idiap
RESEARCH INSTITUTE

ANNUAL   REPORT   2009

CONTENTS

NOTE FROM THE PRESIDENT

"THE FOUNDATION COUNCIL KEEPS A CLOSE WATCH ON IDIAP'S GROWTH CURVE"

Olivier Dumas, President of the Foundation Council of Idiap

"In 2009 Idiap recorded a slight drop in its financial return – as it also did in its costs – but this can by no means be interpreted as a sign of decline. Quite the opposite. In fact, although in a sluggish economic context, the research institute has lost a few projects, at the same time it has acquired a stronger financial basis due to the support of the Municipality of Martigny, the State of Valais and the Swiss Confederation. The Idiap Research Institute can now look to the future with confidence and peace of mind. The creation in 2009 of a showroom, plans for building a technology park and its strategic alliance with EPFL are all indicators of the institute's willingness to open up to the outside, to develop and to strengthen its integration into the Swiss university network.

These actions are encouraging, as are the prospects that they represent. However, the Foundation Council is keeping a close watch on Idiap's growth curve.

Since its creation in 1991, Idiap has seen the number of its employees increase exponentially, even forcing the institution to move in 2007 to the west wing of Centre du Parc on the outskirts of the town. Although this development is to be welcomed because it underlines the progress of research projects and reflects the dynamism of an institute that has succeeded in taking pride of place on the international stage, it still concerns the Foundation Council, which wants Idiap to remain a "reasonable" size.

The elite researchers who currently work in the institute's laboratories willingly admit it: in addition to its high scientific quality, the attraction of Idiap is both its geographical location – a microclimate in the heart of the Swiss Alps – but also and especially its almost family-sized structure. This human-sized organisation allows a good level of communication between the teams to be maintained, encourages synergies and stimulates creativity. The Council wants to safeguard these assets. Combined with the academic advantages that the new strategic alliance with EPFL brings for the senior researchers, they should ensure that Idiap maintains the high level of quality of its research teams in the long term.

I would like to take this opportunity to praise the excellent approach that has led to this strengthening of the partnership with EPFL.

Lastly, Idiap will celebrate its 20th anniversary in 2011. In addition to the many festivities with a scientific theme that are planned, the institute will establish new links with the public, which does not always understand the issues involved in its work.

In the meantime, I would like to thank all Idiap's employees for their work and wish them a creative future!

MESSAGE FROM THE DIRECTOR

"WE ARE BECOMING MORE ATTRACTIVE TO TOP LEVEL SCIENTISTS"

Hervé Bourlard, Director of Idiap

2009 was a transitional period for us, both on a structural and a scientific level.

Last year, we announced our partnership with EPFL and this year, we have really consolidated this joint development plan. Although the cooperation with our colleagues in Lausanne started on the actual day our institute was created in 1991, today we are passing a new milestone. We have succeeded in bringing together two operations, two cultures, two completely different institutions, David and Goliath: an independent research institute with barely 100 employees and an école polytechnique fédérale, which involves 10,000 people and has 250 laboratories.

For Idiap, this represents two major changes: access to academic titles and expansion of the institute's management.

Although they have been supervising PhD students from EPFL for a number of years, until now our senior researchers gained no recognition for doing this. They now have access to academic titles and benefit from better international visibility. Therefore, Idiap is becoming more attractive to top level scientists.

The organization chart has also changed with the arrival of tenure track assistant professors (PATT) and senior lecturers (MER). In the long-term, they will assist me in the management of Idiap to achieve a wider form of management. Here again, this was a challenge. Sharing the helm was not an easy step to take! However, the excellent level of the people selected, the richness brought by this diversity and the fact that this expansion is in the interest of the future of the Institute brought me around to the idea.

This joint development plan is also accompanied by a renewed increase in support from the Swiss Confederation, the State of Valais and the Municipality of Martigny. Thanks to this, the Institute can now contemplate the future with greater serenity: the public proportion of our financing has risen from 26% in 2008 to 34% in 2009. (Cf. page 30)

The Swiss National Science Foundation (SNSF) has also shown its confidence in Idiap once again by attributing to him for a third and final period of four years the "IM2 - Interactive and multimodal management of information systems" centre. During its assessment, SNSF mentioned the scientific quality of Idiap, which was praised as an excellent structural partner.

Through its strategic alliance with EPFL, Idiap strengthened its structures and financing in 2009. This solidity has allowed it to expand its field of vision and to creatively explore the new research areas set out in 2008.

After having expanded, Idiap is now flourishing. I am extremely pleased with this development and I would like to thank all the employees for their commitment and enthusiasm.

Portrait

The Idiap Research Institute, based in Martigny (Valais/ Switzerland), is a non-profit foundation specialised in the management of multimedia information and man-machine multimodal interactions (autres domaines phares?). The Idiap was founded in 1991 by the Town of Martigny, the State of Valais, the Ecole polytechnique fédérale de Lausanne (EPFL), the University of Geneva and Swisscom, and is autonomous but connected to EPFL by a joint development plan.

The Idiap budget, which amounts to more than 9 million Swiss francs, is 65% financed by research projects awarded following competitive processes, and 35% by public funds. (cf. Distribution of sources of financing, page 30)

Whereas it only employed around thirty people in 2001, in 2009 Idiap had around one hundred employees, including 80 researchers (professors, senior researchers, researchers, postdoctoral students and PhD students). All the personnel work at Centre du Parc in Martigny, in the west wing. The institute moved there in August 2007. It now occupies 2,500 m^2 of premises on four floors.

Research areas

Idiap's main research areas are the following:

- Perceptual and cognitive systems
  (speech processing / natural language understanding and translation / document and text processing / vision and scene analysis / multimodal processing / cognitive sciences)

- Social and human behaviour
  (web social media / mobile social media / social interaction sensing / social signal processing / verbal and non-verbal communication analysis)

- Information interfaces and presentation
  (multimedia information systems / user interfaces / system evaluation)

- Biometric user authentication
  (speaker identification and verification / face detection, identification and verification / multimodal biometric user authentication)

- Machine learning
  (statistical and neural network based ML / computational efficiency, targeting real-time applications / very large datasets)

Objectives

Through its activities, Idiap pursues three main objectives:

- Conducting fundamental research projects at the highest level in its preferred areas, thus taking its place among the best on a national, European and global scale.
Idiap benefits from a wide network of partners internationally and works actively with large universities, public and private research centres, etc.

- **Developing recruitment** by helping its interns discover the world of research, by welcoming talented young researchers preparing their PhD and by providing a number of courses at EPFL and in-house.

- **Ensuring technology transfer** through the widest dissemination possible of its research results in the scientific community, but also by forging close ties with the world of industry.

Geographical situation

The Idiap Research Institute is in Martigny, one of the main towns of the canton of Valais, in the French-speaking part of Switzerland in the south of the country. In the heart of the Alps, Valais has an exceptional landscape and a pleasant microclimate, which makes it a very popular tourist destination and a very sought-after place to live.

Martigny is a town of approximately 15,000 inhabitants and is situated close to Montreux, Lausanne and Lake Geneva. Geneva airport is 90 minutes away by train. Martigny is well positioned in the centre of Europe.

Idiap in figures

Human resources (average over the past few years)

13 permanent researchers
 8 postdoctoral
27 PhD students
 9 development engineers
 6 system engineers
10 interns and visitors per year
10 administrative employees
 2 doctorates awarded
30 positions in start-ups on the IdeArk site
23 nationalities represented

Scientific activities

- IM2 National Centre of Competence in Research (interactive and multimodal management of information systems) since 2001

- Participation in 37 research programmes

- Project management in 7 consortiums

- Participation in the economic development strategy of the Canton of Valais through The Ark programme and in particular the IdeArk company

- 166 scientific publications

- Organisation of a number of international conferences

www.idiap.ch

HEURISTICS, OR COMMUNITY RESEARCH

Scientists, students and voluntary contributors from around the world are working together on an artificial intelligence project managed by Idiap.

MASH European project Massive Sets of Heuristics for Machine Learning	
Idiap	**Project coordination** Person in charge: François Fleuret, senior researcher Idiap team: 1 senior researcher, 2 PhD students, 1 dev. engineer
Partners	• Weierstrass Institute for Applied Analysis and Stochastics, WIAS, Berlin (DE) • Centre national de la recherche scientifique (National Centre for Scientific Research), CNRS, Lille (FR) • National Institute for Research in Computer Science and Control, INRIA, Orsay (FR) • Czech Technical University in Prague, CVUT (CZ)
Financing	European seventh Framework Program (FP7)
Schedule	January 2010 - December 2012
Website	http://mash-project.eu

"If we want machines to imitate human behaviour we have to equip them with a system as complex as the human brain."

François Fleuret,
Idiap senior researcher,
coordinator of the project

One of Idiap's new areas of research involves new artificial intelligence techniques using vast families of heuristics. These specialised modules are developed in a collaborative way, and combined using statistical methods. Two projects are following this avenue: the VELASH (Very large sets of heuristics for scene interpretation) national project, which began in September 2009, and the MASH (Massive sets of heuristics) European project.

New generation robots for tomorrow

While today's robots are generally confined to engineering environments, those of tomorrow could well be integrated into other domains, such as home care and medical operations, rescue or even entertainment. Thus, in the majority of projects connected with artificial intelligence, researchers analyse the way in which the human brain processes the data it receives and then attempt to teach the machine to reproduce the way it works. What happens inside our brain when we open a door, switch off the light or dial a telephone number?

Steering an avatar and an automated arm

The MASH project is based on a simple observation: the data processed by human beings are complex and they are complex in the same way for the computer. On the other hand, the structure and reasoning of a human brain are much more complex than those of a computer and combine a large number of modules dedicated to very specific tasks. "All it takes to be convinced", explains François Fleuret, "is the knowledge that some people can have particular brain injuries, which, for example, make it impossible for them to recognise faces, even though they suffer from no other visual deficiency. In summary, if we want machines to be as intelligent as human beings, we have to equip them with an equally rich representation of the world."

While the aim of the VELASH national project is to detect objects within images, the MASH European project has two objectives: to steer an avatar in a three dimensional universe, and steer an automated arm. In December 2012, at the end of three years of work, the machine should be capable of understanding its environment and guiding these two objects alone.
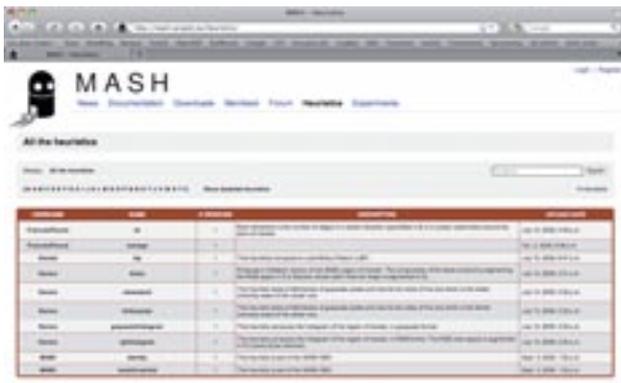
MASH should enable new tools to be developed, which would allow hundreds of people worldwide to work together.

Participation of the global community

What better way to achieve this than by bringing together hundreds of thought processes from hundreds of different brains? "The idea", explains François Fleuret, "is to operate on the same model as the large free software projects, and to combine the efforts of a large number of volunteer contributors, with each bringing his own ideas to the problem." During the first few months, the project is reserved for the five teams of researchers that are partners of MASH. Each team works on the programme, contributes its expertise, and the results are gathered together. Then the project will be opened up to universities and hautes écoles, and lastly to the general public.

"It is a development oriented Web 2.0, with the creation of small communities. Each participant can add modules that he has programmed himself, and our statistical techniques combine the whole to make it an intelligent global system. Experiments take place all the time and feedback is provided continuously", says François Fleuret.

Idiap's work on MASH and VELASH should further scientific progress in Europe, but also give our continent better visibility in this area. The project should also enable new working tools to be developed, which would allow hundreds of people worldwide to work together.
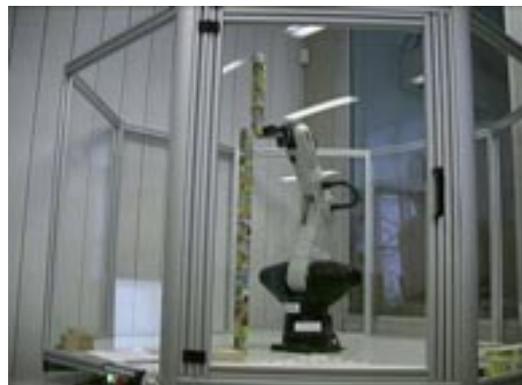
The MASH European project has two objectives: to guide an avatar in a three dimensional universe, and steer an automated arm. In December 2012, thanks to the participation of the global community, the machine should be capable of understanding its environment and guiding these two objects, the avatar and the arm, without human intervention.

ACLD, VIRTUAL ASSISTANT

Designing a system capable of understanding a conversation and finding useful documents in relation to the content of this conversation is the challenge of Andrei Popescu-Belis, senior researcher at Idiap.

AMIDA & IM2 projects Objective ACLD Automatic Content Linking Device	
Idiap	**Project coordination** Person responsible: Andrei Popescu-Belis, senior researcher Idiap team: 1 senior researcher, 1 PhD student, 1 dev. engineer
Partners	• University of Edinburgh (UK) • German Research Center for Artificial Intelligence, DFKI, Saabrücken (DE) • Netherlands Organisation for Applied Scientific Research, TNO, Delft (NL)
Financing	Swiss National Science Foundation European sixth Framework Program (FP6)
Schedule	January 2008 - December 2011
Websites	www.amiproject.org, www.im2.ch

«I can already see myself discussing holiday plans with my family while photos of the destinations discussed, rental prices, flight times, etc. automatically scroll on a screen.»

Andrei Popescu-Belis, Idiap senior researcher, responsible of the project

"The ACLD is a sort of virtual assistant", explains Andrei Popescu-Belis, who is charge of the project. During a work meeting, when the participants are in discussion, they do not have immediate access to their archives, unless they interrupt their discussion and begin a search that is often tedious. Using advanced speech recognition and indexing of multimedia documents technology, the system analyses the discussion and offers content related to the conversation in real time."

Searching multimedia documents

The ACLD system can carry out searches in the project's library (minutes of meetings, reports, scientific articles, etc.), in multimedia documents stored in available databases (slides, audio recordings, videos, etc.) and on the web. The information can be displayed on a shared screen or on the screens of the participants' laptops, thus allowing them to make different choices. "If our system could only provide relevant documents twice during a meeting, that would still be a huge gain", explains Andrei Popescu-Belis. "All the more so since it only takes a quick look at the screen from time to time."

The ACLD system reuses components from the work carried out by the IM2 National Centre of Competence in Research, the third phase of which will last until 2013, and the AMIDA European project, which ended in December 2009. It aims to make human interaction more effective in real time.

Teaching assistance

The context can be that of an intelligent meeting room, but not only that. Idiap is also working on teaching assistance, with teachers being able to enrich their lessons with additional information. The sole requirement is that the system has audio input data, for example a recorded lesson, since it works on the basis of voice recognition.

"The system can also have a personal use", explains Andrei Popescu-Belis. "Imagine that I am planning my holidays with my family, I would benefit from seeing scrolled on a screen, as the conversation progresses, photos of the destinations discussed, webcams, addresses of property agencies and airline companies, the weather forecast, recommendations for vaccinations, etc."

Initial tests at the Rolex Learning Center of EPFL
The IM2 National Centre of Competence in Research, managed by Idiap and financed by the National Science Foundation, ensures the future of the project. "We are currently working on the semantic analysis of what is said, because it often occurs that certain words have several meanings. Through machine learning methods, we are teaching the system to locate clues that can help it to understand the exact meaning of each word by comparing it to hypotheses already generated for neighbouring words."

During this third phase of the project, the task of the Idiap researchers is also to connect their work to real users and applications. "With this aim in mind, we have begun to work in cooperation with the CRAFT laboratory of EPFL, which is working on interactive technologies for teaching assistance. We plan to test this application in a meeting room for students in the new Rolex Learning Center, which has libraries, reading rooms and rooms for group work." A projector fixed to the ceiling will project the image of additional information found by the ACLD onto the table, and the students will be able to indicate their choices directly on the table. Their choices will be detected using a camera also situated on the ceiling.

A The system highlights and indicates in red the words that it recognises in the conversation.

B A tag cloud shows the relative importance of the key words in the conversation allowing the topics covered to be taken in at a glance.

C A search is carried out in the project's document database, which includes extracts of past meetings.

D All the words recognised are also entered into an internet search engine. The titles of the best pages found are displayed.

E Moving the mouse over a result shows why it has been found and clicking on the title opens it with associated software.

SS2-ROB, THE BEGINNINGS OF THE DOMESTIC ROBOT

Teaching it to differentiate between an office and a kitchen, a cup and a pen… These are the first steps in designing a machine, which, in the future, should relieve humans from having to perform domestic chores.

Pascal 2 project Objective SS2-Rob, Semi Supervised learning of Semantic Spatial concepts of a mobile Robot	
Idiap	Person in charge: Barbara Caputo, senior researcher Idiap team: 1 senior researcher and 1 postdoctoral researcher
Partners	Università degli Studi di Milano (IT)
Financing	European seventh Framework Program (FP7)
Schedule	January 2010 - June 2011
Website	www.pascal-network.org

"We are trying to create a robot that understands our environment and our language, so that one day we will simply be able to give it tasks to do vocally."

Barbara Caputo, Idiap senior researcher, responsible of the project

"With SS2-Rob", explains Barbara Caputo, senior researcher at Idiap, "we are working on intelligent robots in order that they understand their environment and can interact with us. If for example I tell my domestic robot to put the Obama T-shirt in the washing machine, it should be able to distinguish it from the others and perform the task without a problem."

Being able to talk to it as to a human being
"To put it plainly, the robot must be able to recognise where it is and know how to get to where the user tells it to go. The difficulty is that we are working on a semantic basis and not a mathematical one. Because no one wants to have to say to their robot: roll along 3 m 46 to the left at an angle of 42 degrees, then lift your left arm onto the grey table in front of you, and take the round object, of a diameter of 8 centimetres… We simply want it to go to fetch the forgotten cup of tea on the desk!" explains Barbara Caputo. That being the case, we have to teach the machine to differentiate a kitchen from an office, a cup from a bowl, etc.

How do we know that an office is… an office?
The human being analyses and understands the situation in a single glance. To also have this overall vision, the machine must have points of reference and understand the relationship between the different elements.
"It must be able to visualize the concept of these rooms and their context, and define which pieces of furniture are found in them and how they are laid out, etc. We have to think of everything. Because if we tell a computer that an office is a room in which there is a table with a chair, a computer and a bookshelf, without any other clarification, when it finds itself in a warehouse full of tables, chairs and computers… it will look for the cup of tea!" In order for the machine to memorise all these variants, the researchers use machine learning.

EMMA, AUTOMATIC MEDICAL IMAGE ANNOTATION

Since they have to deal with digitisation and an increasing number of medical images, professionals want a system capable of archiving and sorting them. The work of Idiap in this area is the most advanced in Europe.

| **EMMA project** Enhanced Medical Multimedia Data Access	
Idiap	Person in charge: Barbara Caputo, senior researcher Idiap team: 3 senior researchers and 2 PhD students
Financing	Hasler Foundation
Schedule	January 2008 - December 2011

The rapid development of new techniques for acquiring medical images and the widespread use of computerized equipment to save, transfer and store digital medical images have highlighted the need to design new methods for managing and archiving these data.

Sorting according to method, anatomy and pathology
The EMMA project works with automatic image recognition. "Firstly the system must be taught how to catalogue images. We provide the machine with models that define the different categories of images, and on this basis the system assigns a label to each of them."

Since 2005, Idiap has taken part in a competition relating to sorting and annotating medical images (ImageCLEFmed). The participants, who are researchers from world-wide research laboratories, have several thousand images submitted to them. They test their algorithms on this database and the aim of their system is to annotate the images according to predefined criteria such as the method (x-ray, radio, etc.), the anatomy (part of the body) or the pathology (injury).

In 2009, seven groups from five different countries participated in this medical annotation task and Idiap won the prize for its algorithms, which provided the most convincing results.
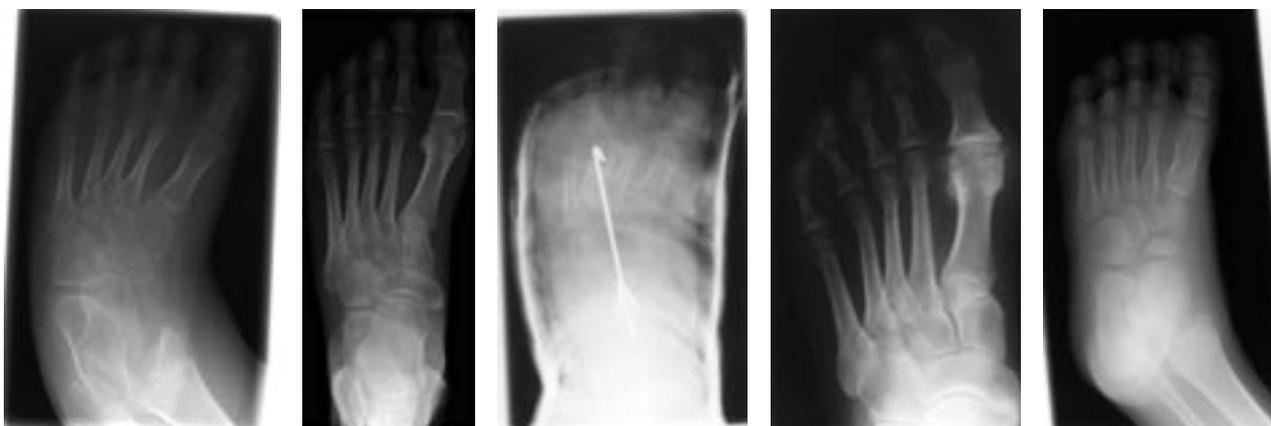
The system on which Idiap is working should soon be capable of sorting images by their type (X-rays, radiography, etc.), by the part of the anatomy that has been examined, and by the pathology detected. The computer is now capable of detecting that these five images are different.

USING MOBILE PHONES TO STUDY SOCIAL ROUTINES

As part of a project carried out in cooperation with Nokia, around one hundred volunteers from French-speaking Switzerland are carrying a mobile phone that records their behaviour. This provides invaluable information for social science researchers.

LS-CONTEXT project	
Large-scale human context discovery for mobile phones	
Idiap	Person in charge: Daniel Gatica-Perez, senior researcher Idiap team: 1 senior researcher, 1 postdoctoral researcher, 1 dev. engineer
Partner	Nokia
Financing	Nokia
Schedule	2009 - 2011

"The mobile phone has become a sort of personal diary that we take everywhere, to which we feed personal information, and that in return could teach us about ourselves."

Daniel Gatica-Perez, Idiap senior researcher, responsible of the project

"Our data for this project is continuously supplied by over 100 volunteers who live in French-speaking Switzerland", explains Daniel Gatica-Perez, senior researcher. "Each of them uses a mobile phone that records their activity. We are conducting this project in collaboration with Nokia Research, Lausanne. The objective is to understand and gain insights about human routines and everyday individual and community behaviour."

Study of social behaviour and urban mobility
Equipped with Bluetooth, GPS, and accelerometer, mobile phones record journeys, their pace, the proximity of other devices, but also data related to the use of the phone, when and how often people make calls, how they use their camera, their music player, etc.

"Nokia is in charge of ensuring anonymity of the data and takes this element of the project very seriously," Daniel Gatica-Perez confirms. "Furthermore, each volunteer can observe and track their own information collected from their phone on the internet."

A large amount of data collected over a long period of time enables researchers to extract highly valuable information about human routines. The research will focus on both social and physical aspects of daily life. The social data will constitute invaluable material for sociologists and other social science researchers. "Furthermore, this data could help answer a certain number of questions on behaviour and social networks, urban mobility, and social habits with regard to mobile phones, etc."

Development of context aware devices

The information related to people's mobility might be used in the future to develop context aware services, the objective being to make mobile phones "smarter" by giving them the capacity to make recommendations. "If your device notes that you go to the cinema every Saturday evening, on a Saturday afternoon it might tell you the programme for movie theaters closest to the place where you are." This recommendation service can be extended to restaurants, shops, etc. "Extracting instantaneous context starts to be feasible task", says Daniel Gatica-Perez smiling. "The challenge will be getting phones to accurately predict what will happen in the future based on months of data used to learn personal habits!"

"In our research, the mobile phone becomes a sort of personal diary that we take everywhere, to which we feed personal information, and that in return teaches us something about ourselves. The result can really give us lots to think about!"

User-labeled location (MIT Reality Mining data)

Celltowers

Home (H), Work (W), Other (O), No data (N)
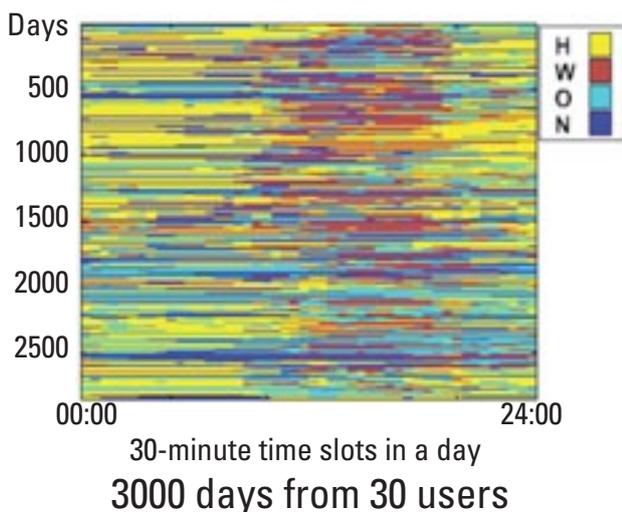
«Going to work around 10am»

Topic 23
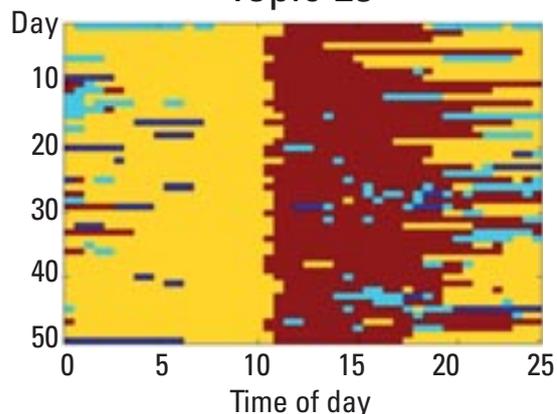
In this example (left), which uses data collected at MIT (Massachussets Institute of Technology) in Cambridge, the abscissa represents the time of the day. On the ordinate, each line corresponds to a day in the life of a person. Working time is shown in red, time spent at home in yellow, and in the blue zones the volunteers are between work and home. The automatic methods aim at extracting interesting patterns of social routines (right). (Image by Kate Farrahi)

A SHOWCASE TO ILLUSTRATE AND CONVINCE

Too often Idiap's work remains a mystery for the general public. Since April 2009, a showroom has been used to demonstrate the research carried out at the Institute to make it more comprehensible to people. An interactive tool that enhances the image of the researchers' work.

It is a place to see, play, discuss and understand… Welcome to the Idiap showroom! As soon as the visitor crosses the threshold, four cameras and eight microphones capture his arrival. His presence is modelled by an avatar in a three-dimensional representation of the room, displayed on a large screen. When he changes position, his avatar moves accordingly and its mouth is animated when he speaks. There is no doubt that he is being monitored. This is a technology developed for research in the field of video surveillance.

Entertaining visit

Florent Monay is in charge and is the main showroom guide. "The showroom is an effective way to illustrate what we do. Instead of sitting visitors in front of a passive presentation, we invite them to interact with the different demonstrations. It is a space that is just as useful to well-informed partners as it is to novices." Four interactive demonstrations are currently accessible on the showroom's computers. With mouse in hand, the public can experiment with some of Idiap's research themes.

Two demonstrations illustrate the various stages of machine learning that vision or speech processing systems may integrate. The first enables an object detector to be created in three stages and the second enables a system that detects voice activity to be obtained. A biometric access control system enables the visitor also to try out the conditions of authentication based on the face and voice. The last demonstration takes the form of a game where a face detector replaces the traditional joystick.

Visitors can also familiarise themselves with the world of Idiap via five short themed films. These presentations describe the Institute's different areas of research in simple and tangible terms. These tools are easy to export and can already regularly be seen at exhibitions or specialised trade fairs.
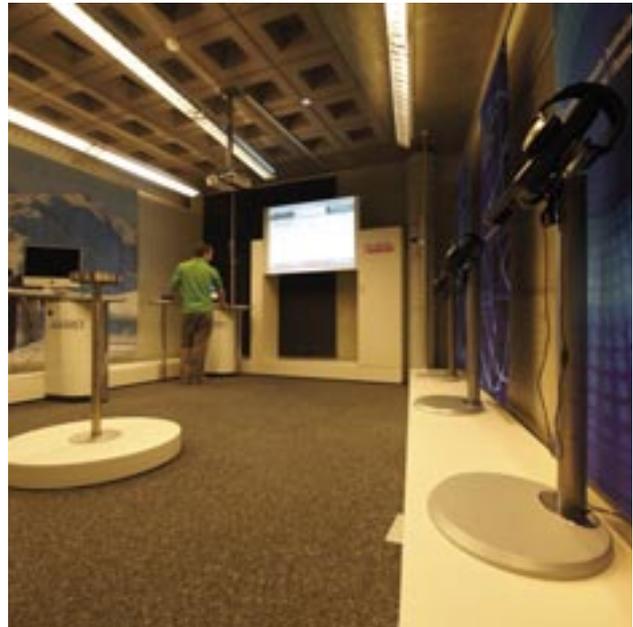
Future developments

The showroom meets a considerable need for visibility and communication. In 2009, it welcomed around thirty groups of visitors. Researchers affiliated to companies as prestigious as Intel, Logitech and NEC have also been able to benefit from this new space. And Florent Monay is convinced that the show is only just beginning! "It is a work in progress. The current demonstrations are still relatively basic. They are going to be enriched by new contributions from researchers so that this space becomes a more representative showcase of our work." The person monitoring demonstration is shortly to be equipped with a voice command system and a new 3D environment. In the longer term, Idiap is considering other suitable ways to illustrate the institute's other research areas, such as analysis of social behaviour and multimedia information systems.

"WE ARE GOING TO BE ABLE TO STRENGTHEN OUR INNOVATIVE CAPACITY"

The strategic alliance with EPFL was announced in February 2008 and took shape in 2009. Idiap's new positioning in the Swiss university network is accompanied by an increase in public financing and beneficial structural changes. Assessment of the operation with Jean-Albert Ferrez, deputy director of Idiap.

What motivated this partnership with EPFL?
This step must be placed in a dual context, that of the Swiss university network and that of Idiap's development. In Switzerland there are cantonal universities, two écoles polytechniques and some smaller institutions with various profiles that Bern recognises and supports. These include Idiap. At the start of the 2000s the size of our institute increased fivefold while its basic financing remained almost the same. This financial imbalance deprived us of exploratory resources. Consequently, the Confederation considered what form of support it could give us and proposed a strategic alliance with our colleagues in Lausanne, accompanied by a fourfold increase in its subsidy for 2008-2011.

**What effects will this alliance
have on your cooperation with EPFL?**
EPFL was one of the initial founders of Idiap and we have worked closely with her for a number of years already. Therefore, this strategic alliance will not fundamentally change our way of working together, but it will further facilitate our cooperation and strengthen our ties on scientific, academic and structural

levels. A number of legal, administrative and geographic obstacles have been overcome to facilitate this partnership.

Will the joint work take place in Lausanne or Martigny?
The people from EPFL have an office, IT structure and access to the laboratories here. The reverse is also true. Idiap's publications are referenced to EPFL and we have access to each other's databases, etc. This environment enables EPFL-Idiap mixed research teams to be formed, which work together between Lausanne and Martigny. Without having merged geographically, it was important for us that the environment allowed total changeability between the two sites.

**Have your research areas
been revised in the context of this partnership?**
Partly. In 2006 Idiap had six research themes: machine learning, speech processing, computer vision, multimedia content management, biometric authentication and multimodal man-machine interaction. In 2008, some new avenues appeared: multilingual processing, new heuristic approaches, social networks, social signal processing and even the use of computers in human communication. EPFL also conducts research activities in areas that are specific to it alone and in some that are common to both of us. Sharing some research areas evidently allows us to work together on projects, but sometimes it has been necessary to avoid redundancy. For example, with the departure of our researcher José Millán, who has gone to the EPFL Neuroprosthetics Centre, we have abandoned the area of cerebral interfaces.

**What will the benefits
of this cooperation be for Idiap employees?**
The alliance stabilises the processes by which Idiap PhD students follow the degree course of the EPFL graduate school. It also opens up new prospects for the researchers. Our senior researchers can now be recognised on an academic level,

teach at EPFL and be granted the title of EPFL lecturer. This is one of the major advantages of this alliance. The processes are underway, but there are still some practical aspects to sort out.

Why is this issue important?

Survival of Idiap relies on its capacity to keep its best researchers in the long term. To achieve this we firstly needed to increase our share of basic financing so that our budget does not depend solely on projects awarded. This has been achieved. Secondly we need a larger scale environment with teaching and academic recognition. Having a lecturer's title is important for a researcher. It lends more weight to his scientific publications and offers him more career prospects.

"Survival of Idiap relies on its capacity to keep its best researchers in the long term. An academic title of lecturer lends more weight to scientific publications and offers more career prospects."

What is your initial assessment of this alliance?

This alliance has pushed us into some beneficial structural reforms. Our autonomy has not suffered from it and our financial continuity has been strengthened. The proportion of our basic financing has reached 34% and now enables us to strengthen our innovative capacity. It has also served as a catalyst to bring Idiap and EPFL researchers closer together, increasing interactions and opportunities for cooperation. Given the state of progress, the academic and structural objectives of the 2008-2011 schedule will be reached to the satisfaction of both institutions. We would like the alliance to continue beyond 2008-2011, which is what we will indicate shortly in our application to Bern for the 2012-2016 subsidy.

ENSURING RENEWAL OF SCIENTIFIC STAFF

Two new positions have appeared in the Idiap organization chart. First tenure track assistant professors (PATT) and senior lecturers (MER) have recently been added to the ranks of the incumbent lecturers. Two offers that should attract the best researchers world-wide and ensure renewal of Idiap's scientific staff.

The American concept of tenure track designates a researcher in the process of appointment to a permanent post. Generally quite young, spotted early and having high potential for advancement, these researchers are integrated into a favourable environment and hired for a six-year trial period. At the term of this period the researcher may or may not be awarded tenure. A bit of a gamble, but also a tool to ensure academic recruitment and renewal of teaching staff for the hautes écoles. Selected and hired by EPFL at the end of 2009, the first EPFL PATT took up his position at Idiap in January 2010 (see page 21). Other tenure track assistant professors should be hired in 2010 or 2011.

At the end of 2009, several Idiap researchers are being assessed under the usual promotion rules of EPFL. During 2010 the institute should therefore gain four or five high-level scientific staff, thus strengthening its stability and visibility and expanding its management team.
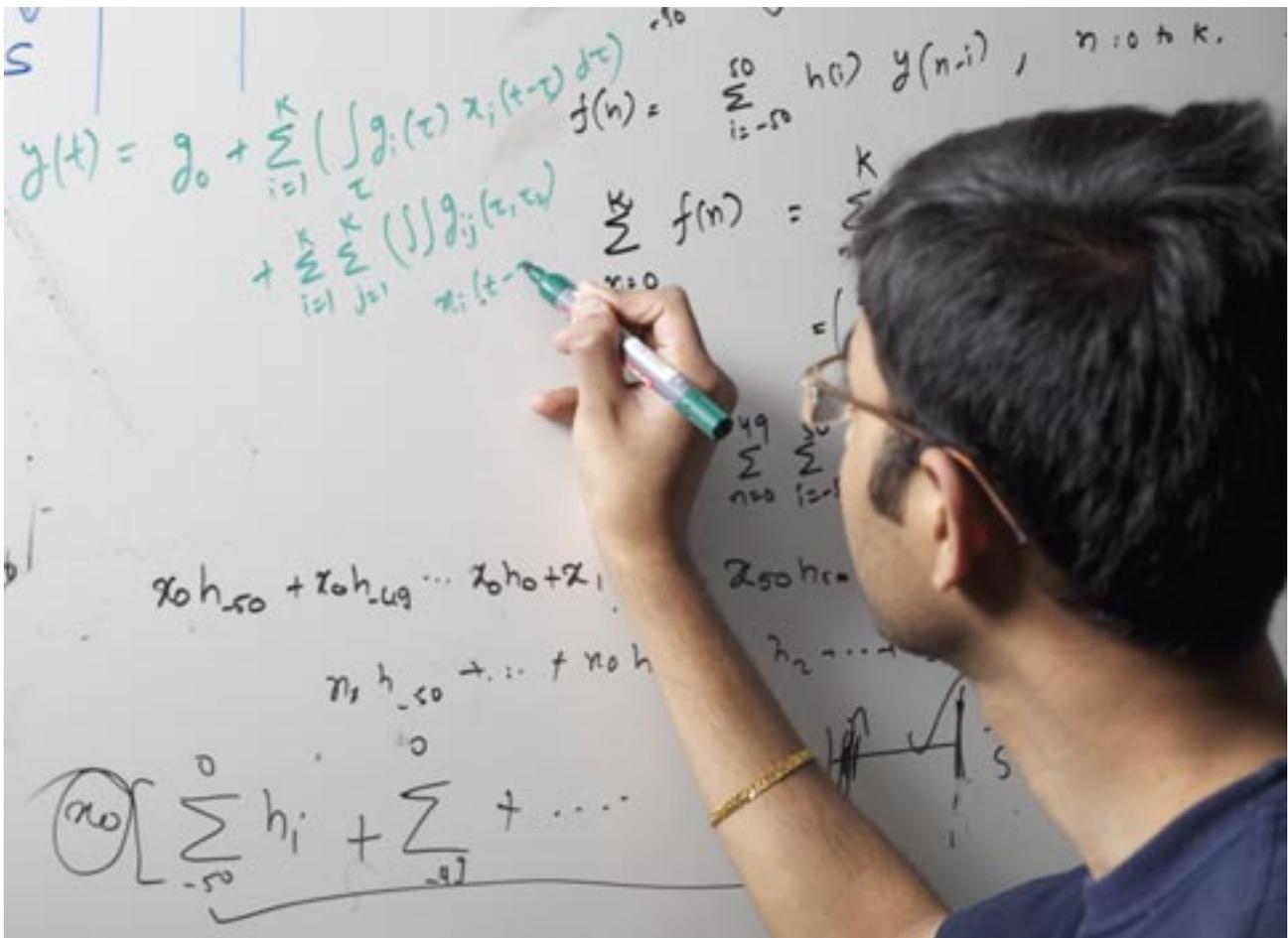
FACES

FIRST TENURE TRACK ASSISTANT PROFESSOR
IDIAP RESEARCHER AND EPFL LECTURER

Volkan Cevher, aged 32, is Idiap's first tenure track assistant professor, an American concept developed to attract the best researchers on the market. Specialised in electrical engineering and a living symbol of the Idiap-EPFL joint development plan, he will be both a researcher in Martigny and will also teach at Lausanne. Profile.

Volkan Cevher was born in 1978 in Ankara (Turkey) and in 1999 obtained a degree in electrical and electronic engineering from Bilkent University in Ankara. He continued his studies at the Department of Electrical and Computer Engineering of Georgia Institute of Technology in Atlanta (USA), where he completed his thesis as a Doctor of Sciences in 2005. This was awarded a prize from the Center for Signal and Image Processing Research of Georgia Institute of Technology for excellence in signal processing research. From 2005 to 2007 he worked as a postdoc at the University of Maryland and then joined the Department of Electrical and Computer Engineering of Rice University.

Volkan Cevher's work on compressive sensing (CS) represents a significant breakthrough in signal processing. His work has enabled signal models to be used that are more structured than the purely parsimonious models used beforehand. Furthermore, his approach, which merges machine learning and CS, offers demonstrable performance guarantees for realistic models that will have fundamental applications, for modelling and compression of natural images for example.

The appointment of Volkan Cevher is part of the scientific partnership agreed between EPFL and the Idiap Research Institute in Martigny in the extensive field of signal processing.

(Source: EPFL, Sarah Perrin)

THE MEN IN THE SHADOWS WHO SWITCH THE LIGHTS BACK ON

They are indispensable and yet we only notice them when there is a problem. The men of the system team look after Idiap's infrastructures and equipment. They are there to help with the slightest problem, develop IT services according to need and provide the resources needed for the work of the Institute's employees. Interview with six guardian angels.

A chair missing, a faulty mouse, a screen to be set up for a demonstration, a recording to be made, an archiving or email problem, a program that does not work properly… We call in the system team. These simple problems are generally solved within an hour but more complex questions can take more time. Sometimes a few hours or even several weeks are needed to resolve a sensitive issue. The helpdesk, which centralises all questions and problems, deals with a number of requests each day. "We are a small team, which has its limits, but we try to set up systems that can meet a maximum of needs", explains Frank Formaz, the head of the team. "However, sometimes there is a discrepancy between the ideas of the researchers and the technical reality. Sometimes just lending a hand can become a really major job! In this case we have to find a compromise between these needs, what is possible for us and our resources."

Disk space multiplied by 1000 in ten years
The six men of the system team are not just simple repairmen. They are very versatile and follow technological innovations, look for novel solutions and implement strategies to ensure that the Institute runs smoothly. The continuity of data is one of their main tasks. "Idiap employees come and go. We are here to make sure their work is reusable", clarifies Frank Formaz. Everything that is produced at Idiap must be able to be found again and used. This is why everything is set up to maintain and safeguard these data. And quality is only a part of the problem: in ten years, the quantity of disk space used has multiplied by 1000! "How do we develop the whole system while maintaining access to old data? Storage is our greatest current challenge", says Frank Formaz smiling.
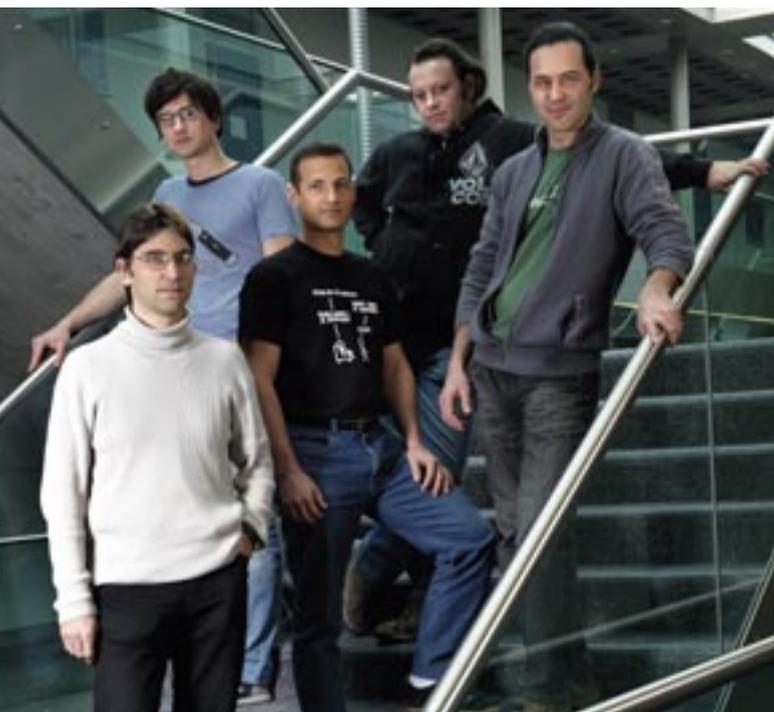
The team's three fields of activity

1. System
- Installation, maintenance and updating the institute's infrastructure (IT, premises, etc.). Maintenance of all the IT services (calculations, storage, email, internet, printing, etc.)

2. Support
- Helpdesk: centralisation and gathering of all user problems or questions (approximately one thousand requests in 2009)
- Helping researchers install equipment and acquisition, processing, saving and circulating data, etc.

3. Internet services
- Essential to Idiap's communication! Regular updates, introduction of new information and creation of new sites to meet new needs.

Frank Formaz, the super coordinator

Aged 36, lives in Fully
At Idiap for 12 years
Training: ETS Engineer
Passions: music, reading and family activities with his three children

"I form the link between needs, people, tools and resources... The binding agent of the group as it were! I am responsible for the centralisation of purchases and ensure that budgets are kept to. I try to avoid being continuously on call because I have to stay abreast of trends to anticipate future problems. Our work is not very visible and users tend to take it for granted. It is really best not to expect any recognition in order to do this job!"

Norbert Crettol, the zen sysadmin

Aged 57, lives in Martigny
At Idiap for 8 years
Training: Arts student, got into IT due to a passion for it
Passions: guitar, stories, mountains, forests and aikido for greater self-knowledge. Ideal for coming down from the virtual clouds of IT!

"I resolve user problems and look after some of the servers and the Internet in particular. I play an interface role between pure technology and the aesthetic window of the webmaster. I am the technician behind the artist! This role in the shadows suits me well."

Cédric Dufour, the man who likes a challenge

Aged 37, lives in Verbier
At Idiap for 4 years
Training: EPF Engineer
Passions: all outdoor activities such as windsurfing, aviation, cycling and skiing

"I look after the institute's central and critical services. We install, maintain, update and debug. I also participate in defining strategies. I like to bring new ideas and take up challenges. Complex problems are right up my alley!"

Tristan Carron, the face of the Helpdesk

Aged 38, lives in Martigny
At Idiap for 7 years
Training: Business Information Manager
Passions: IT in all its forms, from code to games to website design

"I am the only member of the team that everyone knows. This is natural as I am responsible for the helpdesk. I try to meet as many requests as possible to avoid burdening my colleagues who concentrate on other tasks. I like working with people and here, with 29 different nationalities, it suits me perfectly! I have even kept in touch with people who left Idiap 4 years ago."

Vincent Spano, the online artist

Aged 47, lives in Martigny-Croix
At Idiap for 6 years
Training: webmaster
Passions: all board sports and family activities with his three children

"I am the one that opens Idiap's doors to the public. And my greatest pleasure in managing the website is that I am very free to make my own choices! I take care of the content and updates, both for the graphics part and the collection of text. I work in the same office as the project managers, which allows me to keep up-to-date with the Institute's activities."

Bastien Crettol, the researchers' assistant

Aged 29, lives in Sion
At Idiap for 5 years
Training: Business Information Manager
Passions: American literature, cinema, music

"I am here to help the researchers and I have the opportunity to work directly with them on projects. It is not always easy to give the right technical answers, but this connection with the academic environment is very rewarding. It is not a closed world as one might believe. At Idiap, I work with sociable scientists from very different backgrounds."

EMPLOYEES LEAVING AND JOINING

In 2009 the Idiap science team took on eight new talented researchers, two postdoctoral researches, five PhD students and one development engineer. Thirteen left the institute to take on new challenges elsewhere.

THEY ARRIVED IN 2009

First name, last name, position, country of origin, residence

Oya Aran Karakus, Postdoc, Turkey
Trinh-Minh-Tri Do, Postdoc, Vietnam
Charles Dubout, Research Assistant, Switzerland, Renens
David Imseng, Research Assistant, Switzerland, Raron
Gelareh Mohammadi, Research Assistant, Iran
François Moulin, Development Engineer, Switzerland, Vollèges
Dairazalia Sanchez-Cortez, Research Assistant, Mexico
Serena Soldo, Research Assistant, Italy

THEY LEFT

First name, last name, position, country of origin, year of arrival at Idiap, new employer

Silèye Ba, Postdoc, Senegal, 2002, Telecom Bretagne, France
Mike Flynn, Senior Researcher Scientist, England, 2003
Sriram Ganapathy, Research Assistant, India, 2006, John Hopkins University, Baltimore, USA
Sri Venkata Surya Sivaramakrish Garimella, Research Assistant, India, 2007, John Hopkins University, Baltimore, USA
Guillaume Heusch, Research Assistant, Switzerland, St-Légier, 2005, ViNotion, Eindhoven, Netherlands
Joseph Keshet, Postdoc, Israel, 2007, Toyota Technological Institute, Chicago, USA
Stéphanie Lefèvre, Research Assistant, France, 2007, Renault s.a.s, Boulogne Billancourt, France
Francesco Orabona, Postdoc, Italy, 2007, Università degli Studi di Milano, Italy
Elisa Ricci, Postdoc, Italy, 2008, Bruno Kessler Foundation, Povo, Italy
Nicolas Scaringella, Research Assistant, Italy, 2006, Museeka SA, Geneva, Switzerland
Hugues Salamin, Research Assistant, Switzerland, Réchy, 2007, University of Glasgow, Dpt of Computing Science, Glasgow, UK
Samuel Thomas, Research Assistant, India, 2007, John Hopkins University, Baltimore, USA
Johnny Mariéthoz, Dev. Engineer, Switzerland, Chemin-Dessus, 1998, RERO, Martigny, Switzerland

HONOURS

Each year, Idiap awards two awards to its PhD students. The first rewards research and the second a publication. In order to award the Idiap PhD student Paper Award, the candidate is assessed by an in-house committee on the basis of five criteria: his publications, participation in the team, involvement in the project, communication skills and autonomy. For the PhD student Research Award, an initial selection is made by senior members of the institute from amongst the works mainly written by an Idiap PhD student. Members of the International strategic committee then grade the chosen publications, separately and anonymously.

In 2009, the Research award was given to two PhD students, **Dinesh Babu Jayagopi** and **Hugues Salamin** and the Paper award to PhD students **Jie Luo** and **Joel Pinto**.

Jie Luo
recipient of 2009 Idiap PhD student Paper Award

«Who's Doing What»
Joint Modeling of Names and Verbs for simultaneous Face and Pose Annotation (to appear in NIPS 2009)
Jie Luo, Barbara Caputo, Vittorio Ferrari

Dinesh Babu Jayagopi and Hugues Salamin
recipient of 2009 Idiap PhD student Research Award

Joel Pinto (not in the photo)
recipient of 2009 Idiap PhD student Paper Award

«Analysis of MLP Based Hierarchical»
Phoneme Posterior Probability Estimator (to be published in IEEE Trans. on Audio, Speech and Language Processing)
Joel Pinto, Sivaram Garimella, Mathew Magimai Doss, Hynek Hermansky, Hervé Boulard

COMPLETED THESES

Two students completed their thesis in 2009: Guillaume Heusch, from St-Légier (VD) and Nicolas Scaringella. This last one will start work in 2010 at the Museeka start-up in Geneva, for which Idiap has worked on the development of new software for processing music.

- **Bayesian Networks as Generative Models for Face Recognition**
  Guillaume Heusch, 17 novembre 2009
  Thesis director: Dr Sébastien Marcel et Prof. Hervé Bourlard
  Members of the Thesis Committee: Prof. Jean-Philippe Thiran, Prof. Stan Z. Li, Prof. Massimo Tistarelli

- **On the design of audio features robust to the album-effect for music information retrieval**
  Nicolas Scaringella, 29 avril 2009
  Thesis director: Prof. Hervé Bourlard
  Members of the Thesis Committee: Prof. Pierre Vandergheynst, Dr Christof Faller, Prof. Robert Leonardi, Dr Goeffroy Peeters
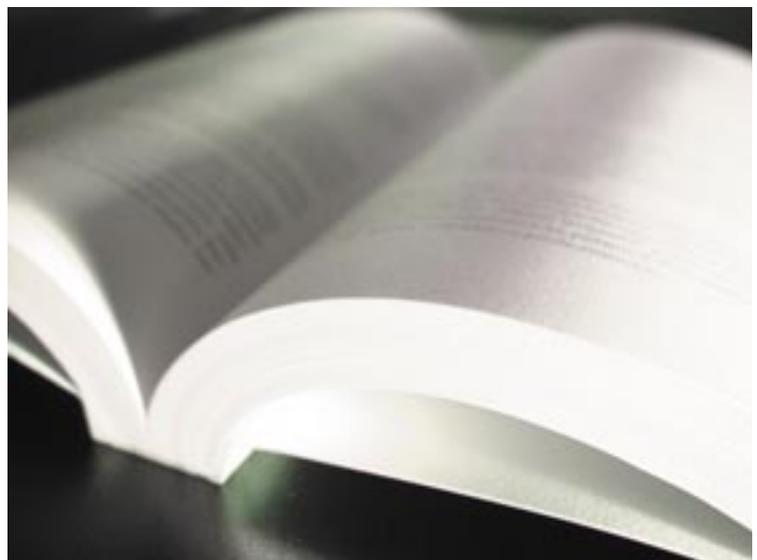
OPERATING ACCOUNT

(Swiss Francs)	2008	2009	%
INCOME			
Town of Martigny	520,000	600,000	6.51%
Canton of Valais	1,200,000	1,000,000	10.85%
Swiss Confederation	900,000	1,510,000	16.38%
TOTAL SUBSIDIES	**2,620,000**	**3,110,000**	**33.74%**
Loterie romande	150,000	150,000	1.63%
EPFL Contribution	103,667	72,000	0.78%
TOTAL DONATIONS - ALLOWANCES	**253,667**	**222,000**	**2,41%**
NCCR IM2 projects	1,477,423	1,331,107	14.44%
Swiss National Science Foundation projects	844,879	965,768	10.47%
European projects	3,415,514	2,452,661	26.60%
CTI projects	79,831	323,097	3.50%
TOTAL PROJECTS	**5,817,647**	**5,072,633**	**55,01%**
Industrial Financing and other income	1,133,392	815,324	8.84%
TOTAL INCOME	**9,824,706**	**9,219,957**	**100.00%**

(Swiss Francs)	2008	2009	%
EXPENSES			
Personnel expenses	6,772,575	6,334,515	68.70%
Training and travel	532,598	502,869	5.45%
Third party expenses	376,882	415,130	4.50%
Computer equipment and maintenance	279,142	199,486	2.16%
Administrative costs	133,017	178,333	1.94%
Promotion and communication	112,552	75,639	0.82%
Rent	710,602	823,187	8.93%
Depreciation	119,794	266,278	2.89%
Other provisions	551,000	397,000	4.31%
TOTAL EXPENSES	**9,588,162**	**9,192,437**	**99.70%**
OPERATING PROFIT / LOSS	**236,544**	**27,520**	**0.30%**

Distribution of sources of financing

- City of Martigny
- Canton of Valais
- Swiss Confederation
- Loterie romande
- NCCR IM2 projects
- Swiss National Science Foundation projects
- European projects
- CTI projects
- EPFL Contribution
- Industrial Financing and other income

Distribution of costs

- Personnel expenses
- Training and Travel
- Third party expenses
- Computer equipment and maintenance
- Administrative costs
- Promotion and communication
- Rent
- Depreciation
- Other provisions
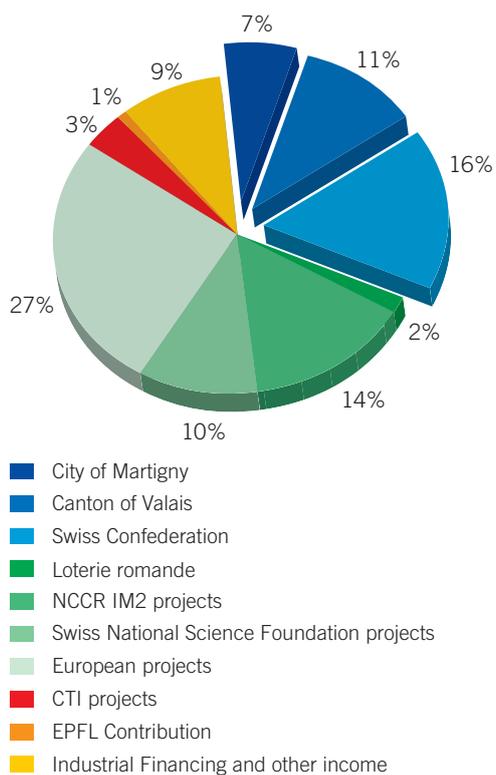
Comments on the 2009 accounts

Idiap ended 2009 with a favourable bottom line. Overall volume, both income and expenses, is down compared to 2008, which was an exceptional year. The reduction of almost one million Swiss francs in the European projects entry is due both to the loss of certain projects and to accounting operations. The 2008 and 2009 sums from the Lottery of French-speaking Switzerland involve two annual tranches for the same project: the showroom (cf. page 15). The payroll represents more than two thirds of the total. The increase in rent reflects the increase in area occupied at Centre du Parc.

The balance sheet reflects the solid situation of the institute, marked on the one hand by considerable liquidity arising from the method of prefinancing European projects, and on the other hand by a reserve for fluctuation in contracts, which takes account of a risk weighting linked to different sources of financing.

Swiss Confederation, Canton, Municipality subsidies

(In thousands of Swiss francs)

YEAR	2008	2009	2010	2011	Total
Confederation	900	1,510	1,795	2,357	**6,562**
Canton	1,200	1,000	900	900	**4,000**
Municipality	550	600	600	650	**2,400**
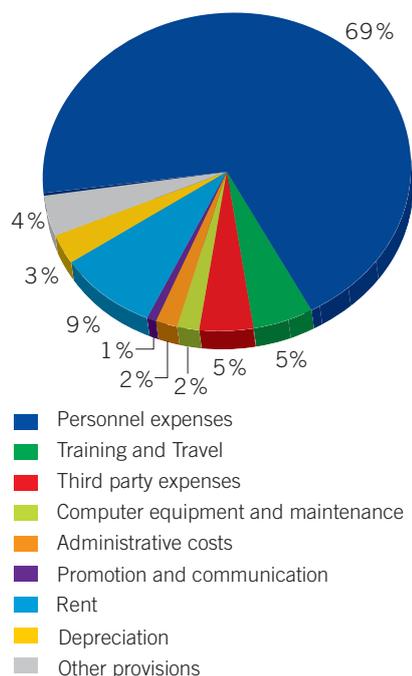
Following the agreement signed with the State Secretariat for Education and Research (SER), which provides for a gradual increase in the federal subsidy, the Canton of Valais and the Town of Martigny have agreed to provide together an almost equivalent amount, in accordance with the distribution given in the table above.

BALANCE SHEET

(Swiss Francs)	31.12.2008	31.12.2009
ASSETS		
Cash	3,951,731.76	2,762,410.81
Accounts receivable	178,849.76	240,450.85
Accrued income and other	360,645.34	520,640.79
TOTAL CURRENT ASSETS	**4,491,226.86**	**3,523,502.45**
Equipment	328,375.00	528,219.05
Financial assets	50,000.00	50,000.00
TOTAL NON-CURRENT ASSETS	**378,375.00**	**578,219.05**
TOTAL ASSETS	**4,869,601.86**	**4,101,721.50**
LIABILITIES		
Accounts payable	297,460.01	277,233.39
Accrued expense	2,905,876.51	1,733,703.29
Provisions	181,000.00	578,000.00
TOTAL FOREIGN FUNDS	**3,384,336.52**	**2,588,936.68**
Share capital	40,000.00	40,000.00
Special reserve	1,200,000.00	1,200,000.00
Retained earnings	8,721.25	245,265.34
Net income	236,544.09	27,519.48
TOTAL OWN FUNDS	**1,485,265.34**	**1,512,784.82**
TOTAL LIABILITIES	**4,869,601.86**	**4,101,721.50**

ORGANIZATION

| FOUNDATION COUNCIL | | AUDITING BODY |

| PRESIDENT OF THE FOUNDATION COUNCIL |

| SUPPORT COMMITTEE |

DIRECTION

| DIRECTOR | | ASSISTANT TO THE DIRECTOR |

| MEMBERS OF MANAGEMENT | | DEPUTY DIRECTOR |

| SENIOR RESEARCHERS *(with EPFL academic recognition)* | | EXECUTIVE OFFICERS |

| SENIOR RESEARCHERS, RESEARCHERS, POSTDOCTORAL RESEARCHERS, PHD STUDENTS, ETC. | | ADMINISTRATION AND SERVICES |
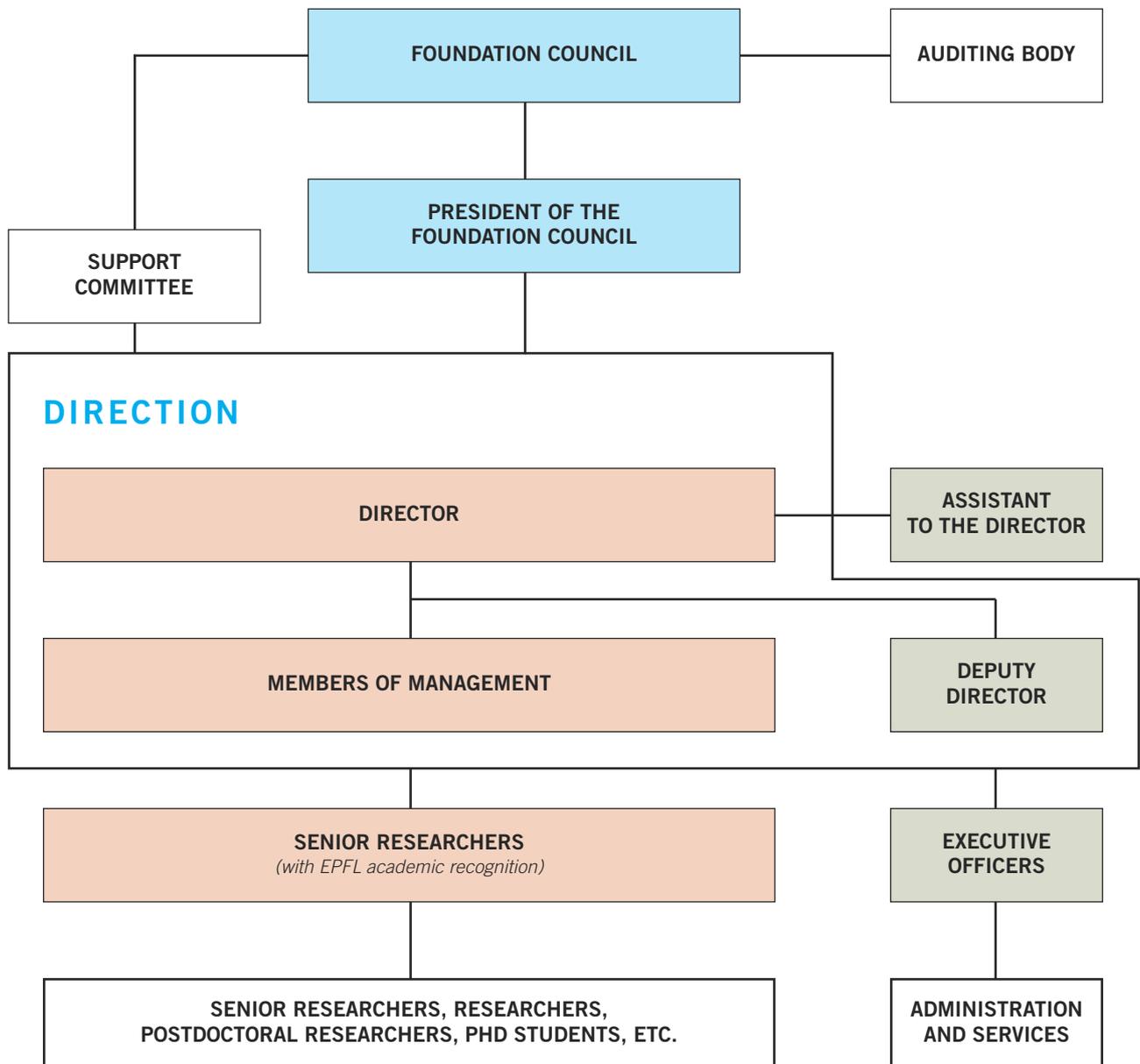
Members of the scientific board
Members of the administrative board

EMPLOYEES

Scientific staff

First name, last name, position, country of origin, residence, arrived

Lakshmi Achuthankutty, Research Assistant, India, 2008
Oya Aran Karakus, Postdoc, Turkey, 2009
Afsaneh Asaei, Research Assistant, Iran, 2008
Constantin-Cosmin Atanasoaei, Research Assistant, Romania, 2008
Venkatesh Bala Subburaman, Research Assistant, India, 2007
Joan Isaac Biel, Research Assistant, Spain, 2008
Hervé Bourlard, Director, Belgium, 1996
Barbara Caputo, Senior Research Scientist, Italie, 2005
Alfred Dielmann, Postdoc, Italy, 2008
John Dines, Senior Research Scientist, Australia, 2003
Trinh-Minh-Tri Do, Postdoc, Vietnam, 2009
Charles Dubout, Research Assistant, Renens, Switzerland, 2009
Stefan Duffner, Postdoc, Germany, 2008
Katayoun Farrahi, Research Assistant, Canada, 2007
Sarah Favre, Research Assistant, Switzerland, Nendaz, 2006
François Fleuret, Senior Research Scientist, France, 2007
Giulia Garau, Postdoc, Italy, 2008
Philip Garner, Senior Research Scientist, England, 2007
Daniel Gatica-Perez, Senior Research Scientist, Mexico, 2002
Hayley Shi-Wen Hung, Postdoc, England, 2007
David Imseng, Research Assistant, Raron, Switzerland, 2009
Dinesh Babu Jayagopi, Research Assistant, India, 2007
Niklas Johansson, Research Assistant, Sweden, 2008
Danil Korchagin, Postdoc, Russia, 2008
Hui Liang, Research Assistant, China, 2008
Jie Luo, Research Assistant, China, 2007
Mathew Magimai Doss, Resarch Scientist, India, 2007
Sébastien Marcel, Senior Research Scientist, France, 2000
Christopher McCool, Postdoc, Australia, 2008
Gelareh Mohammadi, Research Assistant, Iran, 2009

Petr Motlicek, Research Scientist, Czech Republic, 2005
Radu-Andrei Negoescu, Research Assistant, Romania, 2007
Jean-Marc Odobez, Senior Research Scientist, France / Switzerland, Clarens, 2001
Sree Hari Krishnan Parthasarathi, Research Assistant, India, 2007
Hugo Augusto Penedones Fernandes, Research Assistant, Portugal, 2008
Joel Praveen Pinto, Research Assistant, India, 2005
Andrei Popescu-Belis, Senior Research Scientist, France / Romania, 2007
Edgar Francisco Roman Rangel, Research Assistant, Mexico, 2008
Anindya Roy, Research Assistant, India, 2007
Dairazalia Sanchez-Cortez, Research Assistant, Mexico, 2009
Serena Soldo, Research Assistant, Italy, 2009
Nicolae Suditu, Research Assistant, Romania, 2008
Tatiana Tommasi, Research Assistant, Italy, 2008
Fabio Valente, Research Scientist, Italy, 2005
Jagannadan Varadarajan, Research Assistant, India, 2008
Deepu Vijayasenan, Research Assistant, India, 2006
Alessandro Vinciarelli, Senior Research Scientist, Italy, 1999
Majid Yazdani, Research Assistant, Iran, 2008

Development engineers

Philip Abbet, Dev. Engineer, Switzerland, Conthey, 2006
Olivier Bornet, Senior Dev. Engineer, Switzerland, Nendaz, 2004
Maël Guillemot, Dev. Engineer, France, 2002
Christine Marcel, Dev. Engineer, France, 2007
Olivier Masson, Dev. Engineer, Switzerland, Geneva, 2002
Florent Monay, Dev. Engineer, Switzerland, Monthey, 2008
François Moulin, Dev. Engineer, Switzerland, Vollèges, 2009
Alexandre Nanchen, Dev. Engineer, Switzerland, Martigny, 2008
Flavio Tarsetti, Dev. Engineer, Switzerland, Martigny, 2008

Administrative staff

Céline Aymon Fournier, Public Relations, Switzerland, Fully, 2004
Valérie Devanthéry, Ass. Program Manager, Switzerland, St-Maurice, 2008
Jean-Albert Ferrez, Deputy Director, Switzerland, Martigny, 2001
Pierre Ferrez, Program Manager, Switzerland, Verbier, 2004
François Foglia, Program Manager, Switzerland, Saxon, 2006
Edward-Lee Gregg, Financial Assistant, United States, 2004
Sandra Micheloud, Financial Director, Switzerland, Monthey, 2007
Sylvie Millius, Secretary, Switzerland, Vétroz, 1996
Yann Rodriguez, Industrial Relations, Switzerland, Martigny, 2006
Nadine Rousseau, Secretary, Belgium, 1998

System engineers

Tristan Carron, System Administrator, Switzerland, Martigny, 2003
Bastien Crettol, System Administrator, Switzerland, Sion, 2005
Norbert Crettol, System Administrator, Switzerland, Martigny, 2002
Cédric Dufour, System Administrator, Switzerland, Verbier, 2007
Frank Formaz, System Manager, Switzerland, Fully, 1998
Vincent Spano, Webmaster, Switzerland, Martigny-Croix, 2004

Trainees

First name, last name, origin, institution of origin

Idiap trainees generally spend between six and ten months at the research institute. Some are students at EPFL (Ecole poly-technique fédérale de Lausanne) and do this work placement as part of their degree work. Others come as part of student exchange programmes set up within European projects in which Idiap participantes.

Vincent Bozzo, Switzerland, EPFL, Lausanne, Switzerland
Gokul Chittaranjan T., India, University of Florida, Gainesville, United States
Marco Fornoni, Italy, Università degli Studi of Milan, Italy
Nikhil Garg, India, EFPL, Lausanne, Switzerland
Quoc Anh Le, Vietnam, University of Scheffield, England
Benjamin Picard, Belgium, Faculté polytechnique de Mons, Belgium
Jordi Sanchez-Riera, Spain, Polytechnic University of Catalonia, Barcelona, Spain
Xiangxin Zhu, China, Chinese Academy of Science, Beijing, China

Visitors

First name, last name, institution of origin

Visitors are researchers or manufacturers who only spend a few days or weeks at the institute, some to strengthen interinstitutional links and others to get an incite into the work carried out by the institute.

Nicolo Cesa-Bianchi, University of Milan La Statale, Italy
Perttu Laurinen, University of Oulu, Finland
Marc Mehu, Geneva University, Switzerland
Marcello Mortillaro, Geneva University, Switzerland
Nicoletta Noceti, Università di Genova, Italy
Khiet Truong, University of Twente, Netherlands

FOUNDATION COUNCIL

The Foundation Council is responsible for the economic and financial management of the research institute, defines its structures, appoints its director, and generally ensures that the foundation develops successfully by defending its interests. In 2009, following the municipal elections, the Council welcomed a twelfth member in the person of the new President of the Town of Martigny, Marc-Henri Favre.
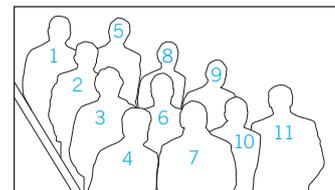
7 **Mr Olivier Dumas,** President
Director of Electricité Emosson SA

11 **Mr Jean-Daniel Antille,** Vice-president
Manager of the regional office
for the economic development of French-speaking Valais

8 **Prof. Martin Vetterli,** Vice-president
Vice-president for international relations,
Ecole polytechnique fédérale de Lausanne (EPFL)

4 **Mr Jean-Pierre Rausis,** Secretary
Managing Director of BERSY Consulting

1 **Mr Josy Cusani**
President of CimArk SA

2 **Prof. Jean-Jacques Paltenghi**
Inter-institutional relations delegate,
Ecole polytechnique fédérale de Lausanne (EPFL)

3 **Mr Pierre Crittin**
Notary

5 **Dr Bertrand Ducrey**
Director of Debio Recherche pharmaceutique SA

6 **Mr Stefan Bumann**
Head of tertiary sector training,
Department of education, culture and sports (DECS)

9 **Mr Daniel Forchelet**
Swisscom Innovations

10 **Mr Jean-René Germanier**
National Councillor

Mr Marc-Henri Favre (not in the photo)
President of the Town of Martigny

Prof. Christian Pellegrini (not in the photo)
Director of the IT department,
University of Geneva

NEW MEMBER OF THE FOUNDATION COUNCIL

Elected in February 2009 to the presidency of the Town of Martigny, Marc-Henri Favre naturally became a member of the Idiap Foundation Council. A chance to ask him briefly to give us his view of the institute.

"BOTH A SPEARHEAD OF THE TERTIARY SECTOR AND A WINDOW ONTO THE WORLD"

Marc-Henri Favre,
President of the Town of Martigny

The position of Idiap in Martigny
"The Town of Martigny is extremely proud to be the home of a research institute of university level and such international renown as Idiap.

On an economic level, the institute represents one of the spearheads of the tertiary sector in our region. This is a role that it shares, in other fields, with Groupe Mutuel and Flagstone Réassurance – for commerce – and with Debiopharm for the pharmaceuticals sector.

The Municipality is very attached to Idiap and demonstrates this by the financial support it grants to it, which is by no means called into question."

Development projects
"Today, almost twenty years after its creation, Idiap is so well rooted in Martigny that it constitutes the basis on which we are going to build and develop a technology centre. Although we are still in the first stages of deliberation and design, we are making good progress and I am looking forward to seeing a "technology park" take form in Martigny.

The IdeArk incubator, in which the town is a shareholder, also comprises an important element in the dynamics of the Idiap Research Institute. As part of The Ark, which promotes

innovation as the driving force behind boosting the Valais economy, IdeArk plays an important role in Martigny. Thanks to IdeArk a number of start-ups have been created, synergies developed, and the reputation of Idiap has grown."

Benefits for the town
"Lastly, interest in Idiap's presence in Martigny is not limited to the financial rewards that it brings. Following the example of Fondation Gianadda in the cultural field, Idiap offers the town a window open onto the world. Due to these two leading foundations, Martigny benefits from global visibility in art, culture and science circles, and in research in particular.

Through its multi-culturalism, Idiap also gives the town a touch of exoticism. I often meet African, South American or Asian researchers in restaurants or in the street and this is always a pleasure for me!

In conclusion, I would like to thank Idiap's management and the whole team for their dynamism, their creativity and their contribution to life in Martigny. I wish them a very successful future!"

ADVISORY BOARD

The Advisory Board is comprised of members of the scientific community chosen by Idiap's management for their exceptional skills and avant-garde outlook. Although their role is strictly advisory, their support and advice is frequently sought and often proves to be invaluable when making decisions on matters of research, training and technology transfer.

Dr Jordan Cohen
Independent Consultant, Spelamode
Half Moon Bay, CA, USA

Prof. Dr Donald Geman
Professor of Mathematics, Johns Hopkins University
Baltimore, USA

Dr John Makhoul
Chief Scientist, Speech and Signal Processing, BBN Technologies
Cambridge, MA, USA

Prof. Nelson Morgan
Director of the International Computer Science Institute (ICSI)
Berkeley, CA, USA

Dr David Nahamoo
Senior Manager, Human Language Technologies, IBM Research
Yorktown Heights, New-York, USA

Prof. Dr Bayya Yegnanarayana
Professor and Microsoft Chair International Institute
of Information Technology (IIIT)
Hyderabad, India

Dr HongJiang Zhang
Managing Director
Microsoft Research Asia Advanced Technology Center
Beijing, China

TOWN OF MARTIGNY

CANTON OF VALAIS

SWISS CONFEDERATION
State Secretariat for Education and Research (SER)

www.loterie.ch

www.swisscom.com

Swiss Power Group.

www.groupemutuel.ch

www.epfl.ch

www.theark.ch

www.ideark.ch

www.snf.ch
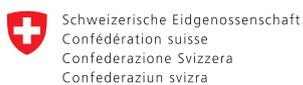
Innovation Promotion Agency CTI

www.bbt.admin.ch/kti

SEVENTH FRAMEWORK PROGRAMME

cordis.europa.eu/fp7

SCIENTIFIC INSERTS

As announced in last year's report, to face its continuous growth, while still fostering internal multi-disciplinary collaborations, Idiap reorganized its internal structuring of its research themes along the following dimensions. Idiap has thus changed/adapted the way it presents itself and describes its current activities, to take into account the new areas of development not only towards human-computer interaction but also towards human-to-human interaction, collaboration, behaviour, and innovation. Thus, after several (13) years of positioning itself under the general theme of "Multimodal human-computer interaction", Idiap decided to officially cover a larger research domain, now referred to as "**Human and Media Computing**".

Articulated around our current activities, "Human and Media Computing" now covers the following research themes:

• **Perceptual and cognitive systems**; Speech processing; Natural language understanding and translation; Document and text processing; Vision and scene analysis; Multimodal processing; Cognitive sciences.

Idiap combines its multi-disciplinary expertise to advance the understanding of human perceptual and cognitive systems, engaging in research on multiple aspects of human-computer interaction with computational artefacts such as natural language understanding and translation, document and text processing, vision and scene analysis, multimodal interaction, computational cognitive systems, and methods for automatically training such systems (see our research efforts in machine learning).

• **Social/human behaviour**: Web social media; Mobile social media; Social interaction sensing; Social signal processing; Verbal and nonverbal communication analysis.

Social Signal Processing is the domain aimed at automatic understanding of social interactions through analysis of nonverbal behaviour.

• **Information interfaces and presentation**: Multimedia information systems, User interfaces; System evaluation.

Information processing by computers must be accompanied by the capacity to present results to users in an efficient and usable way, using human-computer interfaces. In the case of interactive systems, these interfaces must also allow users to enter information in a simple and reliable way, and in the most advanced cases to acquire information from users in a non-disruptive ways. Current research directions at Idiap

focus on multimedia information systems, user interfaces, and the evaluation of interactive systems, explained below in more detail.

• **Biometric person recognition**: Speaker identification/verification; Face detection/identification/verification; Multimodal biometric user authentication.

Conventional means of identification such as passwords, secret codes and personal identification numbers (PINs) can easily be compromised, shared, observed, stolen or forgotten. However, a possible alternative in determining the identities of users is to use biometrics.

Biometric person recognition refers to the process of automatically recognizing a person using distinguishing behavioural patterns (gait, signature, keyboard typing, lip movement, hand-grip) or physiological traits (face, voice, iris, fingerprint, hand geometry, electroencephalogram -- EEG, electrocardiogram -- ECG, ear shape, body odour, body salinity, vascular). Over the last decades, several of these biometric modalities have been investigated (fingerprint, iris, voice, face) and are still under consideration. More recently, novel biometric modalities have emerged (gait, EEG, vascular) mainly due to the development of sensor technologies.

Biometric person recognition offers a wide range of challenging fundamental and concrete problems in image processing, computer vision, pattern recognition and machine learning. It is thus a truly inter-disciplinary research field.

• **Machine learning**: Statistical and neural network based machine learning; Computational efficiency, targeting real-time applications; Very large datasets; Online learning.

Research in machine learning aims at developing computer programs able to learn from examples. Instead of relying on a careful tuning of parameters by human experts, machine learning techniques use statistical methods to directly estimate the optimal setting, which can hence have a complexity beyond what is achievable by human experts.

1. Machine learning and signal processing

Leading researchers: François Fleuret, Barbara Caputo

Machine learning still plays a central place in all Idiap's activities, both as a common tool to solve very large, real-life, real-scale problems, and as a research topic.

Machine learning is applied with great success to research areas such as the automatic analysis of social behaviour, large scale human behaviour modelling, or autonomous cognitive agents, where we pioneer the use of sophisticated multi kernel online learning algorithms for building semantic representations of space. Generic tools developed at Idiap are publicly available through www.torch.ch, and keep being regularly enhanced and updated.

As planned in last year's activity report, new activities have also been initiated in 2009 in the area of "new machine learning algorithms integrating heuristics". Indeed, specifically started in 2009, the MASH project (initiated and coordinated by Idiap) funded by the EU (http://mash-project.eu), and the VELASH project funded by the SNF started recently to investigate the collaborative development of machine learning with very large families of features extractors. They both aim at developing novel tools to allow large groups of individuals to design jointly very complex intelligent systems for computer vision and robotics. With a total workforce of more than 400 person-months, this research initiative will provide key results on the potential of such approaches for the next-generation artificial intelligence.

2. Audio and speech processing

Leading researchers: Hervé Bourlard, Phil Garner, John Dines, Mathew Magimai-Doss, Petr Motlicek, Fabio Valente

Idiap keeps being recognized as one of the key leading institutions in audio and speech processing and retains an expertise in areas such as improved robustness, better modelling of the time/frequency structure of the speech signal, portability across new applications, automatic adaptation, confidence measures, keyword spotting, out-of-vocabulary words, acoustic modelling of temporal dynamics and speaker turn detection (using acoustic features and/or source localization features).
Besides further development, and adaptation to multiple applications, the resulting leading edge software (acoustic feature extraction and continuous speech recognition) are

currently being released through open source libraries:
• "Tracter" Audio processing (incl. beamforming, acoustic feature extraction): http://juicer.amiproject.org/tracter/
• "Juicer" Continuous Speech Recognition: http://juicer.amiproject.org/juicer/

Our research interest in speech processing also involves automatic detection of keywords, i.e. to identify words (or phrases) of interest in unconstrained speech recordings. This is defined in general as Spoken Term Detection (STD), where terms do not have to necessarily be defined a-priori. The STD system developed at Idiap for detecting English spoken terms employs Large Vocabulary Continuous Speech Recognition (LVCSR) system developed within the AMI(DA) project. In 2009, the STD system was further improved by applying a language identification module to avoid processing speech segments pronounced in a non-English language.

3. Image and video processing and analysis

Leading researchers: Daniel Gatica-Perez, Jean-Marc Odobez, François Fleuret, Barbara Caputo

Idiap keeps also being very active, and a recognized leader, in multiple areas of image and video processing and analysis, including: object modelling, algorithm robustness, data fusion (colour, shape, motion) and feature selection, online learning and model adaptation, multi-object tracking (dynamics and data-likelihood modelling), behavioural models, joint tracking and event recognition, computational complexity. In 2009 we laid the foundations for initiating new activities in the area of "autonomous cognitive systems". The PASCAL funded project SS2-Rob (Section 6), granted in 2009 and started in January 2010, aims at designing categorization algorithms for life-long learning of semantic place representation.

Many of the resulting algorithms are made publicly available through the TorchVision library (as part of Torch, http://torch3vision.idiap.ch) also allowing researchers to incrementally build on each other, while optimizing opportunities for collaborations. Some of the 2008 key achievements in image and video processing and analysis include the automatic video monitoring in public spaces, and the pose detection of complex objects in cluttered scenes.

4. Multimodal information management and indexing

Leading researcher: Andrei Popescu-Belis, Mike Flynn

Idiap keeps being active in the areas of content-based information management using multiple data (audio and video) streams, and optimization of user interaction. In 2009, this research activity mainly focused on augmenting the access to meeting archives, both through the Automatic Content Linking Device, described in Section 5.9, and through the design of a question answering prototype system over meeting transcriptions. This system had the goal of passing automatically the BET meeting browser test previously developed at Idiap, in other words of distinguishing automatically between true and false statements referring to important events in a meeting. The system was proven to be quite successful, especially at identifying relevant portions of meetings, even compared to human subjects, and was robust on automatic transcripts compared to reference ones.

5. Biometric person recognition

Leading Researcher: Sébastien Marcel

Idiap keeps working on increasing robustness of person recognition techniques, non-frontal face verification, multimodal fusion, multimodal user authentication (mixture of experts, confidence-based weighting of the different media, etc). In 2009, these efforts were taking place mainly in the context of one SNSF project, as well as the FP7 MOBIO project (with Idiap as coordinator). The resulting research efforts are currently at the basis of several developments and technology transfer success, including one of Idiap's spin-off, KeyLemon (http://www.keylemon.com), enabling users to automatically lock/unlock their laptop based on their facial prints.

In 2009, Idiap established strong relationships with new research institutions and with several companies, such as SAGEM Sécurité, during the preparation of a joint collaborative European project.. This project, positively reviewed and currently under negotiation, should study the vulnerability of biometric systems to attacks at the sensor level, so-called spoofing attacks, and will develop counter-measures. Hence, we expect to start in 2011 a new research direction within the biometric person recognition research theme: Trusted Biometrics under Spoofing Attacks.

6. Social and human behaviour

Leading researchers: Daniel Gatica-Perez, Jean-Marc Odobez

This area is concerned with the automatic analysis of a variety of real-life human behaviours. This activity exploits expertise and synergies between key areas at Idiap including multi-sensor human behaviour capture, machine learning, and perceptual processing. In 2009, the specific work in this domain included three main areas.

In the first place, we extended our work on unsupervised learning of daily routines at large-scale from mobile phone users, which operate on low-level observations obtained from phone sensor data, such as the locations of an individual and who they are in proximity with. We are continuing our collaborative work with Nokia Research Centre, Lausanne, which involves the collection and analysis of real-life data from 150 mobile phone users.

In the second place, Idiap's long-term work on face-to-face interaction modelling in small group meetings, which has produced new methods to model visual attention, turn-taking, and higher-level constructs like dominance (conducted thanks to the European project AMIDA or the NCCR IM2) will be extended thanks to the newly granted HUMAVIPS European project (starting in 2010, see Section 6.2) which seeks to endow robots with simple social skills necessary to deal with small groups of people, including the basic reasoning to pick out a group of humans and determine which ones want to interact with it, or focus attention (and computational resources) on just one person in the midst of other people, voices and background noise, and attracts his attention.

In the third place, there is a growing interest in capturing personal audio logs (multiparty conversations, meetings etc.) through portable devices and analyzing them to model real-life social interactions. One of the main issues here is that recording and storing "raw" audio data can breach privacy of people whose consent has not been explicitly obtained. Along this direction, a PhD student (as part of SNSF project MULTI) is conducting research investigating, extracting, and exploring audio features that are "privacy preserving", i.e., neither reconstruction of intelligible speech nor recognition of lexical content should be possible from these features. However, these features should still carry enough information to perform tasks such as, speech/non-speech detection, speaker diarization, audio scene detection, conversation analysis which in turn aids in the modelling of social interaction.

7. Social signal processing (SSP)

Leading researcher: Alessandro Vinciarelli, Hervé Bourlard

In 2009, Social Signal Processing (SSP) scientific activities have addressed both methodological aspects and experimental works.

On the methodological plan, SSP activities have involved 1) the development of Conditional Random Fields where the relationship between observations and labels is expressed via mixture models (in collaboration with the Institute for Creative Technologies) and 2) the investigation of variability and predictability in behavior as a source of socially relevant perceptions.

On the experimental plan, 1) previously developed approaches for role recognition have been improved by explicitly taking into account conversational dynamics, and 2) variability and predictability measures have been applied to infer socially relevant information from nonverbal voice characteristics.

In parallel with the scientific work, SSP activities have included the coordination of the SSPNet (the European Network of Excellence aimed at establishing a research community in SSP) and the organization of several SSP oriented initiatives, including the First International Workshop on Social Signal Processing, the First International Workshop on Foundations of Social Signals, and the Doctoral Consortium of the International Conference on Affective Computing and Intelligent Interaction. Furthermore, a new SSP research program has been established in IM2 with the goal of exploring, in collaboration with ETHZ and the Affective-Sciences NCCR, effectiveness of delivery in public speech and engagement in social interactions.

8. Multilingual processing of spoken and written information

Leading researchers: Andrei Popescu-Belis, John Dines, Phil Garner, Hervé Bourlard

Multilingual processing is becoming a key research theme in Europe, while being vastly underrepresented in a multilingual country like Switzerland. Building upon their expertise and activities in spoken language processing (and their know-how in multilingual spoken and written information processing), we still believe that Idiap is in a unique position to develop large activities towards multilingual speech-based document retrieval and machine translation technology.

Based on internal activities, and exploiting our internal research network in this area, we could already get an FP7 European project covering some of these aspects: Effective

Multilingual Interaction in Mobile Environments (EMIME, http://www.emime.org). The EMIME project will help to overcome the language barrier by developing a mobile device that performs personalised speech-to-speech translation, such that the a user's spoken input in one language is used to produce spoken output in another language, while continuing to sound like the user's voice. Personalisation of systems for cross-lingual spoken communication is an important, but little explored, topic. It is essential for providing more natural interaction and making the computing device a less obtrusive element when assisting human-human interactions.

In 2009, and in the context of the EMIME project, we have developed a capability in HMM based speech synthesis, in particular in the emerging field of cross-lingual speaker adaptation. Coupled with translation (COMTIS project, see below), this has the persuasive effect of allowing a speaker to appear to speak a different language using his/her own voice. It is worth noting here that this year, both Google and Microsoft have announced plans to develop personalised speech-to-speech translation services signalling increasing interest in this research domain and placing our research at the forefront of developments.

We have also started, in 2009, pilot research on the translation of written language, focussing on a problem that is less targeted in the current statistical machine translation (MT) paradigm: the translation of relationships between sentences. In this direction, we have submitted and obtained an SNSF Sinergia project with two teams in Geneva, called COMTIS (see n. 5 in section 6.2), and submitted a European project which was above threshold but was not funded (this project will be improved and resubmitted when the next call including MT is issued). In the frame of COMTIS, we have mainly focused on the collection of examples of various types of dependencies between sentences (such as pronouns, verb tense) that are problematic to current MT engines, in preparation for their handling within COMTIS.

To further boost our activities in multi-lingual processing, and as part of the SNSF MULTI project, a new PhD student (David Imseng) also started investing different approaches towards fast development/adaption of multi-lingual speech recognition systems, with a particular emphasis on the Swiss national languages.

9. Innovation process and creativity analysis and modelling

Leading researcher: Hervé Bourlard

Although not planned in the initial Joint Development Plan, Idiap is aiming at exploiting and extending his multi-disciplinary know-how described above, towards another new area, where many institutions are currently investing quite a lot of efforts (in infrastructure, new centres, and manpower) and where we believe we could play a key role based on our current activities. This new area concerns the analysis and enhancement (in several directions, including 3D internet collaboration) of innovation/creativity processes through the development of advanced software tools exploiting state-of-the-art knowledge in interactive multimodal information management (IM2), human-to-human communication modelling (AMI/AMIDA), social signal processing (EU SSPNet project), in addition to additional knowledge in data visualization and rendering, cognitive ergonomics, design and art, and social psychology.

1. Realtime large vocabulary speech recognition (Juicer)

Leading researchers: Phil Garner, John Dines

In 2009, Idiap (in collaboration with the universities of Edinburgh, Sheffield and Brno) continued to develop their real-time capable speaker independent large vocabulary speech recognition system. Considerable speed-ups and new capabilities were added, and the system was tested in earnest in the spring of 2009 in the context of an international evaluation run by NIST. Using Juicer, the (AMIDA) team was able to reduce several passes of the previous system into a single pass, in turn allowing more optimisation. The team were ranked first place in the MDM (microphone array based) section. During 2009, the system was further enhanced to enable the system to be run in a real time demonstration context.

2. Speech processing & MLP-based feature extraction

Leading researchers: Fabio Valente, Mathew Magimai.Doss, Hervé Bourlard

In 2009, Idiap contributed significantly to the last year of the US DARPA GALE (Global Autonomous Language Exploitation) project founded by DARPA as part of the Nightingale team lead by SRI international (USA) with partners in ICSI (Berkeley, USA) and RWTH University (Aachen Germany). MLP features trained on 1500 hours of labelled Mandarin speech data (i.e. approximately 600 million speech vectors) have been evaluated in the full multipath RWTH GALE system. Experiments on 4 dataset from GALE evaluation have revealed that Idiap MLP features produced a relative error reduction in the range of 10-15% in all the dataset and they outperformed other MLP front-ends used in the project and evaluated in the same framework. Whenever used in combination with ICSI MLP features, they produced an impressive 20% relative error reduction.

In the context of modelling social interaction (Section 4.6), privacy preserving features, such as, short-term energy, zero crossing rate, kurtosis, spectral slope, cepstral coefficients extracted from linear prediction residual signal was investigated for speech/non-speech detection on multiparty conversation data, and were studied in comparison with "non privacy preserving" state-of-the-art short-term spectral-based features. The studies show that the privacy preserving features when modelled together can yield performance closer to the short-term spectral based features.

Finally, as one of the main conclusions of a finalized PhD thesis (Joel Pinto, defended early 2010), Idiap proposed a hierarchical MLP-based approach to model information present in long temporal context (about 150-250 ms) of phoneme posterior probability estimates for automatic speech recognition. The proposed approach was extensively studied for phoneme recognition, task adaptation, large vocabulary Mandarin automatic speech recognition, and genre adaption. In all the cases, the proposed approach was found to yield better systems than the standard single MLP-based approach.

3. Speaker diarization

Leader researcher: Fabio Valente

Speaker diarization refers to the problem of automatically detecting "who spoke when". Speaker diariziation thus involves determining the number of speakers in a given audio stream and clustering together segments belonging to the same speakers. Typical baseline diarization systems use a combination of acoustic and location features. Idiap has proposed a non-parametric system based on the information bottleneck principle that allows the use of an arbitrarily large number of features and has reported, for the first time, a considerable error reduction (30% relative) over the baseline on several dataset from the diarization international evaluation organized by NIST. Furthermore the proposed system only marginally increases the computational complexity in case of multiple feature streams, resulting in faster-than-real-time diarization.

4. Visual focus of attention

Leading researchers: Jean-Marc Odobez, Daniel Gatica-Perez

In 2009, and as part of its efforts towards automatic analysis and understanding of the human-to-human communication patterns, Idiap has pursued its research on the recognition of gaze (also called Visual Focus of Attention - VFOA) in meetings from head pose information. The joint modelling of people's VFOA, speaker turns, conversational structure, and meeting context (slide presentation) has been extended in three directions. In the first one we studied the role of visual gesturing as an indicator of speaking activity and focus attraction, and showed on 12 hours of data that such visual attractor could be as effective as the speaking status at improving the recognition of the VFOA. Reversely, in the second direction we showed that the VFOA information

(derived from visual information only) could indeed be successfully exploited in a speaker diarization task, as described earlier. In the third direction, we adapted the system to handle gazing at moving targets (typically, people standing to present) and dynamically associate gaze directions with target location. In addition, Idiap has been working on the implementation of a simplified yet robust system for real-time VFOA identification for application in larger system, by optimizing its head pose estimation algorithm. In collaboration with Twente University, this module has been successfully integrated within an application that favours the engagement of participants with a distant access to a meeting, by providing them a better understanding of the social interactions going on in the main meeting room. In the future, this line of research will be continued in the context of the HUMAVIPS project to endow a humanoid robot with capabilities to naturally follow the conversation, monitor people attention, and understand whether it is addressed by the current speaker.

5. Activity monitoring in public spaces

Leading researcher: Jean-Marc Odobez

In 2009, Idiap has developed new approaches towards automatic activity analysis from outdoor or indoor video recordings in metro stations, more specifically: (1) real-time detection of abandoned luggage, and (2) person density monitoring and (3) unsupervised activity discovery from long term recordings.

In the context of the EU CARETAKER project, Idiap has developed two state-of-the-art modules to monitor the activities of passengers in public transportation systems from surveillance cameras. The first one detects abandoned luggage by passengers, and relies on some analysis of the different layers of static and moving objects in the scene, as well as on the detection of potential people remaining close to these objects. The second module detects and counts the number of people present in different camera views, and allows to build statistical models of the metro overall activity, as well as to spot anomalous situation in function of the context (location, week day, time) like exceptional congestions due to overloaded traffic or the presence of a group of loitering people. Integrated within a large-scale system, which was deployed in the Roma and Torino metros, these modules were evaluated by real operators and revealed their high reliability. In particular, the abandoned luggage detector raised the interest of the Roma metro managers, as it proved to be better than commercial systems that they had previously tested, providing much fewer false alarms while still detecting relevant events.

More recently, confronted to the proliferation of video surveillance systems generating large amounts of data, we have investigated novel unsupervised data mining tools that can identify the relevant information they contain from long recordings. Discovered activities are interesting in their own, by revealing proper activities that would have been difficult to define a priori, or can be useful for higher level reasoning. Our approach relies on bag representation of low level features related to location, objects (car, people) and motion, and probabilistic topic models, and allowed to automatically discover the scene structure and object related activities (e.g. car tracks, sidewalks, waiting places, zebra crossing) and identify anomalies (e.g., car stopping at an unusual place, or pedestrian crossing off the zebra).

6. Large-scale human behaviour modelling from mobile phones

Leading researcher: Daniel Gatica-Perez

We have developed methods for unsupervised learning of daily routines at large-scale from mobile phone users, which operate on low-level observations obtained from phone sensor data, such as the locations of an individual and who they are in proximity with, as well as the time of the day when this occurs. Our most recent work addresses the problem of discovering joint routines of location and proximity human routine discovery via probabilistic topic model. We also use methodology to predicting missing multi-modal data. We propose a bag of multi-modal local behaviours that integrates the modelling of variations of location over multiple time-scales, and the modelling of interaction types from proximity. Our representation is robust to characterize real-life human behaviour sensed from mobile phones, which produced data known to be noisy or incomplete. Our methodology is able to discover latent human activities in terms of the joint interaction and location behaviours of 97 individuals over the course of approximately a 10 month period using a massive data set. Using location-based semantic categories (home, work, etc.) and proximity categories that relate to the size of the group a person is physically close to, the human activities discovered with our multi-modal framework involve joint patterns of places and proximity to others at different times, like "going out alone from 7pm-midnight alone" or "working from 11am-5pm with a small group". We have further demonstrated the feasibility of our framework to predict missing location data as well as proximity data. This work is currently extended through new research with Nokia.

7 Idiap@ImageCLEF2009

Leading researcher: Barbara Caputo

Leveraging on the success achieved in 2007 and 2008 in the Medical Image Annotation track, in 2009 Idiap has been invited to join the organizing committee of the ImageCLEF benchmark evaluation campaign. Given the growing impact that publicly available databases and benchmark evaluations have on the scientific community, this represents an exciting opportunity to influence research in selected areas. Specifically, we organized the Medical Image Annotation track, building on our own experience of winners of the last two years, and we started a new track on Robot Vision, leveraging on our experience on autonomous cognitive systems. The Medical image Annotation track was organized for the last time, and consisted of a survey of the previous four years experience. Participants were asked to develop and test their methods on a collection of more than 12,000 fully classified radiographs, labelled according to various levels of semantic. We received a total of 18 submissions, and Idiap ranked second, using the algorithm that did win the 2008 evaluation task. This confirms us in the robustness of our method.

The robot vision task was organized for the first time and attracted a considerable attention. We addressed the problem of topological localization of a mobile robot using visual information. Specifically, participants were asked to determine the topological location of a robot based on images acquired with a perspective camera mounted on a robot platform. A total of 27 runs were submitting, which testify the great interest this task elicited in the community.

8 Realtime multimodal annotation distribution and storage (The "Hub")

Leading researcher: Mike Flynn

The Hub real-time annotation distribution and storage mechanism enables communication between various recognition systems (e.g. automatic speech recognition) and browsers or other processing systems (such as summarizers). The Hub has been extended, during 2009, with a mechanism that facilitates data modelling, called the "Easy Hub" (ezHub). This mechanism allows a principled definition of the entities and relationships that transit to the Hub, and which describe a given type of multimodal data and its processing stages (annotations). The ezHub was instantiated with meeting-related data structures, enabling the Hub to represent all annotations and metadata included in the AMI Meeting Corpus. In addition, we have connected the JFerret framework for designing meeting browsers to the Hub architecture, thus increasing even further the integration of our tools.

9 Automatic content linking device

Leading researcher: Andrei Popescu-Belis

In 2009, within the AMIDA and IM2 projects, Idiap continued the coordination of the Automatic Content Linking Device (ACLD), which was developed into a second, richer, and more flexible version. The ACLD provides the participants to an ongoing discussion (or users listening to a recording discussion or lecture) with documents and fragments of past recorded meetings that are related to the content of this discussion, as well as with websites and potentially any file from the user's computer. The ACLD makes use of Idiap's Large Vocabulary Continuous Speech Recognition system (see above) and multimedia archive contains past meetings recorded in smart meeting rooms and processed using automatic speech recognition, speech segmentation, speaker diarization, and other content abstraction modules. The ACLD makes use of "The Hub" architecture (see above) to integrate its modules. In 2009, a new interface was designed by the ACLD partners, with widgets displaying various aspects of processing (ASR, tag cloud of keyword, and resulting documents and websites). The system was applied to the AMI Meeting Corpus, but also to lectures in computer science. Real-time versions were demonstrated and tested in various conditions. A major achievement was the porting of the ACLD system, together with the Idiap ASR system, on a single Macintosh laptop, which now makes the demonstrator portable: a typical scenario for demonstration has the speaker use the ACLD while describing the system. A version installed in the Idiap Show Room is connected to the database of Idiap publications, and is intended as an assistant for a presenter talking about research at Idiap.

MAIN PROJECTS IN PROGRESS

ACRONYM NAME, NAME	PARTNERS

EUROPEAN PROJECTS

AMIDA
Augmented Multiparty Interaction with
Distance Access

German Research Centre for Artificial Intelligence (DFKI)
International Computer Science Institute (ICSI)
Netherlands Organisation for Applied Scientific Research (TNO)
CSIRO ICT Centre
University of Edinburgh (UEDIN)
Sheffield University (USFD)
Brno University of Technology (BUT),
Munich University of Technology (TUM)
University of Twente (UT)
Philips Consumer Electronics

DIRAC
Detection and Identification
of Rare Audio-visual Cues

Eidgenössische Technische Hochschule Zürich (ETHZ)
The Hebrew University of Jerusalem
Czech Technical University
Carl Von Ossietzky Universität Oldenburg
Leibniz Institute for Neurobiology
Katholieke Universiteit Leuven
Oregon Health and Science University Ogi School of Science and Engineering
University of Maryland, Neural Systems Laboratory

EMIME
Effective Multilingual Interaction
in Mobile Environments

University Edinburgh
Helsinki University of Technology
Nagoya Institute of Technology
Nokia Corporation
University of Cambridge

MOBIO
Mobile Biometry

University of Manchester
University of Surrey
Laboratoire d'Informatique d'Avignon
Brno University of Technology
University of Oulu
EyePmedia
IdeArk

PASCAL2
Pattern Analysis, Statistical Modelling
and Computational Learning

56 sites in the network

SCALE
Speech Communication
with Adaptive Learning

Universität des Saarlandes
University of Edinburgh
University of Sheffield
Radboud University Nijmegen
RWTH Aachen
Motorola Limited UK
Philips
Eurice

DURATION (MONTH/YEAR)	WEB	COORDINATOR	CONTACT
01.06 – 02.10	www.amiproject.org	Idiap Research Institute	Prof. Hervé Bourlard
01.06 – 12.10	www.diracproject.org	Idiap Research Institute	Dr Barbara Caputo
03.08 – 02.11	www.emime.org	University of Edinburgh	Dr John Dines
02.08 – 01.11	www.mobioproject.org	Idiap Research Institute	Dr Sébastien Marcel
03.08 – 02.13	www.pascal-network.org	University of Southampton	Dr François Fleuret
10.08 – 09.12	www.scale.uni-saarland.de	Universität des Saarlandes	Prof. Hervé Bourlard

ACRONYM NAME, NAME	PARTNERS
SSPnet Social Signal Processing Network	Imperial College of Science, Technology and Medicine University of Edinburgh University of Twente Università Di Roma Tre Queen's University Belfast DFKI CNRS Université de Genève Technische Universiteit Delft
TA2 Together Anywhere, Together anytime	EURESCOM - European Institute for Research and Strategic Studies in Telecommunications GmbH, British Telecommunications plc Alcatel-Lucent Bell NV Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. Goldsmiths' College Netherlands Organisation For Applied Scientific Research – TNO The Interactive Institute II Aktiebolag Stichting Centrum voor Wiskunde en Informatica, Ravensburger Spieleverlag GmbH Philips Consumer Electronics BV Limbic Entertainment GmbH Joanneum Research Forschungsgesellschaft GmbH

DURATION (MONTH/YEAR)	WEB	COORDINATOR	CONTACT
02.09 – 01.14	www.sspnet.eu	Idiap Research Institute	Dr Alessandro Vinciarelli
02.08 – 01.12	www.ta2-project.eu	Eurescom	Phil Garner

ACRONYM NAME, NAME	PARTNERS

SNSF PROJECTS

NCCR IM2
Interactive Multimodal Information
Management

Ecole polytechnique fédérale de Lausanne (EPFL)
University of Geneva
University of Fribourg
University of Bern
Swiss Federal Institute of Technology in Zurich (ETHZ)

CODICES
Automatic Analysis of Mexican
Codex Collections

GMface I + II
Graphical Models for Face Authentiction

FlexASR
Flexible Grapheme-Based
Automatic Speech Recognition

UBM
Understanding Brain Morphogenesis

Ecole Polytechnique Fédérale de Lausanne (EPFL)
University of Basel

VELASH
Very Large Sets of Heuristics
for Scene Interpretation

MULTI08
Multimodal Interaction and
Multimedia Data Mining

SNSF PROJECTS (INDO–SUISSE)

KEYSPOT
Keyword Spotting in Continuous Speech

KERSEQ
Kernel Methods for Speech and
Video Sequence Analysis

HASLER FOUNDATION

EMMA
Enhanced Medical Multimedia data Access

DURATION (MONTH/YEAR)	WEB	COORDINATOR	CONTACT
01.02 – 12.09	www.im2.ch	Idiap Research Institute	Prof. Hervé Bourlard
09.08 – 08.11	www.idiap.ch/~eroman/codices.html	Idiap Research Institute	Dr Daniel Gatica-Perez Dr Jean-Marc Odobez
07.07 – 06.09		Idiap Research Institute	Dr Sébastien Marcel
06.09 – 05.12		Idiap Research Institute	Dr Mathew Magimai Doss
10.09 – 09.12		University of Basel	Dr François Fleuret
06.09 – 05.12		Idiap Research Institute	Dr François Fleuret
10.08 – 09.10		Idiap Research Institute	Prof. Hervé Bourlard
10.06 – 09.09		Idiap Research Institute	Prof. Hervé Bourlard
06.06 – 05.09		Idiap Research Institute	Prof. Hervé Bourlard
01.08 – 12.09		Idiap Research Institute	Dr Barbara Caputo

MAJOR PUBLICATIONS / CONFERENCES

This selection, from among the many publications of Idiap illustrates the diversity of our research.

BOOKS, BOOKS CHAPTERS AND JOURNAL PAPERS

Multimodal Signal Processing: Methods and Techniques to Build Multimodal Interactive Systems
Jean-Philippe Thiran, Hervé Bourlard and Ferran Marques
Academic Press, 2009

A Kernel Wrapper for Phoneme Sequence Recognition
Joseph Keshet and Dan Chazan
Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods, John Wiley and Sons, 2009

A Large Margin Algorithm for Forced Alignment
Joseph Keshet, Shai Shalev-Shwartz, Yoram Singer and Dan Chazan
Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods, John Wiley and Sons, 2009

A Proposal for a Kernel-based Algorithm for Large Vocabulary Continuous Speech Recognition
Joseph Keshet
Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods, John Wiley and Sons, 2009

Accessing a Large Multimodal Corpus using an Automatic Content Linking Device
Andrei Popescu-Belis, Jean Carletta, Jonathan Kilgour and Peter Poller
Multimodal Corpora: From Models of Natural Interaction to Systems and Applications, pages 189-206, Springer-Verlag, 2009

Discriminative Keyword Spotting
David Grangier, Joseph Keshet and Samy Bengio
Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods, John Wiley and Sons, 2009

Managing Multimodal Data, Metadata and Annotations: Challenges and Solutions
Andrei Popescu-Belis
Multimodal Signal Processing for Human-Computer Interaction, pages 183-203, Elsevier / Academic Press, 2009

A Novel Criterion for Classifiers Combination in Multistream Speech Recognition
Fabio Valente
IEEE Signal Processing Letters, 16(7):561-564, 2009

A novel statistical generative model dedicated to face recognition
Guillaume Heusch and Sébastien Marcel
Image & Vision Computing, 2009

An Information Theoretic Approach to Speaker Diarization of Meeting Data
Deepu Vijayasenan, Fabio Valente and Hervé Bourlard
IEEE Transactions on Audio Speech and Language Processing, 17(7):1382-1393, 2009

Automatic Role Recognition in Multiparty Recordings: Using Social Affiliation Networks for Feature Extraction
Hugues Salamin, Sarah Favre and Alessandro Vinciarelli
IEEE Transactions on Multimedia, 11(7):1373-1380, 2009

Beamforming with a Maximum Negentropy Criterion
Kenichi Kumatani, John McDonough, Barbara Rauch, Dietrich Klakow, Philip N. Garner and Weifeng Li
IEEE Transactions on Audio Speech and Language Processing, 17(5):994-1008, 2009

Bounded kernel-based perceptrons
Francesco Orabona, Joseph Keshet and Barbara Caputo
Journal of Machine Learning Research, Accepted for pub, 2009

Capturing Order in Social Interactions
Alessandro Vinciarelli
IEEE Signal Processing Magazine, 2009

Classifying Material in the Real World
Barbara Caputo, Eric Hayman, M Fritz and J-O Eklundh
Image and vision Computing, accepted for pub, 2009

COLD: The COsy Localization Database
Andrzej Pronobis and Barbara Caputo
International Journal of Robotics Research, 28(5):588-594, 2009

Contextual classification of image patches with latent aspect models
Florent Monay, Pedro Quelhas, Jean-Marc Odobez and Daniel Gatica-Perez
EURASIP Journal on Image and Video Processing, Special Issue on Patches in Vision, 2009

Discriminative Keyword Spotting
Joseph Keshet, David Grangier and Samy Bengio
Speech Communication, 51(4):317-329, 2009

Implicit Human Centered Tagging
Maja Pantic and Alessandro Vinciarelli
IEEE Signal Processing Magazine, 26, 2009

Multi-layer Boosting for Pattern Recognition
Francois Fleuret
Pattern Recognition Letter, 30:237-241, 2009

On the vulnerability of face verification systems to hill-climbing attacks
Javier Galbally, Chris McCool, Julian Fierrez, Sébastien Marcel and Javier Ortega-Garcia
Pattern Recognition, 2009

Recognizing Human Visual Focus of Attention from Head Pose in Meetings
Silèye O. Ba and Jean-Marc Odobez
IEEE Transactions on Systems, Man, Cybernetics, Part-B, Vol. 39(No. 1):16-34, 2009

Social Signal Processing: Survey of an Emerging Domain
Alessandro Vinciarelli, Maja Pantic and Hervé Bourlard
Image and Vision Computing, 2009

The FEMTI guidelines for contextual MT evaluation: principles and tools
Paula Estrella, Andrei Popescu-Belis and Margaret King
Linguistica Antverpiensia New Series, 8, 2009

Towards Life-long Learning for Cognitive Systems: Online Independent Support Vector Machine
Francesco Orabona, Claudio Castellini, Barbara Caputo, Jie Luo and Giulio Sandini
Pattern Recognition, Accepted for Pub, 2009

CONFERENCE PAPERS

A Multimedia Retrieval System Using Speech Input
Andrei Popescu-Belis, Peter Poller, Jonathan Kilgour, Erik Boertjes, Jean Carletta, Sandro Castronovo, Michal Fapso, Alexandre Nanchen, Theresa Wilson, Joost de Wit and Majid Yazdani
Proceedings of ICMI-MLMI 2009 (11th International Conference on Multimodal Interfaces and 6th Workshop on Machine Learning for Multimodal Interaction), Cambridge, MA, 2009

A theoretical framework for transfer of knowledge across modalities in artificial and cognitive systems
Francesco Orabona, Barbara Caputo, Antje Fillbrandt and Frank Ohl
International Conference on Developmental Learning, 2009

An online framework for learning novel concepts over multiple cues
Jie Luo, Francesco Orabona and Barbara Caputo
Proceeding of The 9th Asian Conference on Computer Vision, Xi'an, China, 2009

Applications of signal analysis using autorégressive models for amplitude modulation
Sriram Ganapathy, Samuel Thomas, Petr Motlicek and Hynek Hermansky
IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2009, WASPAA '09., IEEE, Mohonk Mountain House, New Paltz, New York, USA, pages 341-344, 2009

Arithmetic Coding of Sub-Band Residuals in FDLP Speech/Audio Codec
Petr Motlicek, Sriram Ganapathy and Hynek Hermansky
10th Annual Conference of the International Speech Communication Association, ISCA, Brighton, England, pages 2591-2594, ISCA 2009, 2009

Automatic Out-of-Language Detection Based on Confidence Measures Derived fromLVCSR Word and Phone Lattices
Petr Motlicek
10thAnnual Conference of the International Speech Communication Association, ISCA, Brighton, England, pages 1215-1218, 2009

Automatic Role Recognition in Multiparty Recordings Using Social Networks and Probabilistic Sequential Models
Sarah Favre, Alfred Dielmann and Alessandro Vinciarelli
ACM International Conference on Multimedia, 2009

Automatic vs. human question answering over multimedia meeting recordings
Quoc Anh Le and Andrei Popescu-Belis
10th Annual Conference of the International Speech Communication Association, 2009

Bayesian Networks to Combine Intensity and Color Information in Face Recognition
Guillaume Heusch and Sébastien Marcel
International Conference on Biometrics, pages 414-423, Springer, 2009

Canal9: A database of political debates for analysis of social interactions
Alessandro Vinciarelli, Alfred Dielmann, Sarah Favre and Hugues Salamin
Proceedings of the International Conference on Affective Computing and Intelligent Interaction (IEEE International Workshop on Social Signal Processing), Amsterdam, Netherlands, pages 1-4, 2009

Characterising Conversationsal Group Dynamics Using Nonverbal Behaviour
Dinesh Babu Jayagopi, Raducanu Bogdan and Daniel Gatica-Perez
Proceedings ICME 2009, 2009

Discovering Group Nonverbal Conversational Patterns with Topics
Dinesh Babu Jayagopi and Daniel Gatica-Perez
Proceedings ICMI-MLMI, 2009

Dynamic Partitioned Sampling For Tracking With Discriminative Features
Stefan Duffner, Jean-Marc Odobez and Elisa Ricci
Proceedings of the British Maschine Vision Conference, London, 2009

Error Resilient Speech Coding Using Sub-band Hilbert Envelopes
Sriram Ganapathy, Petr Motlicek and Hynek Hermansky,
12th International Conference on Text, Speech and Dialogue, TSD 2009, Pilsen, Czech Republic, Springer - Verlag, Berlin Heidelberg 2009, 2009

Error Resilient Speech Coding Using Sub-band Hilbert Envelopes
Sriram Ganapathy, Petr Motlicek and Hynek Hermansky
12th International Conference on Text, Speech and Dialogue, TSD 2009, Pilsen, Czech Republic, pages 355-362, Springer - Verlag, Berlin Heidelberg 2009, 2009

Evaluation of Probabilistic Occupancy Map People Detection for Surveillance Systems
Jerome Berclaz, Ali Shahrokni, Francois Fleuret, James Ferryman and Pascal Fua
Proceedings of the IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, 2009

Flickr Hypergroups

Radu-Andrei Negoescu, Brett Adams, Dinh Phung, Svetha Venkatesh and Daniel Gatica-Perez
Proceedings of the 17th ACM International Conference on Multimedia, 2009

Haar Local Binary Pattern Feature for Fast Illumination Invariant Face Detection

Anindya Roy and Sébastien Marcel
British Machine Vision Conference 2009, 2009

Hierarchical Processing of the Modulation Spectrum for GALE Mandarin LVCSR system

Fabio Valente, Mathew Magimai.-Doss, Christian Plahl and Ravuri Suman
Proceedings of the 10thAnnual Conference of the International Speech Communication Association (Interspeech), Brighton, 2009

Hill-Climbing Attack to an Eigenface-Based Face Verification System

Javier Galbally, Chris McCool, Julian Fierrez, Sébastien Marcel and Javier Ortega-Garcia
Proceedings of the First IEEE International Conference on Biometrics, Identity and Security (BIdS), 2009

Implicit Human Centered Tagging

Alessandro Vinciarelli, Nicolae Suditu and Maja Pantic
Proceedings of IEEE Conference on Multimedia and Expo, pages 1428-1431, 2009

Investigating Privacy-Sensitive Features for Speech Detection in Multiparty Conversations

Sree Hari Krishnan Parthasarathi, Mathew Magimai.-Doss, Hervé Bourlard and Daniel Gatica-Perez
Proceedings of Interspeech 2009, 2009

Investigating the use of Visual Focus of Attention for Audio-Visual Speaker Diarisation

Giulia Garau, Silèye O. Ba, Hervé Bourlard and Jean-Marc Odobez
Proceedings of the ACM International Conference on Multimedia, Beijing, China, 2009

Joint Learning of Pose Estimators and Features for Object Detection

Karim Ali, Francois Fleuret, David Hasler and Pascal Fua
Proceedings of the IEEE International Conference on Computer Vision, 2009

KL Realignment for Speaker Diarization with Multiple Feature Streams

Deepu Vijayasenan, Fabio Valente and Hervé Bourlard
10th Annual Conference of the International Speech Communication Association, 2009

Learning and Predicting Multimodal Daily Life Patterns from Cell Phones

Katayoun Farrahi and Daniel Gatica-Perez
ICMI-MLMI, 2009

Learning Large Margin Likelihood for Realtime Head Pose Tracking

Elisa Ricci and Jean-Marc Odobez
IEEE Int. Conference on Image Processing, Cairo, Egypt, IEEE, 2009

Learning Rotational Features for Filament Detection

German Gonzalez, Francois Fleuret and Pascal Fua
Proceedings of the IEEE international conference on Computer Vision and Pattern Recognition, 2009

MDCT for Encoding Residual Signals in Frequency Domain Linear Prediction

Sriram Ganapathy, Petr Motlicek and Hynek Hermansky
Audio Engineering Society (AES), 127th Convention, Audio Engineering Society (AES), Audio Engineering Society, 60 East 42nd Street, New York, New York 10165-2520, USA;, 2009

Measuring the gap between HMM-based ASR and TTS

John Dines, Junichi Yamagishi and Simon King
Proceedings of Interspeech, Brighton, U.K., 2009

Memoirs of Togetherness from Audio Logs

Danil Korchagin
Proceedings International ICST Conference on User Centric Media, Venice, Italy, 2009

MLP Based Hierarchical System for Task Adaptation in ASR

Joel Praveen Pinto, Mathew Magimai.-Doss and Hervé Bourlard
Proceedings of the IEEE workshop on Automatic Speech Recognition and Understanding, Merano, Italy, pages 365-370, 2009

Model adaptation with least-square SVM for adaptive hand prosthetics

Francesco Orabona, Claudio Castellini, Barbara Caputo, Angelo Emanuele Fiorilla and Giulio Sandini
IEEE International conference on Robotics and Automation, 2009

Multi-modal speaker diarization of real-world meetings using compresed-domain vidéo features

Gerald Friedland, Hayley Hung and Chuohao Yeo
International Conference on Audio, Speech and Signal Processing, 2009

Mutual information based channel sélection for speaker diarization of meetings data

Deepu Vijayasenan, Fabio Valente and Hervé Bourlard
Proceedings of International Conference on Acoustics, Speech and Signal Processing, 2009

Mutual Information based Channel Selection for Speaker Diarization of Meetings Data

Deepu Vijayasenan, Fabio Valente and Hervé Bourlard
Proceedings of International conference on acoustics speech and signal processing, 2009

Non-linear mapping for multi-channel speech separation and robust overlapping speech recognition

Weifeng Li, John Dines, Mathew Magimai.-Doss and Hervé Bourlard
Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2009

Out-of-Scene AV Data Detection

Danil Korchagin
Proceedings IADIS International Conference Applied Computing, Rome, Italy, pages 244-248, 2009

Parts-Based Face Verification using Local Frequency Bands

Chris McCool and Sébastien Marcel,
in Mcceedings of IEEE/IAPR International Conference on Biometrics, 2009

Posterior features applied to speech recognition tasks with user-defined vocabulary

Guillermo Aradilla, Hervé Bourlard and Mathew Magimai.-Doss
Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2009

Real-Time ASR from Meetings

Philip N. Garner, John Dines, Thomas Hain, Asmaa El Hannani, Martin Karafiat, Danil Korchagin, Mike Lincoln, Vincent Wan and Le Zhang
Proceedings of Interspeech, Brighton, UK., 2009

Retrieving Ancient Maya Glyphs with Shape Context

Edgar Roman-Rangel, Carlos Pallan, Jean-Marc Odobez and Daniel Gatica-Perez
2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, Kyoto, Japan, IEEE, 2009

Robust Discriminative Keyword Spotting for Emotionally Colored Spontaneous Speech using Bidirectional LSTM Networks

Martin Wöllmer, Florian Eyben, Joseph Keshet, Alex Graves, Björn Schuller and Gerhard Rigoll
IEEE International Conference on Acoustic, Speech, and Signal Processing, 2009

Robust Speaker Diarization for Short Speech Recordings

David Imseng and Gerald Friedland
Proceedings of the IEEE workshop on Automatic Speech Recognition and Understanding, Merano, Italy, pages 432-437, 2009

Robustness of Phase based Features for Speaker Recognition

Padmanabhan Rajan, Sree Hari Krishnan Parthasarathi and Hema A Murthy
Proceedings of Interspeech, 2009

SNR Features for Automatic Speech Recognition

Philip N. Garner
Proceedings of the IEEE workshop on Automatic Speech Recognition and Understanding, Merano, Italy, 2009

Social Network Analysis in Multimedia Indexing: Making Sense of People in Multiparty Recordings

Sarah Favre
Proceedings of the Doctoral Consortium of the International Conference on Affective Computing & Intelligent Interaction (ACII), pages 25-32, 2009

Speaker Change Detection with Privacy-Preserving Audio Cues

Sree Hari Krishnan Parthasarathi, Mathew Magimai.-Doss, Daniel Gatica-Perez and Hervé Bourlard
Proceedings of ICMI-MLMI 2009, 2009

Speech recognition with speech synthesis models by marginalising over decision tree leaves

John Dines, Lakshmi Saheer and Hui Liang
Proceedings of Interspeech, Brighton, U.K., 2009

Steerable Features for Statistical 3D Dendrite Detection

German Gonzalez, Francois Aguet, Francois Fleuret, Michael Unser and Pascal Fua
Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention, 2009

Structure and appearance features for robust 3D facial actions tracking

Stéphanie Lefèvre and Jean-Marc Odobez
IEEE Proc. Int. Conf. on Multimedia and Expo, IEEE, 2009

The more you know, the less you learn: from knowledge transfer to one-shot learning of object categories

Tatiana Tommasi and Barbara Caputo
BMVC, 2009

Topic Models for Scene Analysis and Abnormality Detection

Jagannadan Varadarajan and Jean-Marc Odobez
9th International Workshop in Visual Surveillance, IEEE, Kyoto, Japan, IEEE, 2009

Towards a theoretical framework for learning multi-modal patterns for embodied agents

Nicoletta Noceti, Barbara Caputo, Claudio Castellini, Luca Baldassarre, Annalisa Barla, Lorenzo Rosasco, Francesca Odone and Giulio Sandini
International Conference on Image Analysis and Processing, 2009

Visual Activity Context For Focus of Attention Estimation in Dynamic Meetings

Silèye O. Ba, Hayley Hung and Jean-Marc Odobez
International Conference on Multimedia & Expo, 2009

Visual Speaker Localization Aided by Acoustic Models

Gerald Friedland, Chuohao Yeo and Hayley Hung
ACM Multimedia, 2009

Volterra Series for Analyzing MLP based Phoneme Posterior Probability Estimator

Joel Praveen Pinto, G. S. V. S. Sivaram, Hynek Hermansky and Mathew Magimai Doss
Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2009

Wearing a YouTube hat: directors, comedians, gurus, and user aggregated behavior

Joan-Isaac Biel and Daniel Gatica-Perez
Proceedings of the 17th ACM International Conference on Multimedia, pages 833-836, ACM, 2009

Who's Doing What: Joint Modeling of Names and Verbs for Simultaneous Face and Pose Annotation

Jie Luo, Barbara Caputo and Vittorio Ferrari
Advances in Neural Information Processing Systems 22 (NIPS09), NIPS Foundation, Vancouver, B.C., Canada, MIT Press, 2009

You live, you learn, you forget: continuous learning of visual places with a forgetting mechanism

Muhammad Muneeb Ullah, Francesco Orabona and Barbara Caputo
International Conference on Robotic and Systems, 2009

The complete list, abstracts and full texts are available on the Idiap web site at the following address:
http://publications.idiap.ch