

**The 1st (IM)2 Newsletter**  
What is (IM)2 and who is involved.  
(IM)2's upcoming events.

**The 9 IP of (IM)2**  
A short introduction to the 9 Individual Projects of (IM)2.

Issue N° 00  
February 2002  
Editor: Jean-Albert Ferrez  
[www.im2.ch/nl](http://www.im2.ch/nl) [nl@im2.ch](mailto:nl@im2.ch)



# 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

## The (IM)2 NCCR

(IM)2, the National Center of Competence in Research (NCCR) on Interactive Multimodal Information Management, is aimed at the advancement of research, and the development of prototypes, in the field of man-machine interaction. The NCCR is particularly concerned with technologies coordinating natural input modes (such as speech, image, pen, touch, hand gestures, head and/or body movements, and even physiological sensors) with multimedia system outputs, such as speech, sounds, images, 3D graphics and animation.

The field of multimodal interaction covers a wide range of activities and applications, including the recognition and interpretation of spoken, written and gestured languages, computer vision, and the automatic indexing and management of multimedia documents. Other important related themes are information content protection, data access control, and the structuring, retrieval and presentation of multimedia information.

Multimodal interfaces represent a new, highly strategic, direction for information technologies of the future. Thanks to such interfaces, man-machine interactions will become simpler and, by consequence, more productive. In the near future, multimedia systems equipped with such interfaces will be flexible enough to accommodate a wide variety of users, tasks and environments for which current interaction modalities (such as keyboard, mouse and screen) are insufficient. In a first instance, ideal interfaces would be capable of manipulating more complex and realistic data, including the combination of different forms of data, such as audio and video. The (IM)2 NCCR, headed by the Dalle Molle Institute of Perceptual Artificial Intelligence (IDIAP) in Martigny, combines many partners from a number of university institutions (EPFL, University of Geneva, University of Fribourg, University of Bern, ETHZ), as well as a HES (Sion) and a range of commercial companies. The NCCR also has numerous international contacts, including an agreement for the exchange of young researchers with ICSI in Berkeley, California.

## (IM)2 Constituent bodies

### Director

Prof. Hervé Bourlard, Director of IDIAP.

### Deputy Director

Prof. Murat Kunt, Director of ITS, EPFL.

### Program Manager

Dr Jean-Albert Ferrez, IDIAP.

### Management Board

Aside from Prof. Bourlard, Prof Kunt and Dr Ferrez, the management board of (IM)2 is composed of one representative per institution involved in (IM)2, currently Prof. Thierry Pun (Uni GE), Prof. Rolf Ingold (Uni FR), Prof. Horst Bunke (Uni BE) and Prof. Luc Van Gool (ETHZ).

### IP Heads

See the backpage for a complete list of the Individual Projects (IP) and their heads.

### 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**, Linköping 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; 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.

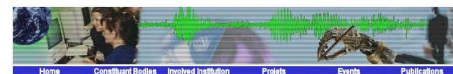
### NCCR Office at SNSF

Dr Stefan Bachmann.

## The (IM)2 web site

[www.im2.ch](http://www.im2.ch)

In addition to this Newsletter, the most important public presence of (IM)2 is its web site [www.im2.ch](http://www.im2.ch). Thanks to our graphic artist and webmaster [Thierry Collado](#), it appears now as a full featured and appealing starting point for collecting information about the NCCR. It will function as a web portal, directing visitors to the relevant parts of the web sites of the institutions involved in (IM)2.



## Events

### (IM)2 Kick-off meeting 6.3.02

On March 6 at EPFL, a joint Kick-off meeting will celebrate the start of the (IM)2 NCCR and of the EPFL Signal Processing Institute (*Institut de Traitement des Signaux, ITS*), a key player in (IM)2. This scientific, informal event aims at bringing together the people who will work within ITS and (IM)2. The program features presentations of ITS, of (IM)2 and its Individual Projects, of the torch software library (see [www.torch.ch](http://www.torch.ch)), of ICSI, Berkeley, and the young researcher exchange agreement being signed between ICSI and (IM)2.

### (IM)2 Inauguration 4.5.02

The official Inauguration of (IM)2 will take place on May 4 in Martigny. This event is synchronized with the (IM)2 Scientific and Industrial Advisory Boards meetings that take place on May 3. The program will be announced later.

### (IM)2 Summer Institute 3-8.10.02

From October 3 to October 8, Verbier will host the first (IM)2 internal workshop. This will be the perfect opportunity to look back on the first months of the NCCR, to identify gaps and overlaps in the various research plans.

# The Nine Individual Projects (IP) of (IM)2

## (IM)2.ACP: Multimedia Information Access and Content Protection

IP Head: Prof. Horst Bunke, Uni Bern  
Partners: UniBE, ETHZ/TIK, IDIAP, Basler Papiermühle

The objective of (IM)2.ACP is to investigate effective methods for the identification or verification of people based on voice and facial characteristics in order to facilitate or protect access to multimedia information. Within (IM)2, these activities will take place in the framework of common efforts toward the research and development of a truly multi-modal (using voice and face characteristics) user identification and authentication system.

## (IM)2.AP: Application project (Smart Meeting Minutes)

IP Head: Prof. Rolf Ingold, Uni Fribourg  
Partners: UniFR, HES-SO, CSEM

The Application Project aims at developing a fully operational demonstrator that integrates the various technologies developed by all (IM)2 partners.

The application that has been chosen is called Smart Meeting Minutes. It deals with recording multiple sound and video tracks of a meeting, in order to produce abstracts that can later be retrieved by content using multimodal interactions.

## (IM)2.DS: Multimodal Information: Deployment, Storage, and Interactive Access

IP Head: Prof. Roger Hersch, EPFL  
Partners: EPFL/LSP, EPFL/LBD, UniL/INFORGE

The aim of (IM)2.DS is to design and implement a database and related information services to support the demonstration of (IM)2 results from other IPs, using an agreed common application framework (the Smart Meeting Minutes).

The common database will provide the shared repository for (IM)2 data, namely multimodality streams (video, voice, presentation data, etc..) and associated metadata (annotations, timestamps, etc.). The database will rely on a performing and scalable storage solution for multimedia data. The database will meet the challenging (IM)2 requirements by providing a set of advanced features, such as easy evolvability, multifacet data support, and visual Web interfaces for multimodal querying of the database.

## (IM)2.IIR: Multimodal Information Indexing and Retrieval

IP Head: Prof. Thierry Pun, Uni Geneva  
Partners: UniGE/CUI, UniGE/TIM, EPFL/LTS, IDIAP

The goals of (IM)2.IIR are to provide multimedia data indexing mechanisms that will facilitate subsequent multimedia data retrieval. With respect to existing state-of-the-art, the emphasis of (IM)2.IIR will be put on:

- multimedia data: how to take into account, and if possible benefit, from the fact that data is not limited to a single medium,
- user interaction: how to benefit from the user actions, ultimately provided in a multimodal manner,
- interoperability: how to interconnect the various IIR tools already developed by the partners of this IP, between medias (ie. between image, text, sound, etc.) as well as between levels (ie. audio signal, recognized speech, high-level semantic interpretation).

## (IM)2.MDM: Multimodal Dialogue Management

IP Head: Prof. Susan Armstrong, Uni Geneva  
Partners: UniGE/TIM, EPFL/LITH, EPFL/LIA, UniGE/CUI

The goal of (IM)2.MDM is to define a framework for multimodal dialogue understanding by a computer program. The IP then aims at integrating this computational dialogue model with lower level language processing tools, in order to provide a generic set of tools for dialogue management by a computer agent. The main goal of (IM)2.MDM is to provide:

- a unified dialogue model for human-to-human and human-to-machine interaction,
- an efficient formalism for the rapid prototyping of multimodal dialogue models, well adapted for a distributed, real time set up,
- a computational architecture allowing an easy integration of autonomous software modules implementing the available models.

## (IM)2.SP: Speech Processing

IP Head: Prof. Hervé Bourlard, IDIAP  
Partners: IDIAP, ETHZ/TIK, UniGE/TIM

The goal of (IM)2.SP is to provide the (IM)2 NCCR with advanced and flexible speech processing modules which can be used as:

- an input mode (voice input),
- an audio indexing tool (requiring large vocabulary, continuous speech recognition systems) turning audio files into text,
- an output mode (requiring speech coding and text-to-speech systems).

## (IM)2.MI: Multimodal Input and Modality Integration

IP Head: Prof. Hervé Bourlard, IDIAP  
Partners: IDIAP

The overall objective of (IM)2.MI is the development of a system for the fusion and efficient decoding of different input modalities. The objective can be decomposed as follows:

- the development of new data fusion algorithms, as well as new decision strategies,
- the development of a multichannel statistical model for the combination of asynchronous input streams,
- the development of an efficient multimodal decoder,
- the implementation of all these algorithms into a common platform.

## (IM)2.SA: Scene Analysis

IP Head: Prof. Murat Kunt, EPFL  
Partners: EPFL/LTS, UniGE, ETHZ, UniBE

The goal of (IM)2.SA is to be able to analyze the content of an image or image sequence in terms of its highest possible semantic content from several modalities. Our solutions will bring the following contributions and progress to the state-of-the-art:

- intelligent image segmentation combining stochastic and structural models,
- efficient, principled and unified grouping method to deal with geometrically regular patterns,
- robust face detection and recognition method under adverse conditions,
- unconstrained automatic handwritten text reading.

## (IM)2.VE: Virtual Entities

IP Head: Prof. Daniel Thalmann, EPFL  
Partners: EPFL/LIG, UniGE/Miralab

The goal of (IM)2.VE is to create a 3D Virtual version of existing scenes that will allow us to take advantage of this ability. This will require capturing the shape and motion of the face and body of people present in the scene, so that they can be recreated as virtual 3D characters who act exactly as the actual people did. We will focus on:

- video-based acquisition, resynthesis and animation of human faces,
- video-based acquisition, resynthesis and animation of human bodies,
- development of 3D worlds including an intelligent virtual camera tool to navigate through Virtual Environments,
- integration.