# IM2 Newsletter

**INTERACTIVE**
**MULTIMODAL**
**INFORMATION**
**MANAGEMENT**

## Contents

## News

**August, 1**
IDIAP, the Leading House of IM2, moves!

The new address is:
IDIAP Research Institute
Centre du Parc
Av. des Prés-Beudin 20
P.O. Box 592
1920 Martigny
Phone numbers unchanged.

**September, 3-4-5**
Swiss-Japanese workshop on Multimedia Content Structuring and Access

**November 15-16**
Site Visit of the IM2 Review Panel

**www.im2.ch**

IM2, c/o IDIAP Research Institute, Centre du Parc,
Av. des Prés-Beudin 20, P.O. Box 592, 1920 Martigny
info@im2.ch - www.im2.ch

## Cover Story

### 4th Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms
28-30 JUNE 2007, BRNO, CZECH REPUBLIC

The fourth workshop on Machine Learning for Multimodal Interfaces, MLMI'07 [1], has taken place in Brno, Czech Republic on June 28-30. The workshop, sponsored by the AMI Consortium and the IM2 NCCR, returned this year to Europe after a 2006 edition in Washington, DC. Launched in 2004 in Martigny, the MLMI series [2] gathers researchers working on machine learning techniques for processing and browsing multimodal input data, in human-human and human-computer interaction settings.

The first two days of MLMI'07 had different thematic orientations. The papers presented on the first day generally had spoken input or audio signals as their major focus, while those presented on the second day focussed on visual input (morning session), or content abstraction and browsing (afternoon session). The two poster sessions also followed a similar distribution. The first day included a special session presenting results from the second PASCAL Speech Separation Challenge [3], and the third day was dedicated to the AMIDA Training Day, followed the next week by the Summer school of the European Masters in Speech and Language [4].

The talks of the two invited speakers were related to the dominant theme of each day. Dr. Nick Campbell [5] from ATR's Media Information Science Laboratories in Kyoto, Japan, described striking analyses of non-verbal features that allow multimodal devices to make sense of human interactive behaviour. Prof. Vaclav Hlavac [6], head of the Center for Machine Perception at the Czech Technical University in Prague, demonstrated several applications of methods for statistical pattern recognition to visual processing tasks.

Overall, MLMI'07 featured 18 oral presentations – two invited talks, two PASCAL talks, and 14 regular talks – and 42 poster presentations, with a total of about 90 registered participants from a dozen countries. Besides the sponsoring projects, AMI/AMIDA and IM2, participants were also related to other large research projects on multimodal processing and browsing, such as CALO and CHIL, and some came from private companies. Local universities were well represented, as well as other European, US and Japanese universities and research institutions. The setting of the workshop, in the conference rooms and lobby of the Hotel Continental, fostered lively interaction among the participants, during the lunch and coffee breaks and at the social dinner, set up by the local organizers in the beautifully restored Rectorate building at the Brno University of Technology.

The oral presentations were recorded using the IM2 Presentation Acquisition System resulting of a collaboration between IDIAP and UniFr, so that video, audio and slides are now available on the MMM server [7]. Revised versions of selected papers will be published in Springer's Lecture Notes on Computer Science series towards the end of 2007 (previous MLMIs appeared as LNCS 3361, 3869, 4299). And of course, plans are already underway for a successful fifth MLMI next summer!

Andrei Popescu-Belis, PC Chair,
and the Organizing Committee:
Hervé Bourlard, Jan Černocký,
Pavel Zemčík and Steve Renals

[1] http://www.mlmi07.org
[2] http://www.mlmi07.org/index.php?id=past_mlmi
[3] http://homepages.inf.ed.ac.uk/mlincol1/SSC2/
[4] http://www.cstr.ed.ac.uk/emasters/
[5] http://feast.atr.jp/nick/
[6] http://cmp.felk.cvut.cz/~hlavac
[7] http://www.idiap.ch/mmm/talk-webcast/mlmi-07/

## IM2 Newsletter

Issue N° 42 / August 2007

Editor : François Foglia
E-mail : info@im2.ch

INTERACTIVE
MULTIMODAL
INFORMATION
MANAGEMENT

## TV-Report on Handwriting Recognition at the University of Bern
### APRIL 2007

At the beginning of 2007, there has been a television report on the handwriting recognition and writer identification activities at the University of Bern. In this report Prof. Horst Bunke gave interviews concerning handwriting recognition for the IM2 smart meeting room project and the Swiss National Library. Together with the two PhD students Marcus Liwicki and Andreas Schlapbach animations illustrating the process of handwriting recognition with hidden Markov models have been created. During an afternoon at the beginning of the year a tv team came to the university for the recording session.
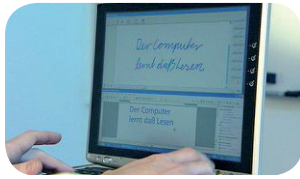
The report begins with a scene where a person writes on the whiteboard using the eBeam system. The data are directly transferred to the computer and appear on the screen, where the person can access the transcription. During the interviews Prof. Bunke states the problems of handwriting re-cognition and writer identification, and the importance of the task for the smart meeting room and other applications.

Some methods for normalization, feature extractions, and classification are briefly described in the report. Persons who are not involved in the field of pattern recognition may easily follow the explanations and understand the main ideas of the algorithms. However, some technical aspects are left out to make the report not too complicated.
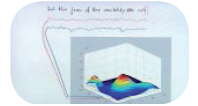
In spite of some minor imprecisions resulting from simplification, the report is clear and technically sound. It is a successful promotion of the handwriting recogniton activities within IM2.

The report has been broadcasted on March 28 in the Nano-series on 3Sat. More information about the television report can be found on http://www.3sat.de/nano/cs-tuecke/105981/index.html.
Marcus Liwicki, University of Bern

## BananaScreen
### just your smile, all you need to log in…

**www.bananasecurity.com**

Tired of locking your computer before leaving? Tired to enter your password so many times? BananaScreen has the solution.

This fun and easy application locks your computer when you move away from it. When you come back, BananaScreen unlocks the computer in less time than it would take to enter your password. All you need is a correctly configured webcam and BananaScreen. The software which is still in beta version works with most of the webcams on Windows XP systems. BananaScreen is freely available for download on the www.bananasecurity.com website. A MacOS version will be available soon.

The software has been downloaded more than 25'000 times. Several articles have been published in newspapers and in hundreds of blogs on the web. BananaScreen also got the "100% Clean" Award from Softpedia.

Due to the numerous positive feedbacks and the encouraging support, new products based on BananaScreen's technology are beeing designed.

By the end of the year, a company will probably be created. Some interesting contacts have already shown a great interest in this new venture.

The success of BananaScreen is a good example that biometrics, usually associated with high security like physical access control, may also lead to more fun and convenient applications. BananaScreen has been developed in the framework of a technology transfer project between the IDIAP Research Institute and the HES SO (Haute Ecole Superieur de Suisse Occidentale) in Sierre, through the "Business Experience" program. The goal of this program is to give HES students the opportunity to develop entrepreneurship skills by bringing an innovative idea or technology to the market with the support of professional coaches. On IDIAP's side, the technology provider, the project was coordinated by Dr Yann Rodriguez, a former IM2 PhD student who recently finished his thesis on Face Detection and Recognition under the supervision of Dr Sebastien Marcel and Dr Samy Bengio.

## New IM2 Financial Manager
### JUNE 2007

Sandra Micheloud replaces Pierre Dal Pont who retired in June. She has completed, in 2002, a BS in International Hospitality and Tourism management. She worked for 13 years for Swisscom where she held several positions. After her Bachelor of Science she entered Syngenta Crop Protection in Monthey as financial controller. During that time she also took evening courses to broaden her theoretical knowledge in Swiss and international financial norms, taxes, law, economy and financial controlling. She joined IDIAP in February 2007 to take over t he position of Financial Manager.

The IM2 management wishes all the best to Pierre Dal Pont and welcomes Sandra Micheloud in the IM2 venture.

## IDIAP Participates in the CLEAR 2007 Evaluation Initiative
### BALTIMORE, MD, USA, 8-9 MAY

The CLEAR evaluation workshop (www.clear-evaluation.org) is an international initiative to evaluate systems that automatically recognize events, activities, and their relationships from video and audio. Since 2006, CLEAR brings together projects and researchers working on related technologies to participate in a common international evaluation. Many computer vision related tasks such as face detection and tracking, person detection and tracking, and head pose estimation are the targets of the evaluation. In 2007, Jean Marc Odobez and Sileye Ba from IDIAP participated in the CLEAR workshop (Baltimore, MD, USA, 8-9 May) evaluating algorithms for two head pose estimation tasks and coordinating one of them.

In the task coordinated by IDIAP, the evaluation data consisted of eight single-view meetings recorded in the IDIAP multi-sensor meeting room with people's head pose ground truth generated with a magnetic field sensor recorded in the IM2 context. To automatically solve this task, the
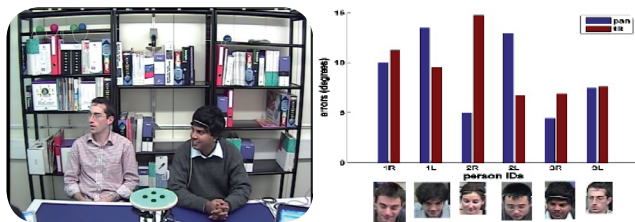


*Fig 1. Sample image from the IDIAP head pose evaluation data (right) and IDIAP's method pose estimation errors for individuals for the 1st task (left)*

IDIAP team applied the probabilistic head pose tracker developed during Silèye Ba's thesis (funded by IM2 and directed by Jean-Marc Odobez and Hervé Bourlard) . In addition to IDIAP, the team of the University of Karlsruhe, Germany, participated in the task. IDIAP's results, which were the best among the two competing teams, are quite competitive. The average estimation error is about 8.8° in head pan estimation and 9.4° in head tilt estimation. Analyzing results more in detail, we can observe that there is important performance variability over individuals (due

to variations in people appearances and sitting attitudes) as can be seen in Fig. 1. Depending on individuals, the head pan and tilt errors vary from 4 degrees to 15 degrees.

The IDIAP team also participated in a second, multi-view head pose estimation task in CLEAR, which consisted in estimating, from four cameras, the head poses with respect to a global room coordinate of people recorded in a multi-sensor lecture room at the University of Karlsruhe (see Fig 2).

Five teams participated in this task: University of Karlsruhe, Technical University of Catalonia, Spain, Centro per la Ricerca Scientifica e Tecnologica, Italy, University of Illinois at Urbana Champaign, USA, and IDIAP. For this task the IDIAP team developed a technique that first tracks location and head poses independently for each of the camera views, then represents head poses with respect to the global room coordinates, and finally fuses the estimates provided by each of the four cameras to obtain a single estimate. The IDIAP system produced overall results of 15° error in head pan estimation, which is quite good given the small resolution of



*Fig 2. UKA CHIL head pose evaluation data*

the images, where face features such as nose or eyes are hardly visible. The work of Odobez and Ba on modeling of visual attention is partly funded by IM2.VP and contributes to other IM2 IPs through collaboration with IM2.MPR.

## Venture Leader prize 2007

Have you ever been ill in a foreign country and found it frustrating that the doctor couldn't understand you? MedSLT addresses the problem of a lack of qualified human interpreters in hospitals, especially in emergency rooms. MEdSLT offers a cost-effective alternative in the form of a tool for helping physicians perform a verbal diagnosis on a foreign patient. The adopted approach is unique in that it reuses general language resources thanks to the specialization method used. This allows the system to offer tailor-made solutions in terms of the languages and diagnostic domains needed.

It is thanks to this SNF project headed by Prof. Pierrette Bouillon and developed at the Multilingual Language Processing Department of the Translation and Interpretation School (Geneva University) that Marianne Starlander, currently a PhD candidate, won the Venture Leader pri-



*Marianne Starlander*

ze 2007 (http://www.venturelab.ch/fr/vLeaders. asp). She thus travelled to Boston end of June together with her 20 fellow prize-winners including the co-founder of Koaba (Herbert Bay) also an IM2 related project.

During these ten days, the venture leaders' team had the opportunity to be in the heart of the dynamic Boston area, meeting successful serial entrepreneurs and visiting start-up incubators. Networking was an important part of the stay, but the main goal was a four-day tailor made course in business at the Babson College, MA.

"Winning the prize has been a very enriching and encouraging experience" says prize winner Marianne Starlander, "it really opened many doors to me and made me much more self-confident. I really think this opportunity should be given to more researchers".

UNIVERSITÉ DE GENÈVE

# Upcoming Events

## IDIAP moves!          *2007-08-02*

Facing constant growth for the 7th year in a row, the IDIAP Research Institute had already added the Dalle Molle Pavillon (2002) and a floor of the UBS building (2004) to the original Villa Tissières. As future expansion was not possible anymore downtown Martigny, the institute relocates in the Centre du Parc. The Centre, well-known to the IM2 crowd for having hosted Winter institutes, MLMI'04 and several site visits, offers modern offices over three floors, and expansion opportunities both in the current building and in the immediate surroundings. Spin-off companies Cinetis and IdeArk come along, benefiting from an area dedicated to start-ups. The move is scheduled for August 2007, but should not have any major impact on the activities of IDIAP and its services.
The new address is:

IDIAP Research Institute
Centre du Parc
Av. des Prés-Beudin 20
P.O. Box 592
1920 Martigny
Phone numbers unchanged

## Swiss-Japanese Workshop

*on Multimedia Content Structuring and Access*

http://www.im2.ch/jp-ch-workshop-2007

As a key research theme in Information and Communication Technologies in general, the workshop organized by IDIAP, in collaboration with EPFL and UniGe, will serve as a forum for the dissemination of state-of-the-art research, development, and implementations of multimedia systems, technologies and applications, also including overviews of current large project initiatives involving some of the workshop participants. Presentations and panel discussions will encompass the following research themes and project overviews:
Joint audio/video (multimedia) processing, Multimedia structuring and indexing, also including multimedia (hierarchical, multi-tier) annotation standards, Retrieving and accessing multimedia information, Next-generation multi-modal search in multimedia data collections, Multi-lingual issues in all the above.
The workshop will take place in Switzerland (Martigny, Lausanne, Geneva) for 3 days (September 3-5, 2007).

# Partner News

## Most Cited Paper Award

IM2 is proud that one of its papers in multimodal speech and face identification recently got the Most Cited Paper Award from Elsevier. Papers for this distinction have been determined solely on the basis of the highest number of cites received for all journal articles published between the years 2004 and 2006 [data culled from SCOPUS reports (www.scopus.com) created on January 11, 2007].

*C.Sanderson*

The winning paper is "Identity verification using speech and face information," by C. Sanderson and K.K. Paliwal, Digital Signal Process. 14 (2004), pp. 449–480.
IM2 congratulates Drs. Sanderson and Paliwal for accomplishing this great achievement.

# Selected publications

Generating Usable Formats for Metadata and Annotations in a Large Meeting Corpus
*A.Popescu-Belis and P.Estrella*
Proceedings of ACL 2007 (45th Int. Conf. of the Association for Computational Linguistics) Interactive Poster and Demonstration Sessions, pp. 93-96, Prague, Czech Republic

Person Authentication using Brainwaves (EEG) and Maximum A Posteriori Model Adaptation
*S.Marcel and J. del R.Millán*
IEEE Trans. on Pattern Analysis and Machine Intelligence, 29(4): 743–748.

Face Detection with Boosted Gaussian Features
*J.Meynet, V.Popovici and J.Thiran*
Pattern Recognition, Vol. 40, Nr. 8, pp. 2283-2291, 2007

Mixtures of Boosted Classifiers for Frontal Face Detection
*J. Meynet, V. Popovici and J. Thiran*
Signal, Image and Video Processing, Vol. 1, Nr. 1, pp. 29-38, 2007

Reliability-based decision fusion in multimodal biometric verification systems
*K.Kryszczuk, J.Richiardi, P.Prodanov,A.Drygajlo*
EURASIP Journal on Advances in Signal Processing, Volume 2007 (2007), Article ID 86572, 9 pages

Multimodal Biometrics for Identity Documents (MBIOID)
*D.Dessimoz, J.Richiardi,C.Champod,A.Drygajlo*
Forensic Science International, Volume 167 (2007), pp. 154-159

Human-Centered Computing: Toward a Human Revolution
*A.Jaimes, D.Gatica-Perez, N.Sebe, T.S.Huang*
IEEE Computer, May 2007

Quasi text-independent speaker verification based on pattern matching
*M.Gerber, R.Beutler, and B.Pfister*
In Proceedings of Interspeech. ISCA, 2007

A Cognitive and Unsupervised MAP Adaptation approach to the Recognition of the Focus of Attention from Head Pose
*J-M.Odobez and S.Ba*
Proc. of the IEEE International Conference on Multimedia and Expo (ICME'07), Bejing, July, 2007

Combining On-Line and Off-Line Systems for Handwriting Recognition
*M.Liwicki and H.Bunke*
9th Int. Conf. on Document Analysis and Recognition, 2007

Analysis of Head Mounted Wireless Camera Videos for Early Diagnosis of Autism
*B.Noris, K.Benmachiche, J.Meynet, J-P.Thiran and A.Billard*
In Proceedings of the International Conference on Recognition Systems

Dimensionality Reduction using Adaptive Approximation
*E.Kokiopoulou and P.Frossard*
IEEE Int. Conf. on Multimedia & Expo (ICME), Beijing, China, July 2007

Spoken Handwriting Verification using Statistical Models
*A.Humm, J.Hennebert and R.Ingold*
Accepted for publication in Proc. of 9th International Conference on Document Analysis and Recognition (ICDAR'07), Curitiba, (BR), September 23 - 26, 2007

Analysis of multimodal binary detection systems based on dependent/independent modalities
*O.Koval, S.Voloshynovskiy and T.Pun*
In Proceedings of the IEEE 2007 International Workshop on Multimedia Signal Processing, Chania, Crete, Greece, October 1-3 2007

Automatic image annotation with relevance feedback and latent semantic analysis
*D.Morrison, S.Marchand-Maillet, E.Bruno*
Workshop on Adaptive Multimedia Retrieval (AMR 2007). Paris, 2007

Contrasting the Automatic Identification of Two Discourse Markers in Multiparty Dialogues
*Popescu-Belis A. and Zufferey S.*
To appear in the Proceedings of SIGDIAL 2007 (8th SIGdial Workshop on Discourse and Dialogue), Antwerp, Belgium, 10 p.