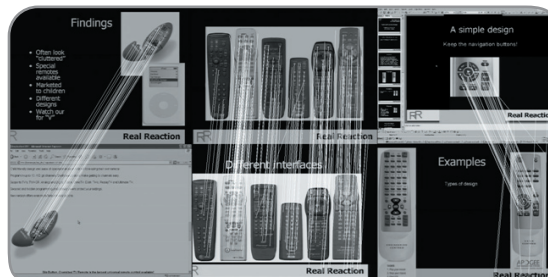

Figure 2: Shown on the left-hand side is and example object that is detected in two different slides, in the middle multiple objects of the same category found in two different slides, and on the right-hand side the matching of an object part to the full object displayed in another slide.

Currently, we have a prototype system in place. Our ongoing activities include improving the matching to make it more robust against false matches in text areas. In addition, work is progressing to scale the process up to large slide collections.

[1] BIWI Homepage: *http://www.vision.ee.ethz.ch*
[2] SURF – Speeded Up Robust Features:
http://www.vision.ee.ethz.ch/~surf

The IDIAP 15th Anniversary Workshop

To celebrate 15 years since its inception, IDIAP held a workshop on 12-14th September 2006 at the Hotel du Parc, Martigny. Videos of the talks and slides are available from *http://www.idiap.ch/ws15/.*

Over these years, the central themes of interest to IDIAP – machine learning, Information Retrieval, Speech Processing and Computer Vision -- have matured, leading to many theoretical and practical advances. The workshop addressed these areas from the viewpoint of key progress made and current techniques and applications. Invited speakers put into context the achievements and challenges in their fields, coupled with young researchers who had the opportunity to discuss their work. Programme highlights included:

- **Chris Bishop** (Assistant Director, Microsoft Research UK) presented key developments in machine learning, concentrating on the introduction of probabilistic methods and graphical models.
- **Nelson Morgan** (ICSI Berkeley) gave a fascinating historical overview of the field of speech recognition pointing out key advances in the field, particularly in the use of statistical models.
- **Thomas Hofmann** (Google) took the perspective that, by curious coincidence, the birth of the internet largely coincided with the birth of IDIAP. The internet has heralded a revolution

in information and communication and will continue to provide a platform for intense research on methods for information retrieval.
- **Patrick Perez** (IRISA/INRIA Rennes) discussed computer vision theory and applications, with entertaining demonstrations and applications to tracking and image analysis. Much progress in the field was made by the application of stochastic methods to approximate inference in large scale systems.

IDIAP director Herve Bourlard gave an entertaining history of the institute, from its small beginnings to its current day involvement in several large scale international projects. IDIAP is well positioned to take advantage of the continued growth in man-machine interaction.

The major social event of the workshop was a visit to the St Bernard Dog Museum, and a chance to relax during the IDIAP open afternoon of demonstrations and discussions with staff members.

Overall, the workshop speakers gave the impression of tremendous growth in research areas of interest to IDIAP over the last 15 years, both from theories of machine learning and related methods to their application to solving real world problems. From a historical perspective, it is likely that this is very much still the beginning, with a great many interesting challenges still lying ahead.

INTERACTIVE
MULTIMODAL
INFORMATION
MANAGEMENT

Advancement of women

Portrait of Dr Barbara Caputo, first recipient of the IDIAP fellowship for female researcher

I received a laurea degree in Physics from the university of Rome, Italy, «La Sapienza'» in 1998 and a M.S. degree in Biomedical Imaging from the same institute in 2000. In 2004 I received a PhD in computer science from the Royal Institute of Technology, Stockholm, Sweden.

In October 2005 I was awarded the 2005 IDIAP Female fellowship; I joined IDIAP in August 2006, and since then I've been working on visual recognition methods for cognitive systems. Prior to joining IDIAP, I was a research scientist at the University of Erlangen-Nurnberg, computer science department, in the pattern recognition group headed by Prof. H. Niemann (1999-2001). During 2002, I was visiting researcher at the Smith-Kettlewell Eye Research Institute (SKERI) in San Francisco, California, USA. From 2003 till June 2006 I was a research associate at the Computational Vision and Active Perception (CVAP) laboratory at the Royal Institute of Technology (KTH), Stockholm, Sweden.

I am the author/co-author of more than 30 international publications; My M.Sc. thesis won the annual best thesis award for the year 2000 from the University of Rome La Sapienza.

I currently am the coordinator of two national research projects sponsored by the Swedish national research council. Since 2003 I have been actively participating in several national and European projects on cognitive vision and cognitive systems.

Milène Di Franco, 13, reports on her scientific adventure

Within the framework of the advancement of women in science and in collaboration with the AVPEHP (association valaisanne des parents d'élèves à haut potentiel), IM2 supported the participation of a 13 year old girl, Milène Di Franco, in a scientific holiday camp, called "how it works". This camp which took place from August 13 to 20 at La Chapelle d'Abondance in France represented an extraordinary opportunity for young children between 11 to 14 year old to discover the world of sciences. On her return, Milène was very enthusiastic. Here is her report:

My name is Milène Di Franco, I am 13 years old. Thanks to IM2 and AVPEHP I took part in a great one week scientific adventure, with 50 kids of which 9 were girls.

During this week, I had the choice to participate in different workshops and activities; robotics, crystallography, the construction of a rocket, to know better the hardware of a computer, to create a game or an audio CD. For instance, I chose to learn the different modes of communication between animals or to rebuild a Rubik's cube (my best score: 1min 46). Also I studied the setup of a website. This activity was very interesting, but difficult at the same time. Two days were dedicated to visiting the FUTUROSCOPE. It was great and I discovered a lot of scientific concepts and experiments. For instance, I loved all the demonstrations in 3D.

Apart from scientific topics, I swam and learnt the theory of first aid. I have to say that some days later I put into practice some of these notions. My new knowledge was very useful in an urgent situation.

In conclusion, I recommend this camp. I met people, discovered and enjoyed a lot in a fun atmosphere. Once again, thank you very much for having given me the opportunity to take part in such an interesting week.

MLMI 07 - www.mlmi07.org
BRNO, JUNE 28-30, 2007

Preliminary Announcement

The fourth MLMI workshop is coming to Brno in the Czech Republic, following successful workshops in Martigny (2004), Edinburgh (2005) and Washington, DC (2006). MLMI brings together researchers from the different communities working on the common theme of advanced machine learning algorithms applied to multimodal human-human and human-computer interaction. The motivation for creating this joint multi-disciplinary workshop arose from the actual needs of several large collaborative projects, in particular IM2, which contributed to initiate the first MLMI in June 2004.

MLMI-07 will follow on directly from the annual conference of the Association for Computational Linguistics (ACL/EACL 2007, Prague, June 25-27). MLMI-07 will feature special sessions and satellite events such as the Summer school of the European Masters in Speech and Language (http://www.cstr.ed.ac.uk/emasters/).

The call for papers, posters and demos will be issued in November 2006. The expected deadlines are February 23, 2007 for papers and a month later for posters and demos. In common with the previous MLMI workshops, revised versions of selected papers will be published in Springer's LNCS series (cf. LNCS 3361 and 3869).

Brno is the second largest city in the Czech Republic and the capital of Moravia. The local organizers are members of the Faculty of Information Technology (http://www.fit.vutbr.cz) at Brno University of Technology, founded in 1899. Brno can be easily reached by direct flights from Prague, London and Munich and by trains or buses from Prague (200 km) or Vienna (130 km).

Organizing Committee *(to be extended)*
• Honza Cernocky, Brno (organization co-chair)
• Andrei Popescu-Belis, Geneva (program chair)
• Steve Renals, Edinburgh (special sessions)
• Pavel Zemcik, Brno (organization co-chair)

IM2 Newsletter

Upcoming Events

5th Site Visit of the Review Panel   *16-17.11.06*

The 5th Site Visit of the IM2 Review Panel will take place on November 16-17, 2006, in Martigny. This year, the focus will be on:
- Assessment of the start of the new phase and achievements since the last review
- Implementation of the new management structure
- Education and training activities since the beginning of the NCCR
- Presentation of projects

This year, we welcome three new members who have agreed to join our panel: Prof. Tat-Seng Chua - National University of Singapore, Prof. Helen Mei-Ling Meng - The Chinese University of Hong Kong and Prof. Yu-Jin Zhang - Tsinghua University.

The entire list of the review panel members is: Prof. A. Steger, (Chair), Dr. G.B. Beretta, Dr. B. Byrne, Prof. S.-F. Chang, Prof. T.-S. Chua, Prof. B. Hirsbrunner, Prof. R. Jain, Prof. H.M. Meng, Prof. C. Opitz and Prof. Y.-J. Zhang

New seminar series at ETHZ

At ETHZ, the Computer Vision Laboratory and the Speech Group of the TIK lab, have started a joint seminar series on MultiMedia and MultiModal signal processing for the upcoming winter semester. These `m&m seminars' are intended to strengthen the knowledge about and appetite for multimodal processing in the two groups, both of which participate in IM2. The NCCR has been the direct inspiration for this initiative. The seminars will be announced at *http://www.vision.ee.ethz.ch/teaching/ mm_seminar.en.html* and are wide open to all who are interested to come!

Challenging Brain Computer Interfaces: Neural Engineering Meets Clinical Needs in Neurorehabilitation
ROME ITALY, NOVEMBER, 9-10, 2006

The past few years have witnessed advances in Brain-Computer Interfaces (BCI), but how far is this new field from clinical practice? The goal of the workshop is to draw the current and future scenarios involving the application of advanced neural engineering techniques to interpret brain signals for clinical use in the rehabilitation context.

The presentations will consist of a series of invited talks and poster presentations.

Some of the major groups in BCI pursuing clinical applications of this technology will report their experience. Also, the view of clinicians involved in neurorehabilitation programs will complete the picture. Finally, the European MAIA project will report their achievements in non-invasive brain-controlled robots. Altogether the workshop will address how ultimate neural engineering techniques could meet the challenge of neurorehabilitation.

http://www.maia-project.org/workshop-2006.php

NIPS 2006

- NIPS'06 Workshop on Learning to Compare Examples
  http://www.idiap.ch/lce
  (Organised by Samy Bengio and David Grangier)

- NIPS'06 Workshop on Advances in Models for Acoustic Processing
  http://www.idiap.ch/amac
  (Organised by David Barber and Taylan Cemgil)

Official website of the NIPS conference:
http://nips.cc

Partner News

INTEL gift

Intel gift to the Computer Vision and Multimedia Laboratory, University of Geneva: Intel Corporation, Santa Clara, CA, USA, has granted the amount of $50'000 to the CVML to support research in the area of «Models for personal semantic augmentation of digital media». This gift will contribute to support the work of Dr. Serhiy Kosinov, Viper Group, CVML.

Major publications

The AMI meeting transcription system: Progress and performance
T. Hain, L. Burget, J. Dines, G. Garau, M. Karafiat, M. Lincoln, J. Vepa and V. Wan
proceedings of NIST RT'O6 Workshop, Washington, D.C., 2006

A Novel Gaussian Sum Smoother for Approximate Inference in Switching Linear Dynamical Systems
D. Barber and B. Mesot
to appear in Advances in Neural Information Processing Systems (NIPS), volume 20, 2006

Distributed SVM applied to Image Classification
E. Kokiopoulou and P. Frossard
IEEE Int. Conf. on Multimedia & Expo (ICME), Toronto, Canada, July 2006

Monocular Tracking with a Mixture of View-Dependent Learned Models
T. Jaeggli, E. Koller-Meier and L. Van Gool
IV Conference on Articulated Motion and Deformable Objects (AMDO 2006), July 2006

Tracking the Multi-Person Wandering Visual Focus of Attention
K. Smith, S. Ba, D. Gatica-Perez and J.-M. Odobez
in Proc. Int. Conf. on Multimodal Interfaces (ICMI), Banff, Nov. 2006

Elcano: a tangible personal multimedia browser
H. Chiquet, F. Evéquoz and D. Lalanne
19th ACM Symposium on User Interface Software and Technology, demo paper UIST 2006, Montreux, Switzerland, October 15-18, 2006

Automatic extraction of geometric lip features with application to multi-modal speaker identification
I. Arsic, R. Vilagut and J. Thiran
Proc. of ICME 2006, Toronto, Canada, July 2006

Low-Level Grounding in Multimodal Service Robot Conversational System using Graphical Models
P. Prodanov, A. Drygajlo, J. Richiardi and A. Alexander
Intelligent Service Robotics, Springer, 2006.

Asymmetric Learning and Dissimilarity Spaces for Content-based Retrieval
E Bruno, N. Moenne-Loccoz and S. Marchand-Maillet
International Conference on Image and Video Retrieval 2006, 13-15 July, Tempe, AR, USA

An Analysis of Quantitative Aspects in the Evaluation of Thematic Segmentation Algorithms
M. Georgescul, A. Clark and S. Armstrong
The 7th SIGdial Workshop on Discourse and Dialogue, 144-152, Sydney, July 2006

Archivus: A multimodal system for multimedia meeting browsing and retrieval
M. Ailomaa, A. Lisowska, M. Melichar, S. Armstrong and M. Rajman
Proceedings of the Interactive Presentations of COLING/ACL 2006, Sydney, Australia, July 17-21, 2006