idiap
RESEARCH INSTITUTE

ANNUAL REPORT 2014

Valais
Wallis
Digital

CONTENTS

A MESSAGE FROM THE PRESIDENT

SCIENTIFIC AND HUMAN ADVANCEMENT

Olivier Dumas, President of the Foundation Council, Idiap

Visionary and therefore inherently avant-guard, bold, enthusiastic, and tenacious—the Idiap Research Institute was founded in Valais in 1991. A few thousand miles away, Mikhail Gorbachev resigned as the people's democracies of Eastern Europe broke apart. It is almost 25 years since our institute was born, and in that time it has played its part in great scientific and human advancements.

An Internal Audit with Inspiring Findings
On the eve of this anniversary year, far from resting on its countless laurels, Idiap has continued on the path of a different kind of quest—that for perpetual development. In 2014, the institute chose to conduct an audit. Administrative structure, scientific quality in broad terms, technology transfer, the relationship with our partner EPFL—all these elements were examined by a commission of international experts. A 10-page report was submitted, and its conclusion is inspiring (see pages 10–11): Idiap is an internationally renowned research institute, is highly effective, and has been able to adapt to changes in its environment, and is—besides—a great place to work.

No wonder then that Brussels has tasked the institute's director—Hervé Bourlard—in collaboration with other European partners, with the job of roadmapping those future research scenarios that correspond to the challenges that our society will, tomorrow, be facing (see page 15).

Links with the World of Business
The center of excellence IdeArk dedicated to multimodal interaction and the management of multimedia information, and led by Idiap Deputy Director François Foglia, amply fulfills its role as an interface with the business world. Numerous start-ups are developing business opportunities working closely with our researchers in Martigny. In the same vein, Idiap has begun approaching businesses present in the canton in order to place its specialist skills—in, for example, signal analysis, visual recognition, or robotics—at their disposition and to make its own contribution to resolving those difficulties found in production facilities.

Raising the Profile of Science in Valais
The ties that bind Idiap and EPFL have always been very close. EPFL is one of the founders of Idiap, along with the city of Martigny and the University of Geneva, and many of our senior researchers are now delighted to occupy roles as professors or lecturers and research supervisors at the prestigious Lausanne institution. Idiap welcomes the creation of the EPFL pole Valais-Wallis since its launch will, without a shadow of a doubt, raise the profile of a scientific Valais, which many are unaware of despite the fact that numerous institutions like Idiap have been operating successfully out of the canton for some considerable time.

More than twenty years after its birth, Idiap has changed. Valais too. The digital revolution and the concerns surrounding energy that have emerged in the "Old Country" echo those of the planet itself—concerns that, tomorrow, will urge on the fortunes of the institute, and of the canton too.

A MESSAGE FROM THE DIRECTOR

"DESPITE THE GLOOM IN EUROPE, WE HAVE BEEN ABLE TO DIVERSIFY, STRENGTHENING OUR POSITIONING"

Hervé Bourlard, Director, Idiap

Each year, at this time, the writing of our annual report demands a certain retrospection. This year, 2014, two key ideas stand out and synthesize the twelve previous months: our engagement in the regional economy and the expansion in the scope of our sphere of international activity.

Highlighting Our Work at the National and Local Levels
Our research groups—led by our best scientists—have once again endeavored to secure national or European projects, continuing their quest for innovation and their efforts to remain connected to the concerns of our society, all in the spirit of facilitating technology transfer. They are becoming more and more independent and complementary. Their work has always been recognized on an international scale, but also—and this is quite recent—it is also being highlighted at the national and local levels. This is happening not only thanks to the positions occupied by our researchers, such as Daniel Gatica-Perez, recently made an EPFL professor—and let us take this opportunity to congratulate him here—but also thanks to the transformation of certain of our innovations into regional reality.

With this in mind, Idiap is indeed very proud to present its new spin-off, BioWatch SA—the result of the participation in ICC'2014 (see page 18) of a PhD student from our research group specializing in biometrics. This new company, the ambition of which is to create a watch strap that authenticates its owner using the venous system of the wrist—more reliable than fingerprints, will certainly have a role to play in the fast-growing market for connected watches.

Maintaining Our International Positioning, despite the Gloom in Europe
Such regional success stories are all the more appreciated following the cruel twist of fate that occurred on 9 February 2014, by which the Swiss gave their popular approval to the initiative for the introduction of immigration quotas; in the process cooling somewhat the atmosphere in the ranks of European research, it must be said.

Our institute has, for example, been denied certain projects despite our tenders being awarded maximum possible points… A state of affairs that adds itself to the resizing of the lots called for by the EU. With increasingly reduced financial packages, shorter and shorter schedules, and a—often difficult to define—"priority to innovate", our scientists can no longer develop their ambitious research visions over the long term. Our institute is, thus, managing numerous research projects simultaneously, and these must be regularly renewed. Far from taking umbrage at this development, however, we have responded. We have diversified our research, better positioned ourselves vis-à-vis application areas rather than areas of fundamental research, and strengthened our strategic positioning in the field of data processing (multimedia and social media)—a discipline that lies at the heart of a great number of research projects. And, above all, we have ourselves created conditions ideal for a substantive approach.

The New Biometrics Center—A Milestone in Our Diversification
The founding—in the heart of our institute and by one of our senior researchers, Dr. Sébastien Marcel—of the Swiss Center for Biometrics Research and Testing, and its recent expansion to the European level, is certainly the most obvious recent example of Idiap's diversification (see page 17). Through this new initiative, Idiap—expert for over twenty years in biometric recognition (face, voice, etc.), and host to numerous regularly consulted reference databases—is not only positioning itself as a key player in Europe in the field of biometrics, it is also committing to becoming, increasingly, only one of the links in the long chain that connects research to practical applications that serve end users. With this center—which includes, from the outset, industrial partners in its approach—the Idiap Research Institute is truly taking the destiny of its own scientific discoveries into its own hands, and thus reaching a new milestone in its evolution.

My thanks and congratulations go to all Idiap's employees—researchers, developers, and administrative staff. Thank you all for your enthusiasm, for your professionalism, and for the friendly atmosphere you help to create. Here's to 2015!
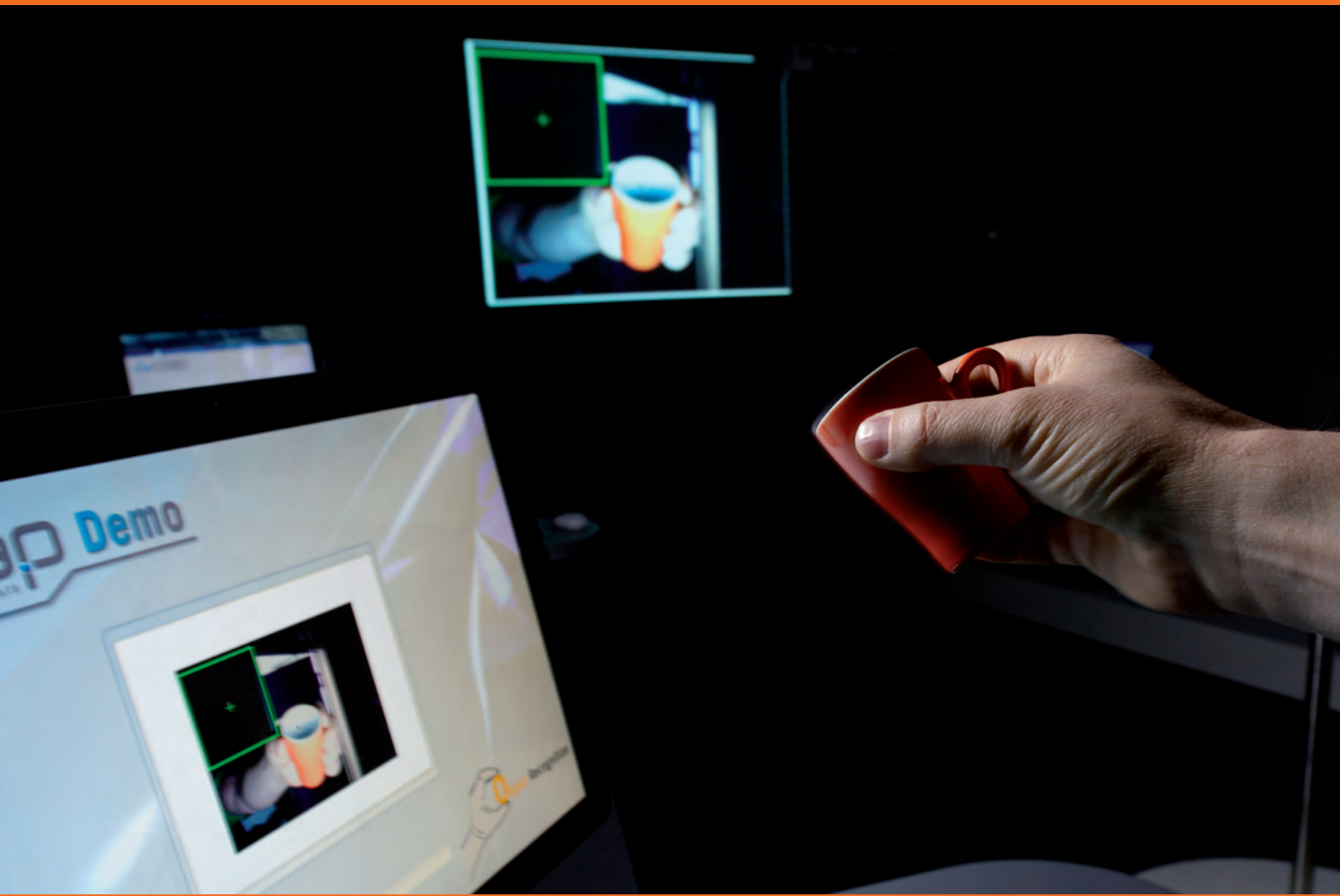
AN INSTITUTE ON A HUMAN SCALE, BUT WITH AN INTERNATIONAL REPUTATION

With over one hundred employees, and research domains in touch with current challenges, Idiap is committed to scientific progress in the service of humanity.

In 1991, at its foundation, the Idiap Research Institute first established its mission of promoting scientific progress in the service of the well-being of humanity. Today, over twenty years later, society's best interests remain at the heart of the institute's activities.

A Profusion of Technological Products
The beginning of the 21st century has witnessed the arrival of a profusion of new technological tools. On the one hand, these tools have made possible considerable improvements in efficiency and comfort. On the other hand, they disrupt people's habits, leaving some users impoverished, and others weary of constant system modifications. It is within this context that Idiap works, essentially to improve man–machine relationships, and to optimize human communication.

A National and International Network
In Switzerland, Idiap works mostly with the two federal institutes of technology, and with other universities of applied sciences and arts and cantonal universities at which research takes place. Idiap is currently working on several European projects, and maintains close links with its numerous partners, including those in France, England, and Germany. Across the Atlantic, Idiap and the International Computer Science Institute (ICSI) in Berkley, California, share a privileged partnership.

STRUCTURE

• Non-profit foundation

• Independent institution (but connected to EPFL through a joint development plan)

• Specializes in the management of multimedia information and man-machine, multimodal interactions

• Founded in 1991 by:
   • The city of Martigny
   • The State of Valais
   • EPFL
   • The University of Geneva
   • Swisscom

FINANCING

50% public funds

50% research projects

Research

Education

Technology transfer

To conduct fundamental research projects at the highest level in its chosen fields of expertise, thereby assuring its position among the best, nation-, Europe-, and worldwide. On the international stage, Idiap benefits from a broad network of partners and collaborates actively with, among others, public and private research centers, and large universities.
Competitive research projects guarantee 50 percent of the institute's financing.

To support the educational process by helping its interns discover the world of research. Idiap welcomes talented young researchers working toward their PhDs, and provides them with numerous courses at EPFL.
One in two scientific contributors is a research assistant.

To ensure technology transfer not only by promoting the widest possible dissemination of its research results into the scientific community, but also—and most importantly—by forging close ties with the world of industry.
Technology transfer generates 10 percent of Idiap's budget. Thanks to the incubator—The Ark, Idiap contributes to the creation of numerous successful start-ups.

IDIAP IN FIGURES (YEAR 2014)

Human Resources

2	professors
2	senior scientists (MER)
13	permanent senior researchers
23	postdocs
37	research assistants
20	system engineers and development engineers
10	administrative staff
29	trainees (average/year)
4	visitors

140 **people in total**

96 **full-time equivalents**

In total, 34 nationalities are represented at Idiap

Rest of Asia: 5%
Rest of the world: 3%
America: 15%
Switzerland: 30%
Rest of Europe: 24%
India: 12%
France: 11%
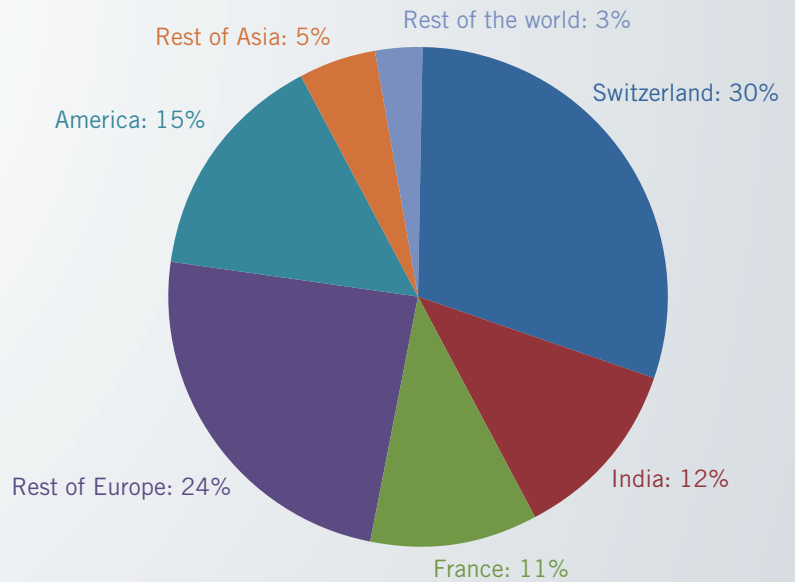
Scientific activities

- Participation in sixty-two research programs
- Project management in twenty-two consortia (Switzerland, Europe, industrial projects)
- Participation in the economic development strategy of the Canton of Valais through the program—The Ark and, in particular, the incubator IdeArk
- Forty-seven posts in seventeen IdeArk start-ups
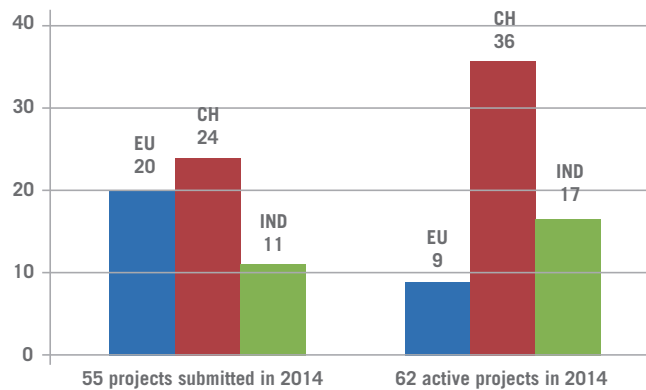
Submission and financing of research projects
In 2014 Idiap submitted 55 projects (20 to European agencies, 24 to Swiss funding agencies and 11 industrial projects). At the same time the institute managed 62 active projects.

	55 projects submitted in 2014	62 active projects in 2014
EU	20	9
CH	24	36
IND	11	17

Publication type	Count
Thesis completed	6
Conference articles	88
Book chapters	8
Scientific articles	28
Internal communication reports	3
Research reports	22
Internal research reports	103

Publications
In 2014, Idiap researchers contributed to 258 published works.

5 RESEARCH THEMES & TEN APPLICATION DOMAINS

HUMAN–HUMAN, HUMAN–MACHINE & HUMAN–ROBOT INTERACTION
Voice-controlled devices, voice–voice translation systems, navigation systems

HOME AUTOMATION (DOMOTICS)
Multi-sensor activity analysis, human-behavior analysis, adaptation to human behavior (e.g., heating), home safety and security

EXPLOITATION OF RICH MULTIMEDIA ARCHIVES
Semantic indexing, object recognition, audio–video content filtering, YouTube, document analysis, analysis of cultural heritage media

ENTERTAINMENT AND GAMES
Multi-lingual gaming, remote-family games ("togetherness")

COLLABORATIVE AND CREATIVE SYSTEMS
Smart meeting rooms, video conferencing, multimedia indexing and access, cross-lingual collaboration

CITIES OF THE FUTURE (SMART CITIES)
Ecology, environmental management, pollution reduction, traffic reduction, better use of road networks, noise reduction

MANAGEMENT OF MOBILE SYSTEMS
Signal processing for mobile platforms, mobile social networks, modeling of human activity to anticipate needs

SECURITY AND RISK MANAGEMENT
Access control (for physical or virtual spaces), speaker identification, face identification, video surveillance, natural-risk modeling, safety of public and private spaces

ENERGY MANAGEMENT
Heterogeneous energy-networks, sensor networks, modeling of human activity to anticipate needs

HEALTH AND BIO-ENGINEERING
Smart patient (data-) management, protheses (auditory, corporal, etc.), bio-systems modeling, interfaces for the disabled, capture and indexing of medical documents

Central wheel

- Application areas -

- Research Themes -

PROJECTS

Research Themes:
- **A** Perceptual and cognitive systems
- **B** Social / human behavior
- **C** Information interfaces and presentation
- **D** Biometric person recognition
- **E** Machine learning

Application areas:
- Human-human human-machine & human-robot interaction
- Exploitation of rich multimedia archives
- Collaborative and creative systems
- Mobile computing
- Energy
- Healthcare & bio-engineering
- Security and risk management
- Smart cities
- Entertainment
- Home automation (domotics)

Numbered markers: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10

A Perceptual and cognitive systems
(automatic speech recognition, computer vision, handwriting recognition, multimedia document processing, robotics, machine translation)

B Social/human behavior
(web and mobile communication, social interaction, social signal processing)

C Information interfaces and presentation
(multimedia information systems, personalization and contextualization)

D Biometric person recognition
*(speaker recognition, face recognition)*

E Machine learning
(statistical modeling, neural networks, mathematical models)

RESEARCH GROUPS

Ten groups, each led by one or more top-level scientists, divide the institute's research activities between them. The group responsible for speech- and sound processing is led by the institute's director—Hervé Bourlard—and three senior researchers. The remaining groups are all headed by senior researchers.
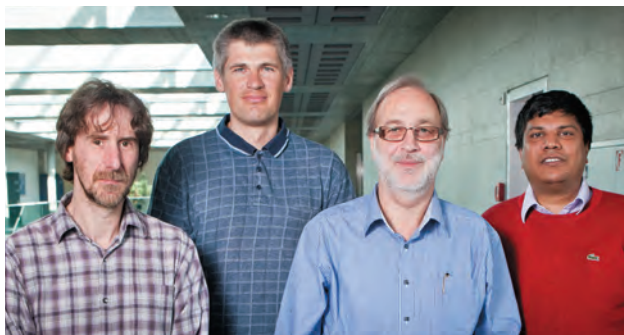
Speech & Audio Processing

Prof. Hervé Bourlard (second from right)**, Dr. Phil Garner, Dr. Petr Motlicek, Dr. Mathew Magimai Doss**
H-index*: HB: 49, PG: 19, PM: 11, MMD: 19

Speech processing has been one of the mainstays of Idiap's research portfolio for many years. Today it is still the largest group in the institute, and Idiap continues to be recognized as a leader in the field. The expertise of the group encompasses statistical automatic speech recognition (based on hidden Markov models or hybrid systems exploiting connectionist approaches), text-to-speech, and generic audio processing (covering sound source localization, microphone arrays, speaker diarization, audio indexing, very-low-bit-rate speech coding, and perceptual background noise analysis for telecommunication systems).

Social Computing

Dr. Daniel Gatica-Perez
H-index: 42

Social computing is an interdisciplinary domain that integrates theories and models from mobile and ubiquitous computing, multimedia, machine learning, and social sciences in order to sense, analyze, and interpret human and social behavior in daily life, and to create devices and systems that support interaction and communication. Current lines of research include ubiquitous sensing of face-to-face interaction, behavioral analysis of social video, crowdsourcing, and urban data-mining using smartphones and mobile social networks.

Computer Vision & Learning

Dr. François Fleuret
H-index: 24

The scientific objective of this group is the development of new statistical learning techniques, mainly for computer vision, with a particular interest in their computational properties. Application domains include image recognition and scene analysis, the tracking of people and biological structures, and the understanding of images in general.

Perception & Activity Understanding

Dr. Jean-Marc Odobez
H-index: 31

This group conducts research into human–human activity analysis using multimodal data. This entails the investigation of fundamental tasks such as the representation, detection, segmentation, and tracking of objects and people, the characterization of their state, and the modeling of sequential data and the interpretation of that data in the form of gestures, activities, behavior, or social relationships. These investigations take place through the design of principled algorithms that extend models from computer vision, statistical learning, or multimodal signal processing. Surveillance, traffic analysis, analysis of behavior, human–robot interfaces, and multimedia content analysis are the main application domains.

*The h-index attempts to quantify the productivity and impact of a given researcher based on the quotation pattern and frequency of that researcher's publications. The higher the index number, the more the citation level is important.

Robot Learning & Interaction
Dr. Sylvain Calinon
H-index: 28

The Robot Learning & Interaction group focuses on human-centric robot applications. The scientific objective is to develop probabilistic approaches for encoding movements and behaviors in robots evolving in unconstrained environments. In these applications, the models serve several purposes (recognition, prediction, online synthesis), and are shared by different learning strategies (imitation, emulation, incremental refinement, or exploration). The aim is to facilitate the transfer of skills from end users to robots, or between robots, by exploiting multimodal sensory information and by developing intuitive teaching interfaces.

Artificial Cognitive Systems
Dr. Barbara Caputo
H-index: 29

The Artificial Cognitive Systems group works on the development of multimodal learning algorithms to enable artificial agents to act autonomously in realistic settings, with a special emphasis on their ability to—autonomously—detect knowledge gaps and fill them with open-ended learning strategies. The focus of this work is on designing algorithms that are principled and computationally efficient and that provide robust performance in very realistic settings while at the same time providing theoretical guarantees on expected behavior. The group is funded by national and international funding agencies.

Applied Machine Learning
Dr. Ronan Collobert
H-index: 24

This group is interested in computer algorithms that can "learn" a behavior in order to perform a given task of interest, in contrast to algorithms with behavior constrained by hand-crafted rules. This research is driven by real-world applications involving large amounts of data. Domains of interest include natural language processing, computer vision, and audio processing. A particular emphasis is placed on generic machine-learning tools, which require minimum a priori knowledge of the data (such as deep-learning techniques), and on unsupervised learning techniques, which can leverage inherent semantics from large-scale, structured data.

Computational Bio-Imaging Group
Dr. Michael Liebling
H-index: 17

This group focuses on research into computational imaging and the analysis of biomedical images. This includes developing algorithms for image deconvolution and super-resolution in optical microscopy, three-dimensional tomography reconstruction from projections, and—more generally—combining unusual sensing devices and approaches with computational methods to produce images ideally suited to the observation and quantification of complex and live biological systems.

Biometric Person Recognition
Dr. Sébastien Marcel
H-index: 28

Biometrics refers, in computer science, to the automatic recognition of individuals based on their behavioral and biological characteristics. The Biometric Person Recognition group investigates and develops novel image-processing and pattern-recognition algorithms for face recognition (2-D, 3-D, and near-infrared), speaker recognition, anti-spoofing (attack detection), and emerging biometric modes (EEG and veins). The group is geared toward reproducible research and technology transfer, using its own signal-processing and machine-learning toolbox.

Natural Language Processing
Dr. Andrei Popescu-Belis
H-index: 29

The Natural Language Processing group studies how the semantic and pragmatic analysis of texts can improve the execution of two important tasks—machine translation and information retrieval. In particular, the group studies how the recommendation of information from networked multimedia repositories can be improved by analyzing a network's structure and the information content of its nodes.

IDIAP UNDER THE MICROSCOPE OF THE WORLD'S RESEARCH ELITE

In September 2014, seven of the scientific world's leading figures conducted an extensive audit of Idiap at the request of the institute's director—Hervé Bourlard. The seven experts were unanimous in highlighting Idiap's qualities, and also indicated some avenues that could lead to even further progress.

In 2014, Idiap was just two years away from celebrating its 25th anniversary and about to begin the process of renewing its public funding for the period 2017–2020. "It seemed the ideal moment for us to review the institute as a whole," says Bourlard, "including our research activities, our organization, and our relationship with EPFL." A rare process of reflection, welcomed by the audit committee, which also congratulated Professor Bourlard on his visionary and passionate management approach.
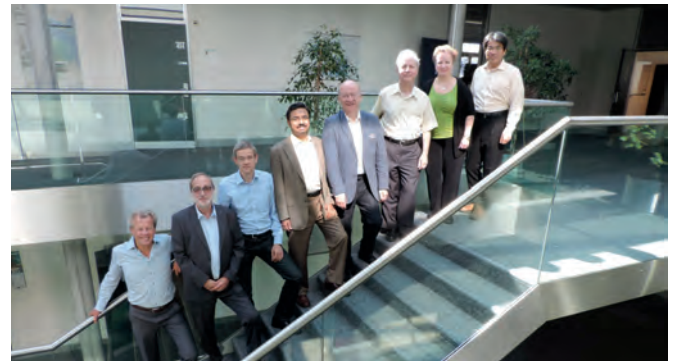
RESEARCH ACTIVITIES

Scientific Projects Provide Half of the Institute's Budget

The number of projects undertaken by the institute is growing while their size and duration have decreased, large projects such as IM2 or AMI/AMIDA having come to an end. Currently, more than half of Idiap's annual budget is drawn from national and international research projects. That Idiap, in this context, continues to increase its research budget is a positive sign.

Quality Resources and an Open-Source Philosophy

The best specialist journals publish articles by Idiap researchers, which reflects the degree of excellence of the institute. Its researchers also contribute in a remarkable manner to the creation of quality resources, particularly databases and software, often open-source in nature.

Constant Development of New Research Domains

In parallel to consolidating its pioneering activities in the field of speech recognition, now also applied to dialects—such as that of Upper Valais, for example—without lexicons, Idiap pursues its development activities in new research domains including robotics and biomedical imaging.

The Method—Preparatory Research, Presentations, and Meetings

In addition to documents provided by Idiap, the committee carried out the audit thanks to three days spent on site at the institute, from 3 to 5 September 2014. On the agenda, presentations by researchers, managers, and members of the scientific college, and face-to-face interviews with young researchers, postdocs, and graduate students. A few weeks later, the committee submitted a detailed report highlighting Idiap's strengths, and made suggestions regarding further improvements.

MANAGERIAL AND STRUCTURAL ACTIVITIES

The Collegial Atmosphere Helps New Arrivals

The friendly, collegial atmosphere, a culture of excellence, and the quality of the infrastructure present combine to make the Idiap research institute a great place to work. Students were particularly pleased with the help they received during their settling-in period in Martigny, and with the various measures aimed at encouraging social interaction.

High-Profile and Attractive

The 18 posts advertised over the last year attracted more than 500 applications. This illustrates the excellent visibility and attractiveness of Idiap, where no fewer than 34 different nationalities rub shoulders.

RELATIONSHIP WITH EPFL

Significant Participation in EPFL's Academic Mission

Creating course materials, teaching, publishing scientific articles—Idiap's researchers play an important role in the academic mission of EPFL. Not to mention that they are training, at any one moment, over 40 PhD students.

Idiap—A Unique Setup That Deserves a Higher Profile

The diversity of its scientific domains and the scope of its mission (research, technology transfer, and training) make Idiap a unique setup that deserves greater recognition from its various partners. The committee recommends, in particular, that Idiap researchers be appointed to academic positions at EPFL.

RECOMMENDATIONS

Structure, Relationship with EPFL, Development, and Reputation

In concluding its report, the audit committee presented certain recommendations for improving Idiap's productivity and impact.

- To complete its organizational structure, Idiap could establish an advisory group—including scientists—tasked with planning future infrastructure investments.
- An intensification of the relationship with EPFL can be achieved by securing EPFL academic appointments for Idiap researchers and by establishing better connectivity between the two entities.
- The committee encourages Idiap to continue its exploration of new application domains (medicine, energy, etc.) and to create a partnership with those European organizations that are likely to facilitate Idiap's integration into major projects in the long term.

KNOWLEDGE AND TECHNOLOGY TRANSFER

Effective Solutions to Meet the Needs of Industry

Adapting research findings to the needs of the industry (technology transfer) is part of Idiap's mission. The institute is particularly effective in establishing influential practice (developer groups, the IdeArk business incubator, the International Create Challenge, etc.), to the great satisfaction of its partners, including industrial giants such as Nokia, Logitech, Yahoo, Samsung, and HP.

At the Service of the International R&D Community

Attentive to local, cantonal, and federal economic development, Idiap also places its expertise at the service of the international R&D community. Highly active in the transfer of knowledge between academic and industrial institutions, Idiap creates and makes available a significant number of professional software packages–a service that a traditional university structure could not hope to deliver.

Innovation Driver

By positioning itself as an innovation driver, Idiap is adopting a strategy that benefits Swiss industry. The recently created biomedical image processing research group and Swiss Center for Biometrics Research and Testing illustrate the institute's intention of remaining in tune with contemporary economic and societal needs.

JOIN IN CONSERVING THE COLLECTIVE MEMORY OF VALAIS

The project "Valais*Wallis Digital"—conceived and designed by Idiap in the frame of the bicentenary of the canton's entry into the Swiss confederation—invites *Valaisannes* and *Valaisans* to share their memories and archives (photos, videos, texts, etc.) with fellow citizens via its internet platform. If it is to succeed, the project needs to mobilize the entire canton.

1. Visit www.valais-wallis-digital.ch and create a user account

The platform is free to use and open to everyone, as is searching the database and consulting files.

Creating a user account allows you to submit your archives, add additional information to them, add comments to others' files, and contact other users of the platform.

2. Dust off your archives and search for treasures hidden away in your attic

Whether it's family photos, texts, or audio or video files, every item may be of historical interest. Don't let them be forgotten. Share them and help preserve the collective memory of Valais.

3. Digitize them

Only digitized documents can be submitted to the internet platform. To do this, you can either digitize them yourself using, for example, a scanner or a digital camera, or call a company that specializes in digitizing data, such as Martigny's *Cinetis*—a Valais*Wallis Digital project partner.

4.Share them on www.valais-wallis-digital.ch

Once your files have been digitized, connect to the online platform. Using the "My Documents" feature, you can now upload them, add comments, and share them with the online community.

While helping to conserve the Canton of Valais' heritage, you will also be directly contributing to research carried out at Idiap in the domain of digital sciences. This includes managing and indexing large multimedia databases created as a result of mass collaborative or participative projects of this and other kinds. This type of research has an impact that goes far beyond the Valais*Wallis Digital project—aiming to develop new technology and tools for addressing the major societal challenges of the future, such as energy management, health care, and the environment.

A Set of Cards Distributed by Migros in "Mania" Form

To encourage mass participation in the project, a set of cards has been created in collaboration with Migros Valais and the Academy of Contemporary Arts (EPAC) at Saxon. Featuring 200 years of the history of the "Old Country"—illustrated via as many cards—"Valais Digital Mania" offers a fun and interactive approach to the highlights of the period 1815-2015, each card presenting, in an entertaining way, one of 200 of the canton's historic events.

What's more, a free mobile application—designed to facilitate interaction with the Internet platform—provides collectors with more historical detail by offering, once a card is scanned, additional information concerning the event in question.

INEVENT—INDEXING MULTIMEDIA INFORMATION

The result of a three-year European project, the portal *inEvent* helps its users find a path through the jungle of multimedia data available on the Web. The Idiap-led project should help introduce the term "hyper-event" into our daily vocabulary.

Music videos, movie clips, mass production of leisure and hobby videos, streaming, online TED-type conferences, distance learning … Every minute, no fewer than 300 hours of video are uploaded to YouTube. And in the worlds of academia and business in particular, the use of video conferencing is becoming more and more widespread.

While systems such as Google are good at recognizing and sorting text, the indexing of audio and video content is—in comparison—a decidedly sketchy affair. How can this be remedied? The challenge is all the greater because, in addition to automatic transcriptions of videos, online comments and social network "likes" can sometimes be appended to the contents of the recorded events.

Multi-Dimensional Indexing of Multimedia Events Thanks to "Hyper-Events"

Conducted from November 2011 to October 2014, the European FP7 project "inEvent" aimed to provide users with multi-dimensional indexing for easier access to, and navigation through, these vast multimedia databases. Coordinated by Idiap, *inEvent* brought together academic and industrial partners, including giants IBM and Radvision, but also Klewel, an Idiap spin-off. The project's goal was to create a system that not only indexes multimedia experiences, but analyzes them from the inside—hence the name *inEvent*—and forges complex and pertinent links between them.

Called "hyper-events", by analogy with the term "hypertext", these links enrich multimedia content and make multi-dimensional searches possible: identifying, for example, conferences addressed by the same speaker and other events touching on the same theme, linking events taking place at the same location, or proposing similar content.

An Innovative Graphic Interface

"*InEvent* builds intelligent databases," says director of Idiap and *inEvent* project manager, Hervé Bourlard. "This project is a real tour de force of innovation. It includes the automatic segmentation of video, speech transcription, the identification of speakers, and semantic audio and video interpretation—all our areas of expertise." The design of the interface, with its keyword clouds, makes it very easy to identify the most relevant results—an intelligent approach that was recognized with an award at the ACM Multimedia Grand Challenge 2013, the world's leading conference on multimedia processing.

Another strength of *inEvent* is how it takes into account the emotional dimension of information. Imagine a sequence showing someone falling in the street. Is it dramatic or funny? Would a European viewer perceive it in the same way as an American or Chinese viewer? "The events of January 2015 in Paris showed us that not all cultures experience everything in the same way—humor being a particular example. With this in mind, it's easy to see how our system could become more and more useful in the future."

The system has been enthusiastically received and major partners have already expressed an interest. "They're unanimous in recognizing *inEvent's* potential for the world of business, but also for education and social media."

www.inevent-project.eu

Available in an Open-Source Format

The future direction *inEvent* takes depends not only on the project's industrial partners, but also on the wider world of software development, since inEvent project leaders have chosen to release code to the public domain in an open-source format. By doing so, they have made a significant contribution to our increasingly media-oriented society, calling on the talents of external developers to ensure further system improvements. Perhaps, in a not-so-distant future, the term "hyper-event" will be on the lips of all the world's digital media users.

A CONCEPTUAL REVOLUTION NAMED PHASER

Idiap is working on a system of automatic speech recognition that should go far beyond the present state of the art. More robust and simpler than current systems, the model could also be more compatible with the mechanisms of the human auditory system as we understand them, leading to an impact far beyond speech recognition alone. The project, called PHASER, is funded by the Swiss National Science Foundation (SNSF).

© Fotolia- lassedesignen

How to exploit vast databases to build a speech recognition system that remains robust in the face of all kinds of variables? A system whose effectiveness will not be reduced either by the different accents of speakers, or by the recording conditions? Since June 2014, under the supervision of Prof. Hervé Bourlard, Idiap researchers (Dr. Afsaneh Asaei and research assistant Pranay Dighe) have been working on answering these questions in the context of a project funded by the SNSF: PHASER—Parsimonious Hierarchical Automatic Speech Recognition.

A Change in Paradigm

"We're supporting a paradigm shift in speech recognition. So far—and for over forty years now—every solution (commercial or research) has operated largely on purely statistical systems," explains Afsaneh Asaei, postdoctoral researcher at Idiap. In these models, only certain specific characteristics of the recorded data are extracted.

"By exploiting the recent resurgence of neural networks (deep neural networks)—a discipline never actually abandoned by Idiap—and the almost unlimited availability of memory and computational resources, the system proposed by PHASER aims to be able to conserve and fully exploit all data."

Combining Two Strategies to Improve Effectiveness

To build a system of automatic speech recognition using a large amount of data, the Idiap researchers combine two strategies. The first promotes simplicity and a parsimonious representation of information. This parsimony constraint usually leads to simpler and more robust systems.

"The second strategy favors structure and hierarchy. For example, to understand the content of a book, you first look at the jacket, then the title, and then the table of contents before reading all the details. This strategy is exploited by (human) cognitive systems. So it's very interesting to design a machine capable of carrying out tasks in a similar way."

An Award-Winning Innovation

PHASER therefore promises a particularly effective system of automatic speech recognition that implements two principles: simplicity or parsimony on the one hand, hierarchical structure on the other. Idiap's combination of these two key components represents an entirely new proposition. Indeed, one of the first research papers to present the premises of this approach was awarded the Best Doctoral Paper prize in 2011 at a conference (ICASSP—International Conference on Acoustics, Speech and Signal Processing) organized by IEEE (the Institute of Electrical and Electronics Engineers), an international professional association.

ROCKIT—ROADMAPPING TOMORROW'S EUROPEAN RESEARCH

In the shape of the European "ROCKIT" project, Idiap and its partners have been tasked with guiding research into human–human and human–machine interaction.

Draw up an inventory of current research, develop a strategic roadmap that identifies societal and industrial challenges, and thus identify those areas of research and development that require a particular focus between now and 2020. These are the aims of the ambitious European FP7 program-funded project ROCKIT (Roadmap for Conversational Interaction Technologies), launched in early 2014 and due to run for a period of two years. The stakes are high—the project must define the societal challenges that the European Union will face and assist the union in the development of a roadmap involving academia, industry, and the political milieu.

"The domain addressed by ROCKIT is natural language and conversation agents, or more generally anything that concerns multi-sensor interaction between humans and the devices of tomorrow: speech, translation, gestures, robotics, etc."

explains Bourlard—the project's director at the institute. ROCKIT, coordinated by the University of Edinburgh, naturally has a seat at the table for Idiap—one of six project partners—given that the institute's research themes are so closely related to this field.

"Our task was to design five research development scenarios for the next ten years" (see below). Another aspect of the project is to bring together all the key players from the areas concerned under a single European research and innovation alliance that also involves commercial partners from the sectors concerned.

©David Tilly

1. Adaptable Interfaces for all
Like a chameleon, interfaces should be able to adapt automatically to all their users—including children, the elderly, or disabled users—and to all circumstances. You receive a call while your hands are busy with a task? The phone detects this and automatically switches to visual mode, allowing you to pick up the call just by looking at the phone.

©Trebola

2. Smart personal assistants (SPAs)
A booming sector for technology, smart personal assistants help us in our everyday lives, while automatically adapting to the user, the context of use, and the sensors available, etc. While, today, they already provide appropriate responses to a whole range of questions, tomorrow they will be able to capture our emotions and, thus, guide users according to their preferences (travel, shopping, etc.).

ROCKIT partner institutions

University of Edinburgh—School of Informatics (GBR)
Idiap Research Institute (CHE)
Europe Unlimited (BEL)
LT-Innovate, the Forum for Europe's Language Technology Industry (Europe)
Saarland University (DEU)
Vodera (GBR)

© Alexander Vasenin

3. Active access to complex, unstructured information
Who has never felt lost among the increasingly large and heterogeneous information flows that are being endlessly generated? Dynamic, interactive systems will help users orient themselves. Better still—being proactive—they will sort out the relevant data to be transmitted, and find information more rapidly, without the need to explicitly launch a search tool.

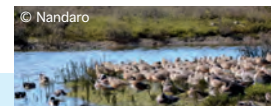©Doug Janson

4. Communicative robots
The development of robots, including "social robots", that interact with humans and adapt to their environment opens up broad new perspectives—robots as carers for the elderly or disabled, replacing humans to carry out difficult or dangerous tasks, carrying out surgical operations, etc. Many applications also concern children, providing educational tools and games, and also therapeutic support, particularly in cases of autism.

© Nandaro

5. Shared collaboration and creativity
This scenario aims to deepen communication—multidisciplinary, multi-language, and "multi-social"—between humans and make it more effective. It is built on the strengths of the collective, which itself boosts creativity and facilitates problem solving. Social platforms and networked games are just two examples of systems that encourage us to work together, whether to share information within a multidisciplinary group or to play or boost our collective and individual creativity, while translation systems abolish the language obstacle. The result? Individuals are enriched by their contact with the collective group.

BIOMETRY: AN ESSENTIAL SECURITY TOOL

Even if biometric recognition is today being integrated into more and more of the appliances that we use in our everyday lives, its reliability still raises many questions. A pioneer in the field of image and speech recognition, Idiap has hosted—since April 2014—the Swiss Center for Biometrics Research and Testing, and its European spin-off, EAB-CITeR.

No more passwords: the future is biometric identification, whether by fingerprint, iris, or face recognition, to mention only the most common methods currently available. Just as these technologies are becoming increasingly important in the fight against criminal activity, several projects—coordinated by Idiap at the European level (MOBIO—Mobile Biometry, TABULA RASA, and BEAT—Biometrics Evaluation and Testing)—are seeking to increase their reliability.

With the financial support of the Canton of Valais and the City of Martigny, and with the aim of positioning itself as European leader in the field, Idiap officially opened—in April 2014—the Swiss Center for Biometrics Research and Testing. Its objective: to enable the scientific community to carry out research, in a coordinated manner, by exploiting the institute's widely recognized expertise and the significant quantity of biometric data and software gathered at Martigny as a result of past, present, and future projects. "Only two similar structures exist in the world: one in the US and one in China. We will, in fact, be collaborating with the Americans," says Sébastien Marcel, senior researcher at Idiap and director of the center.

Research Funded by Corporate Contributions
The contacts established during the first meetings organized within the framework of the new center quickly convinced the sector's major European players to unite behind a common project, and so EAB-CITeR (the European Cooperative Identification Technology Research Consortium) was born.

Inspired by the American model (US-CITeR), the new competence center is based at Idiap under the control of the recently created Swiss center. It promotes a cooperative research model, uniting scientists, industry, and government agencies. "In practical terms, the institutions that join the project bring with them companies or other interested partners, who make financial contributions. The sums collected provide basic funding for research, the themes and projects of which are decided by the contributors. This mixed, public–private approach offers unparalleled opportunities for collaboration and cooperation between industry and researchers in order to advance biometric security technology." Thanks to initial financial support from the Canton of Valais and the City of Martigny, the center aims to be self-financing within one to two years.

Assessing the Compliance Levels of Programs and Applications
In addition to the research carried out at the center, aspects vital to the field—such as assessing results obtained and evaluating products going to market—are also top priorities. Joining itself to the European BEAT project, an online platform has been made available to researchers and companies enabling them to test their algorithms and software on the numerous biometric databases stored at Idiap. Besides the advantage of encouraging knowledge sharing within the scientific community, the tool also makes it possible to issue compliance certificates confirming the conformity of the programs tested, the validity of the results obtained, and the performance of the applications evaluated.

The center also works on spoofing—a group of attacks that are becoming more and more sophisticated, and that include the falsification of fingerprints, or facial reproduction with the aid of masks. "It's a question of identifying vulnerabilities and developing appropriate countermeasures. Some research paths are oriented toward the detection of vital signs, such as the blinking of the eyes, heart rate, or the use of new biometric modalities such as vein patterns, which are considered tamper-proof."

A CHALLENGE IN THE FORM OF AN ACCELERATED BUSINESS INCUBATOR

For the third consecutive year, Idiap's International Create Challenge (ICC) has enabled researchers to transform their ideas into commercial prototypes. The three weeks of immersion at the institute have once again proved a resounding success. The 2014 edition rewards three projects that may, tomorrow, improve our daily lives.

From 26 August to 16 September 2014, Idiap hosted—in the form of a total immersion program—seven teams of aspiring entrepreneurs. In total, twenty scientists, designers, and prospective businessmen and women from Switzerland, the United States, Greece, Italy, or Turkey worked intensively on their respective projects.

Brand-new this year, a crowdfunding site to generate support for all projects submitted in the framework of ICC'2014: funding.idiap.ch. While the principal goal of such crowdfunding remains the visibility afforded to the projects, this new approach has also boosted the ability of young entrepreneurs to promote their products. And "certain investments indicate that the region has no shortage of investors and that this is an interesting approach," says François Foglia, deputy director of Idiap.

The jury of the 2014 edition—composed of leading figures from the worlds of research, venture capital, and innovation—distinguished three projects that combined innovation and high scientific quality.

BioWatch Secures Smartwatches by Reading the Veins in the Wearer's Wrist
BioWatch is the world's first smart, biometric watch that uses authentication based on the wrist's venous system. Goodbye passwords! Thanks to a sensor situated in the wrist strap, BioWatch identifies its owner with a reliability rate of 99.5 percent; better than fingerprints, or even voice- or iris-recognition. Inspired by the work of Joe Rice, a pioneer in biometric recognition using veins, the project is principally driven by Pedro Tome, postdoctoral researcher, and Matthias Vanoni, research assistant at the institute.

Horus—Personal Assistant for the Visually Impaired
How do you choose matching clothes when you are visually impaired? Or read a bus timetable, recognize a friend, or see a crosswalk? Under the watchword "The invisible becomes audible", the Horus system looks like a small box that fits all types of glasses. The small camera it contains monitors the environment, analyzes it, recognizes images, and transmits information to the wearer orally. Horus can also store information, such as faces for example, according to a user's needs. A true revolution for the visually impaired.

Anemomind—All Sails to the Wind
In November 2014, Alan Roura, the only Swiss sailor competing in *La Route du Rhum,* sailed with Anemomind—the "Digital Advisor for Smart Sailors"—onboard. The system processes data from a vessel's navigation instruments (wind speed, water temperature, etc.), stores and analyzes them, provides real-time diagnostics, and offers comparisons with previous readings, and even the possibility of sharing data. An idea born from the fact that currently used systems are complicated to calibrate and provide information that is difficult to use. Notably, Anemomind—the electronics of which were designed at the HES-SO Valais–Wallis—will be manufactured in the canton (Collombey-Muraz).

www.biowatch.ch

http://horus.technology/en/

www.anemomind.com

THE HASLER FOUNDATION—SUPPORTING SWISS RESEARCH AND INNOVATION

The Hasler Foundation, whose incentive programs address the worlds of both research and business, regularly supports scientific projects and projects involving significant innovation, both at the national level. Since 2010, the foundation has contributed financially to around twenty of Idiap's projects.

Based in Bern, the Hasler Foundation (Haslerstiftung) invests several million Swiss francs each year in encouraging the development of information and communication technologies. It does so "for the good and the benefit of the Swiss cultural and professional milieus," and with the aim of ensuring for the country "a role of leader in the fields of science and technology."

Since 2010, around twenty Idiap projects have benefited from the financial support of the Hasler Foundation. Much of this support—so important for the institute—comes in the form of an independent research program set up to aid non-commercial institutions, providing funding for individual science and technology projects for up to three years.

Supporting Innovative Projects

In addition to fundamental or basic research, the foundation also supports the transfer of high-potential research findings when creating a spin-off is envisaged. This challenge—although extremely rewarding, both for the research milieu and for the young entrepreneurs concerned—is always difficult to overcome. For this reason, the foundation offers the future entrepreneurs engaged in the projects it supports additional resources for a maximum of 18 months in order to help the project survive this difficult phase.

The latest such measure granted to Idiap has encouraged the creation of the institute's latest spin-off, recapp IT AG. The project, funded to the tune of CHF 250,000, is supervised by Dr. David Imseng—recent graduate of EPFL/Idiap, research assistant at the Martigny institute, and co-founder of recapp IT AG. "For us, this aid is essential and comes at just the right moment," says the researcher from Upper Valais. "Everyone wins. The Hasler Foundation because it knows that the money will be used well because it's being managed by a recognized and experienced research institute; Idiap because innovation and technology transfer are among its missions; and finally me, because I can put a team together and completely focus on the project."

Passionate about multilingual speech recognition—he dedicated his doctoral thesis to the subject—Imseng and recapp IT AG are developing a mobile application to record, archive, and reuse the content of meetings. With, as its starting point,

the experience gained and the results obtained from the Valais parliament-based "Mediaparl" project (see AR12, pages 10-11) recap IT AG wants to further push the development of this tool, all the while focusing primarily on the Swiss market, and its languages and dialects.

Apple and Microsoft Interested

Already attracting interest from giants such as Apple and Microsoft, the application has also charmed other cantonal parliaments and is fêted wherever it is presented. Winner of the International Create Challenge 2013, finalist in Banque Cantonale du Valais' Prix Créateurs BCVS 2014, the application won grand prize in 2014 in LAUNCH—the UC Berkeley Startup Competition—where it competed with some 200 projects worldwide.

With the support of the Hasler Foundation, of Idiap, and of the business incubator The Ark, the recapp IT AG team is now working on its finished product. "We're a Swiss start-up, which uses Swiss technology, and makes a Swiss product." recapp IT AG—an innovation on a red field with a white cross.

recapp IT: www.recapp.ch

recapp IT AG's application recognizes languages, automatically generates text from the spoken word, and indexes keywords. Via a very rare feature, it can even manage speech that combines several languages, a situation so typical of many domains in Switzerland. It is already attracting interest from several cantonal parliaments, as it facilitates the work of those charged with recording, archiving, and making available the content of parliamentary sessions and debates.

From the Telephone to Information and Communication Technology
The Hasler Foundation—named the Hasler Works Foundation at its inception in 1948—was originally a corporate foundation. Its founder, Gustav Hasler (1877-1952), wanted to ensure that the Hasler Works, which had developed from the Federal Telegraphy Works, would continue to be run in the same spirit after his death. The foundation was also tasked with supporting "the telephony and telegraphy branches in Switzerland."
Today, free of any commercial interests, the foundation promotes education, research, and innovation in the field of information and communication technologies via annual investments of several million Swiss francs.
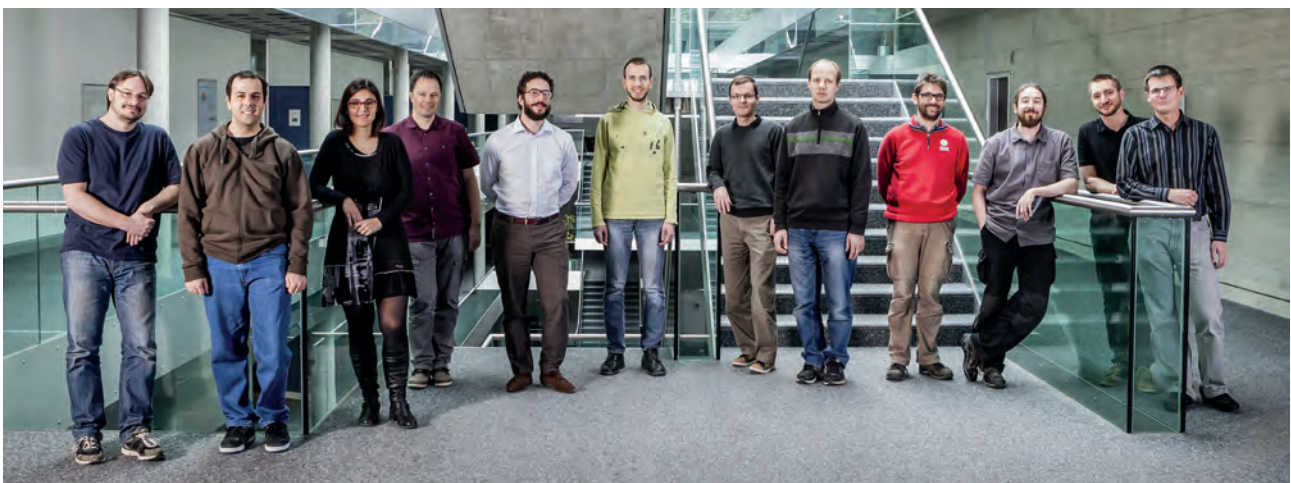
www.haslerstiftung.ch

THE TRANSMISSION CONNECTING THE MOTOR OF RESEARCH TO THE WHEELS OF INDUSTRY

Roles and responsibilities: the institute's development team continues to grow and transform. Spotlight on the behind-the-scenes workers who help convert Idiap's innovations from algorithms into products.

Without them, researchers' findings would never go beyond the realm of a set of mathematical formulas. Without their contribution, Idiap's innovations—even the most amazing—would never leave the field of research. They are the development engineers. People who—as the team's head, Olivier Bornet, observes aptly—"take research notes with them to meet with industry, so that in the end everyone can use the product." Engineers in their own right, they also contribute important—and often new—elements to Idiap scientists' findings, for example developing fully functional prototypes, which sometimes highlight certain shortcomings that need to be addressed.

Bornet took charge of the group in March 2014. A native of Nendaz, he knows the place well, as it is here that he took his first steps as a developer back in 1994—at the time, alone in holding such a position at Idiap.

Stimulating and Convivial Teamwork
Little by little, the department is expanding with the arrival of new colleagues—a logical evolution since technology transfer is among Idiap's missions, alongside research and training.

For several years, the development engineers were scattered throughout the site, each working individually for one researcher or another. Since late 2006, however, they have been working as a team, quartered on the fourth floor of the Centre du Parc. "This form of organization makes us more versatile; we're no longer tied to one particular project; rather we're linked to several," notes Bornet. "This creates a stimulating working atmosphere. Our connection goes beyond a mere working relationship: most of the developers take their breaks together and eat together at noon, alongside researchers and other employees of the institute."

Showcasing the Research Carried Out at the Institute
The development team currently has 14 members and a range of profiles. Some followed the HES (universities of applied sciences) route—at Sion or Yverdon, for example—others are from EPFL, and four have (EPFL) PhDs. "The presence of PhDs on the team is an important asset. They facilitate communication with the researchers, with whom they share similar knowledge and skills."

The whole team pulls toward the same goal: showcasing the many technologies produced by the institute's researchers in order to capture the interest of the industrial world. The creation of demonstrators is part of that mandate. Demonstrators like "IDEA", which creates a real-time audiovisual representation of a room and its occupants—represented by animated avatars—who we see moving around and talking. Many of these demonstrators are gathered in the showroom of the institute: what better way to show business leaders, school children, or members of the public the fruits of the research conducted at Idiap.

A Team Capable of Meeting the Most Diverse Demands

The development team's skills are relevant to every one of the institute's research themes. "We're able to meet the demands of all our researchers, whatever their project." From the creation of a database to the realization of an industrial product, there is no end to the requests that the group receives. These requests often come from Idiap start-ups (Klewel, Koemei, KeyLemon, recapp IT AG, etc.), but can also come from other companies in the form of calls for scientific collaboration with the institute. "Our role is obviously not to compete with local IT businesses, but when a project is highly specialized we're often the only ones with the necessary skills. Keep in mind that Idiap helps showcase research findings and develop collaborations with industry. But it never commercializes those findings itself."

Among the various projects that the team is currently working on, one comes from the French-speaking regional rehabilitation clinic (la Clinique romande de réadaptation—CRR / Suva). "The idea is to provide doctors with digital voice recorders, equipped with a voice recognition system adapted to specific medical vocabulary and able to handle the different accents encountered." Another project—running in collaboration with the Valais-based company ES Concept—aims to adapt, in real-time, advertisements displayed on billboards at sports events to the countries where those events are broadcast on TV.

Competence and Enthusiasm

In the team, the task of submitting industrial projects has—since 2014—fallen to development engineer Dr. Florent Monay. In collaboration with the researchers, Monay is responsible for drafting each proposal, and for its technical management once the project has been accepted. In this way, his task is to determine what solutions should be used to best meet the demands of industry.

Finally, two new developers joined the group in the last year (see box). "The members of our team are as competent as they are passionate," says Bornet. "And they are constantly improving how the team performs."

Hugues Salamin
33. Swiss. Lives in Dorénaz.
Idiap development engineer since 1 June 2014.
Specialist in artificial intelligence and automatic statistical learning.

"I already knew of Idiap as it was where I started my PhD in 2007, after my Master's at the Zurich polytechnic school. In 2009, I went—along with my supervisor (Alessandro Vinciarelli, now a professor at the University of Glasgow)—to Scotland, where I completed my doctorate, followed by a postdoc. The moment an opportunity at Idiap presented itself, I applied; the opportunity to come back "home" was too good to miss. I'm delighted to be here, especially as the work here is very interesting."

Frédéric Dubouchet
38. Swiss. Lives in Lens.
Idiap development engineer since 2 August 2014.
Specialist in graphics processing unit (GPU) computation.

"After a Bachelor's degree from the engineering college in Geneva, I worked at CERN for around 10 years while doing my Master's in parallel. I confess that I didn't know of Idiap before a friend told me that the institute was looking for development engineers. As I was dreaming of a move to Valais, particularly because I'm a keen skier, I applied straight away."

FACES

SYLVAIN CALINON, RESEARCHER

SYLVAIN—ROBOTS IN THE BLOOD

Sylvain Calinon heads the new Idiap research group dedicated to robot learning and interaction. A meeting with a researcher and "robot whisperer".

Since joining Idiap in spring 2014, Sylvain Calinon has shared his office with a singular individual. The size of a man, with a strong torso and long arms capable of supple, fluid motion; eyes moving behind its glasses and a quivering handlebar moustache.

Baxter—for that's the individual's name—is a robot. And he is the center of attention and the object of all Sylvain's care and research. The face of the young Swiss, trained at EPFL, lights up when he talks about Baxter. Never missing an opportunity to put his robotic charge to the test, he gives Baxter an object to grasp and begins to explain how the robot works.

His Profession—A Daily Passion
At 34, following five years at the Istituto Italiano di Tecnologia (IIT) in Genoa, Sylvain—a native of Yverdon—joined Idiap. Upon his arrival, the institute created a new research group: Robot Learning & Interaction (see page 9). "Given Idiap's competencies, I'm looking forward to seeing how many of them I can apply to the robot. Vision, motion, machine learning, verbal and non-verbal interaction: all of these can contribute to a robotics project. Moreover, thanks to the teams responsible for project management, development, and technology transfer, we have here an ideal environment in which to conduct our research projects."

Robotics is Sylvain's principal interest, his profession, and his hobby: a passion that infuses his daily life to the point that he talks more willingly of his chosen field than of himself. "From the beginning of my studies, it was clear that this is what I would do." A vocation established even earlier perhaps, as he recently found—among his childhood books—a volume entitled *The Robots.*

European Project in the Underwater Milieu
The research group Robot Learning & Interaction, directed by Sylvain Calinon, has already been awarded a major European project—DexROV, which begins in 2015 and will run for three and a half years. The project aims to create remote-controlled underwater robots, equipped with articulated arms, for the inspection and repair of pipelines and for oceanography, geology, and underwater archeology activities.
DexROV, which brings together seven academic and industrial partners, represents—for Idiap—incoming investment of nearly one million Swiss francs.

Science Fiction—An Unexpected Source of Inspiration
A robot like Baxter, capable of memorizing actions and reproducing them while adapting to its environment, immediately brings back memories of robotic characters from the movies. "The work of scientists is a source of inspiration for science fiction writers, and vice versa. We need to work faster—people are expecting to encounter in reality what they've seen on the big screen."

The development of a robot and its behavior requires a great deal of collaboration between scientists. Sylvain, who adores this aspect of his work, sometimes takes time out during his business trips to watch movies about robotics and artificial intelligence. For Sylvain, our vision of the robot is cultural: at times destructive, at times the savior of mankind. It goes without saying to which category belong the robots that he develops "In SMEs, they'll do uninteresting or dangerous work. They could also contribute to tasks in direct collaboration with people, as well as to rehabilitation, as exoskeletons attached to the body. Will they contribute to making the world a better place? I hope so. And, rest assured, we'll always need people. Who else would repair robots and teach them new tasks?"

A PASSION FOR PEOPLE

After ten years in the United States, returning to Switzerland Michael Liebling chose Idiap. The physicist, passionate about teaching, will head the institute's new research group devoted to computational bio-imaging.

From California straight to Switzerland. Michael Liebling had almost forgotten how harsh the winters of his homeland could be. A troublesome point when all his belongings were on a slow cargo ship somewhere on the ocean between the US and Europe.

But it would take more than that to make a hole in Michael's good humor, delighted as he was to join Idiap after over a decade stateside; an adventure that took him from the California Institute of Technology (Caltech) to the University of California at Santa Barbara (UCSB). "I was supposed to stay a year," Michael smiles. He left for the United States after obtaining his science doctorate from EPFL, supported in his postdoc studies by a young researcher's grant from the Swiss National Science Foundation. That year-long trip would last a decade.

Seduced by the Dynamism of Idiap

It's the summer of 2013, and Michael takes advantage of his vacation in Switzerland to visit Idiap. "I had always kept an eye on what was going on there, and was well aware of the institute's level of excellence. During my visit, I was seduced by Idiap's extraordinary dynamism. I also liked the idea that the institute not only leads large international projects, but also projects that have a local impact, like those relating to Valais' parliamentary proceedings or the bicentenary of the canton joining the confederation." Michael, 38, will form and lead the institute's new computational bio-imaging group (see page 9). "The challenge is to find the right people. I know they're out there. I just need the intuition to pick them out."

Fascinated by the computerized reconstruction of images, but also by medicine—he almost embarked on medical studies after high school—at UCSB the young physicist built up an interdisciplinary laboratory integrating biology. "I remember the very first meeting of my postdoc at Caltech: the biologists were talking among themselves; I listened to them, noting the words they used regularly. I was building lists of terms to learn."

Combining Research and Teaching

That California lab, which develops cell-imaging tools used especially for monitoring the embryonic development of the heart, is emblematic of the "Liebling method": building bridges between disciplines, empowering students, and passing on his knowledge in a manner both rigorous and inventive. If necessary, at the end of a demonstration, make something out of gelatin and food coloring? Michael didn't hesitate. When one feels, as strongly as Michael Liebling does, that "science is exciting", if a method works, you use it. "I asked the students themselves to create the images that we were going to use and to keep the biological samples alive. I think the more control we have over how we do things, the more effective we are."

His approach earned him the *Excellence in Teaching* award, conferred annually on the campus by the American conglomerate Northrop Grumman. "Nice recognition," commented Michael modestly, ever quicker highlighting the qualities of others than his own. Happily combining research and teaching, he admits to having spent a lot of time preparing his lectures. "It's very exciting: student's expectations are so varied. I combined a range of approaches so that everyone could get something out of them."

Michael's sharing approach seems second nature. "Collaborating with others is one of the nicest things one can do. I think I've never given a specialist conference presentation: I want my lectures to be as accessible to an engineer as they are to a biologist or a physicist."

Using fluorescence microscopy, Michael Liebling uses filters similar to those in this color chart to isolate the light coming from each type of tissue or cell, which have been previously marked with different dyes.

DANIEL GATICA-PEREZ, NEW EPFL PROFESSOR

"TEACHING IS A NATURAL EXTENSION OF RESEARCH"

Head of Idiap's Social Computing research group, Daniel Gatica-Perez also teaches IT Social Media at EPFL, where he has recently been appointed professor—a recognition shared.

Congratulations on your recent appointment as an EPFL professor. What does this change in status mean to you?
Even if the title of professor is bestowed on a personal level, I don't take it as such. For me, this appointment is a recognition of the years of work carried out by the Social Computing group that I lead. It is, therefore, also a recognition of the scientific importance of this field of research. More broadly speaking, it also recognizes the value of Idiap. I'm hopeful that the increased visibility that this appointment affords us will strengthen our existing synergies with other institutes and researchers, and even create new ones.

You have been teaching for almost ten years, alongside your research career. Is it easy to wear these two hats?
I see teaching as a natural extension of research. So—yes, I have no difficulty in conducting both these activities in parallel. Furthermore, they enrich each other. I learn a lot from the contact I have with the doctoral students who take my classes at EPFL. What I get out, I share with my research group and vice versa. The two activities are perfectly complementary.

Your research lies at the crossroads of computing and the social sciences. Is interdisciplinarity important to you?
Absolutely. A connection with other disciplines is central to my vision of research. I favor a transversal approach. I like the idea of tapping into other domains to find the answers to the questions we pose ourselves.

How, specifically, is this reflected in your research?
Our "SenseCityVity" project is a good example that combines urban planning, sociology, psychology, and social media. The aim is to mobilize citizens as factors of social change through the use of smartphones to help people understand socio-urban problems in certain cities and communities. We lead the project in collaboration with Mexican researchers, and with the support of CODEV—EPFL's Cooperation & Development Center.

You never hesitate to rely on communities, such as young people or students, for the production of research data. What are the strengths of such "crowdsourcing"?
From a scientific standpoint, it enables us to have a very large database, much richer than if we used traditional survey metho-dologies. The diversity of the data produced also guarantees us

a more representative result. But, beyond this, it's a question of involving citizens in a project by raising their awareness of the concrete benefits they can derive from it. Upstream, con-textualizing the process is also crucial if we are to identify, as accurately as possible, the needs of the communities we are reaching out to.

Using your research as a lever for improving the well-being of communities is a constant factor in your work?
I place the human being at the center of my research, which is also one of the fundamental principles of Idiap. My projects always aim to produce a positive and concrete impact on people and their quality of life, whether it be at a local, a cantonal, or a national level. This is part of my responsibility as a researcher.

What are your principal projects for the future?
The SenseCityVity project is generating some very interesting results. At the end of the pilot phase, conducted in Mexico, I would like to roll the project out in Valais. And—once again in the spirit of mobilizing citizens—it's very close to my heart to develop new crowdsourcing projects in Latin America or in developing countries, such as India or Kenya.

MY INTERNSHIP AT IDIAP MADE UP MY MIND

For over three months, Matthieu Mabillard swapped his position as an IT apprentice at Migros Valais for that of an intern at Idiap. A look back at how enriching the experience was and at the path traveled by this young native of Valais with his passion for politics and psychology.

First computer specialist hired by Migros Valais and the first intern to link up with Idiap's development team—Matthieu Mabillard enjoys a challenge. And that agrees with him rather well. At 19, he has an open mind but clear ideas. "I'll complete my apprenticeship in July. Then I'll complement it with a technical qualification one-year before starting a Bachelor's HES degree in computer science, most probably in Yverdon."

Computing—Of Course

Matthieu Mabillard caught the computing bug early and quickly developed a real interest in how programs operate. More than technology, it is mainly computing and software that feed his curiosity. "At the age of seven or eight, I was already trying to understand how the family computer worked. How to change a password, how to set options for folders… In a way, to understand what was going on behind the scenes." When the crucial moment came to choose his career, Matthieu had no hesitation. With his love of mathematics and logic, two options were open to him: follow his father into the world of finance, or pursue a career in IT. The second option won. And there are no regrets. "Whatever I chose, I'd have stayed in touch with figures. The only difference being that the *1*s and the *0*s don't have the same meaning from one field to the other."

Self-Taught

Matthieu, in the final year of his apprenticeship, had had little opportunity to flex his programming muscles—despite his passion for the discipline—during his time with Migros, where his daily schedule and that of his four colleagues runs to the tempo of maintaining the company's IT equipment (600 workstations in Valais) and providing user support.

The internship at Idiap came as a breath of fresh air, and Matthieu seized—again without hesitation—this unique opportunity to rub shoulders with the world of research and development to which he aspires. The atmosphere into which he plunged for three and a half months immersed him in the workings of an internationally renowned research institute, where scientists and researchers from throughout the world work side by side. "I was more used to asking people which valley they came from. At Idiap I had to broaden my horizons and think in terms of countries or even continents."

During the internship, Matthieu was actively involved in the creation of the Valais*Wallis Digital application for Android-based devices—a boon for a young man who describes himself as a self-taught programmer. "Without the opportunity to practice programming regularly in my job, I had to learn about the subject mainly by myself, topping up the basic information I'd been taught in courses. I deepened my understanding of C++ and Java by a lot of reading and amused myself by creating some fun programs at home, like a vocabulary revision program or a remote control application connecting my PC and my smartphone."

A Passion for Politics

Although his days are largely occupied by computing, Matthieu still does not forget to stay curious and open to the world around him. Interested in psychology and communication, as well as in the law, it is to politics that he devotes much of his free time. Not so much as a committed activist, more as an attentive observer. "I refuse to be labeled as a member of one party or another. What interests me are the debates and the discussions, whether on television or in the press." As in computing, Matthieu is trying to understand the hidden workings—exploring what lies behind the interviews, screen presence, and appearances.
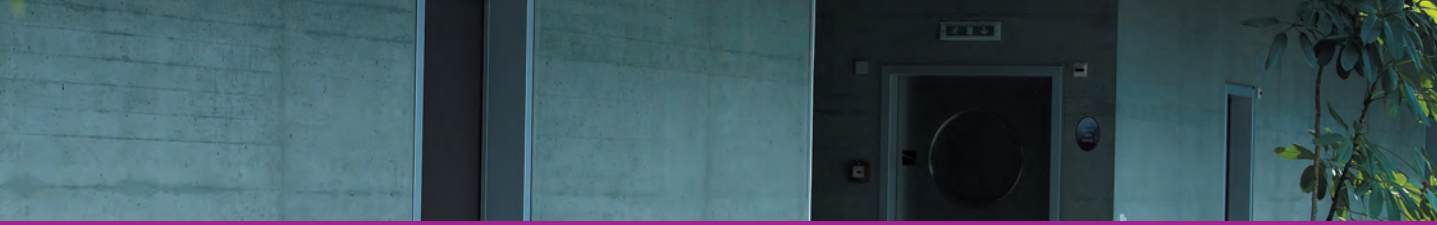
JOINING AND LEAVING

In 2014, seventeen members of staff left and eighteen joined the team. The new arrivals include two researchers, two development engineers, seven research assistants, and one industrial relations collaborator.

JOINING US IN 2014

First name and last name, position, country of origin

Sergio Calabretta, industrial relations, Italy

Sylvain Calinon, researcher, Switzerland

Tatjana Chavdarova, research assistant, Macedonia

Tiago De Freitas, research assistant, Brazil

Subhadeep Dey, research assistant, India

Frédéric Dubouchet, development engineer, Switzerland

Charles Dubout, scientific collaborator, Switzerland

Sucheta Gosh, postdoc, India

Ivana Humawan, postdoc, Australia

Michael Liebling, senior researcher, Switzerland

Ngoc Quang Luong, postdoc, Vietnam

Parvaz Mahdabi, postdoc, Iran

Skanda Muraldihar, research assistant, India

Xiao Pu, research assistant, China

Dhananjay, Ram, research assistant, India

Mario Rosario, engineer, Portugal

Hughes Salamin, development engineer, Switzerland

Pedro Tome, postdoc, Spain

MOVING ON IN 2014

First name and last name, position, country of origin, joined Idiap

Sara Bahaadini Beigy, research assistant, Iran, 2013

Chidansh Amitkumar Bhatt, postdoc, India, 2012

Ronan Collobert, researcher, France, 2010

Charles Dubout, scientific collaborator, France, 2014

Nesli Erdogamus, postdoc, Turkey, 2012

Marco Fornoni, research assistant, Italy, 2009

Paul Gay, research assistant, France, 2012

Mohammad Javad, Taghizadeh, research assistant, Iran, 2010

Leonidas Lefakis, research assistant, Greece, 2010

Thomas Meyer, research assistant, Switzerland, 2010

Gelareh Mohammadi, postdoc, Iran, 2009

François Moulin, development engineer, Switzerland, 2009

Francisco Pinto, postdoc, Portugal, 2013

Yann Rodriguez, industrial relations, Switzerland, 2002

Mario Rosario, engineer, Portugal, 2014

Samira Sheikhi, research assistant, Iran, 2010

Gyorgy Jozsef Szaszak, postdoc, Hungary, 2012

AWARDS

Internal

Each year Idiap awards prizes to two of its research assistants. The first award is for research, the second for papers published. Candidates for the Research Prize are evaluated, by an internal commission, against five criteria: papers published, teamwork, commitment to a project, communication skills, and autonomy. For the Publication Prize, the institute's leadership draws up an initial list of candidates from papers of which an Idiap research assistant is the primary author. The members of the Advisory Board then evaluate, independently and anonymously, the papers selected.

In 2014 the Research Prize was awarded to **Kenneth Funes Mora** for his excellence in research on gaze estimation, contribution to data and software release, and animation of the Idiap and EDEE/EPFL PhD community. The Research Publication Prize went to **Laurent Son Nguyen** for his outstanding scientific article entitled "Hire Me: Computational inference of hirability in employment interviews based on nonverbal behavior".

External

This year, Idiap wishes to acknowledge its researchers' brilliant participation at international conferences. The quality of their research was rewarded with several distinctions.

Alexandre Heili and Jean-Marc Odobez

Best Paper Award at the IEEE Int. Workshop on Performance Evaluation of Tracking Systems (PETS), January 2014
Parameter Estimation and Contextual Adaptation for a Multi-Object Tracking CRF Model

L. El Shafey, E. Khoury and S. Marcel

BTAS Best Paper Award at the International Joint Conference on Biometrics (IJCB) 2014
Audiovisual gender recognition in uncontrolled environment using variability modeling techniques

Sylvain Calinon

Excellence in Reviewing Award, Robotics and Autonomous Systems (Elsevier), 2014

Kenneth Funes Mora

Laurent Son Nguyen

THESES COMPLETED

Six students completed their theses in 2014: Laurent El Shafey, Marco Fornoni, Alexandre Heili, Thomas Meyer, Ramya Rasipuram, and Samira Sheikhi.

- **Scalable Probabilistic Models for Face and Speaker Recognition**
  Laurent El Shafey, 24 March 2014
  Thesis directors: Prof. Hervé Bourlard and Dr. Sébastien Marcel
  Members of the thesis committee: Prof. Jean-Philippe Thiran, Prof. Josef Kittler, Dr. Jan Černocký, and Prof. Pascal Fua

- **Saliency-Based Representations and Multi-component Classifiers for Visual Scene Recognition**
  Marco Fornoni, 26 September 2014
  Thesis directors: Prof. Hervé Bourlard and Prof. Barbara Caputo
  Members of the thesis committee: Prof. Colin Jones, Prof. Jean-Philippe Thiran, Prof. Vittoria Murino, and Prof. Danijel Skočay

- **Human Tracking and Pose Estimation in Open Spaces**
  Alexandre Heili, 20 June 2014
  Thesis director: Dr. Jean-Marc Odobez
  Members of the thesis committee: Prof. Colin Jones, Dr. François Fleuret, Dr. Patrick Pérez, and Dr. Tao Xiang

- **Discourse-Level Features for Statistical Machine Translation**
  Thomas Meyer, 8 December 2014
  Thesis directors: Prof. Hervé Bourlard and Dr. Andrei Popescu-Belis
  Members of the thesis committee: Prof. Jean-Philippe Thiran, Dr. Martin Rajman, Prof. Bonnie Webber, and Prof. Martin Volk

- **Grapheme-Based Automatic Speech Recognition using Probabilistic Lexical Modeling**
  Ramya Rasipuram, 1 October 2014
  Thesis directors: Prof. H. Bourlard and Dr. M. Magimai Doss
  Members of the thesis committee: Prof. J.-Ph. Thiran, Dr. K. Knill, Prof. S. Renals, and Dr. J.-M. Vesin

- **Inferring Visual Attention and Addressee in Human Robot Interaction**
  Samira Sheikhi, 11 July 2014
  Thesis director: Dr. Jean-Marc Odobez
  Members of the thesis committee: Prof. David Atienza Alonso, Dr. Daniel Gatica-Perez, Prof. Britta Werde, and Prof. Frederic Lerasle

OPERATING ACCOUNTS

(Swiss francs)	2013	2014	%
INCOME			
City of Martigny	700,000	700,000	6.72%
Canton of Valais	1,620,000	1,720,000	16.52%
Swiss Confederation	2,400,000	2,550,000	24.49%
Loterie Romande	250,000	50,000	0.48%
NCCR IM2 projects	1,273,628	476,648	4.58%
Swiss National Science Foundation projects	1,489,392	1,427,889	13.71%
Hasler Foundation projects	314,717	371,045	3.56%
European Commission projects	1,132,694	1,210,890	11.63%
The Ark projects	189,288	173,797	1.67%
CTI projects	328,864	607,217	5.83%
EPFL contribution	72,000	72,000	0.69%
Industry financing	473,367	366,736	3.52%
Other income / extraordinary income	394,213	685,926	6.59%
TOTAL INCOME	**10,638,162**	**10,412,148**	**100.00%**
EXPENSES			
Personnel expenses	7,709,161	7,879,097	75.7%
Education and travel	458,200	355,766	3.4%
Third-party expenses	93,062	302,540	2.9%
Computer equipment and maintenance	241,310	246,284	2.4%
Administrative costs	193,520	233,001	2.2%
Promotion and communication	60,413	50,079	0.5%
Rent	817,621	818,888	7.9%
Depreciation	184,563	310,853	3.0%
Exceptional expenses	97,881	0	0.0%
Provisions	750,000	200,000	1.9%
TOTAL EXPENSES	**10,605,731**	**10,396,508**	**99.8%**
OPERATING PROFIT / LOSS	**32,431**	**15,640**	**0.2%**

Distribution of sources of financing
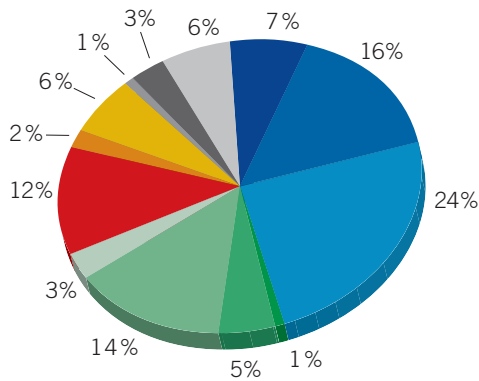
- City of Martigny
- Canton of Valais
- Swiss Confederation
- Loterie Romande
- NCCR IM2 projects
- Swiss National Science Foundation projects
- Hasler Foundation projects
- European Commission projects
- The Ark projects
- CTI projects
- EPFL contribuiton
- Industry financing
- Other income / extraordinary income

Distribution of costs
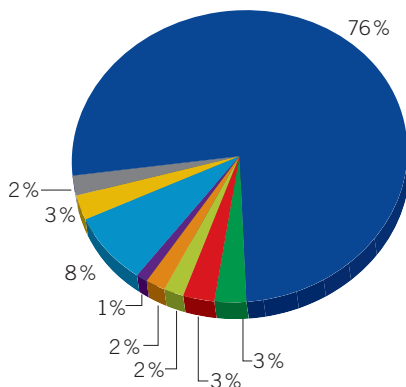
- Personnel expenses
- Education and travel
- Third-party expenses
- Computer equipment and maintenance
- Administrative costs
- Promotion and communication
- Rent
- Depreciation
- Provisions

Management comments

It is a pleasure to be able to note that, for the third consecutive year, Idiap has managed a budget of over 10 million Swiss francs.

Also, the proportion of public investment has once again met the strategic objective, set by the institute's management, of less than 50 percent.

The success rate with regards to European projects is, however, lower than in the past. This can be explained by the result of the federal referendum of 9 February 2014, which continues to generate uncertainty regarding the position of Swiss research in the European H2020 program, and a certain ostracism of Swiss partners.

Staff costs (salaries + social charges) make up the bulk of expenditures, at around 80 percent.

Idiap closes its 2014 financial statements with a surplus of CHF 15,640.

Federal, cantonal, and municipal subsidies

(In thousands of Swiss francs)

YEAR	2012	2013	2014	2015*
Confederation	2,393	2,400	2,550	2,495
Canton	1,120	1,320	1,720	1,720
Municipality	650	700	700	700

* Budget

BALANCE SHEET

(Swiss francs)	31.12.2013	31.12.2014
ASSETS		
Cash	3,614,085.58	3,172,241.55
Accounty receivable	220,492.98	315,159.11
Accrued income and other	985,382.87	502,413.88
TOTAL CURRENT ASSETS	**4,819,961.43**	**3 989,814.54**
Equipment	340,658.79	515,818.64
Financial assets	10,000.00	10,000.00
TOTAL NON-CURRENT ASSETS	**350,658.79**	**525,818.64**
TOTAL ASSETS	**5,170,620.22**	**4,515,633.18**
LIABILITIES		
Accounts payable	307,725.05	300,607.00
Accrued expenses	2,921,723.22	2,058,214.07
Provisions	750,000.00	950,000.00
TOTAL FOREIGN FUNDS	**3,979,448.27**	**3,308,821.07**
Share capital	40,000.00	40,000.00
Special reserve	1,000,000.00	1,000,000.00
Retained earnings	118,739.57	151,171.95
Net income	32,432.38	15,640.16
TOTAL OWN FUNDS	**1,191,171.95**	**1,206,812.11**
TOTAL LIABILITIES	**5,170,620.22**	**4,515,633.18**

ORGANIZATION

FOUNDATION COUNCIL

SCIENTIFIC ADVISORY BOARD

DIRECTOR
Hervé Bourlard

DEPUTY DIRECTOR
François Foglia

SCIENTIFIC COLLEGE
F. Fleuret, D. Gatica-Perez, J.-M. Odobez

ADMIN COMMITTEE

RESEARCH COMMITTEE

Admin

Services

FINANCE & HUMAN RESOURCES
Ed Gregg

INDUSTRIAL RELATIONS
Sergio Calabretta

SYSTEM & INFRASTRUCTURE
Frank Formaz

DEVELOPMENT ENGINEERS
Olivier Bornet

PROGRAM MANAGERS
Léonore Miauton

SPEECH & AUDIO PROCESSING
Hervé Bourlard
Phil Garner
Mathew Magimai Doss
Petr Motlicek

COMPUTER VISION & LEARNING
François Fleuret

SOCIAL COMPUTING
Daniel Gatica-Perez

PERCEPTION & ACTIVITY UNDERSTANDING
Jean-Marc Odobez

ROBOT LEARNING & INTERACTION
Sylvain Calinon

ARTIFICIAL COGNITIVE SYSTEMS
Barbara Caputo

APPLIED MACHINE LEARNING
Ronan Collobert

COMPUTATIONAL BIO IMAGING
Michael Liebling

BIOMETRIC PERSON RECOGNITION
Sébastien Marcel

NATURAL LANGUAGE PROCESSING
Andrei Popescu-Belis

EMPLOYEES

Oya Aran Karakus, Research Fellow, Turkey, 2009

Afsaneh Asaei, Postdoc, Iran, 2008

Sara Bahaadini Beigy, Research Assistant, Iran, 2013

Chindansh Bhatt, Postdoc, India, 2012

Joan Isaac Biel, Research Assistant, Spain, 2008

Hervé Bourlard, Director, Switzerland, Saxon, 1996

Sylvain Calinon, Researcher, Switzerland, Gimmel, 2014

Gulcan Can, Research Assistant, Turkey, 2013

Olivier Canévet, Research Assistant, France, 2012

Barbara Caputo, Senior Researcher, Italy, 2005

Tatjana Chavdarova, Research Assistant, Macedonia, 2014

Ivana Chingovska, Research Assistant, Macedonia, 2011

Ronan Collobert, Researcher, France, 2010

Tiago De Freitas, Research Assistant, Brazil, 2014

Subhadeep Dey, Research Assistant, India, 2014

Pranay Dighe, Research Assistant, India, 2013

Trinh-Minh-Tri Do, Postdoc, Vietnam, 2009

Charles Dubout, Scientific Collaborator, France, 2014

Elie El Khoury, Postdoc, Lebanon, 2011

Laurent El Shafey, Research Assistant, France, 2010

Nesli Erdogmus, Postdoc, Turkey, 2012

Marc Farras Font, Postdoc, Spain, 2011

François Fleuret, Senior Researcher, France, 2007

Marco Fornoni, Research Assistant, Italy, 2010

Kenneth Funes Mora, Research Assistant, Costa Rica, 2011

Philip Garner, Senior Reseacher, England, 2007

Daniel Gatica-Perez, Senior Researcher, Mexico, 2002

Paul Gay Research Assistant, France, 2012

Sucheta Ghosh, Postdoc, India, 2014

Manuel Günther, Postdoc, Germany, 2012

Maryam Habibi, Research Assistant, Iran, 2011

Alexandre Heili, Research Assistant, France, 2010

Ivana Himawan, Postdoc, Australia, 2014

Pierre-Edouard Honnet, Research Assistant, France, 2012

Rui Hu, Postdoc, China, 2013

David Imseng, Postdoc, Switzerland, St. German, 2009

Cijo Jose, Research Assistant, India, 2013

Ilja Kuzborskij, Research Assistant, Lithuania, 2012

Alexandros Lazaridis, Postdoc, Greece, 2012

Rémi Lebret, Research Assistant, France, 2012

Leonidas Lefakis, Research Assistant, Greece, 2010

Joël Legrand, Research Assistant, France, 2012

Michael Liebling, Senior Researcher, Switzerland, St Sulpice, 2014

Ngoc Quand Luong, Postdoc, Vietnam, 2014

Srikanth Madikeri, Postdoc, India, 2013

Mathew Magimai Doss, Researcher, India, 2007

Parvaz Mahdabi, Postdoc, Iran, 2014

Sébastien Marcel, Senior Researcher, France, 2000

Thomas Meyer, Research Assistant, Switzerland, Martigny, 2010

Gelareh Mohammadi, Research Assistant, Iran, 2009

Petr Motlicek, Researcher, Czech Republic, 2005

Skanda Muralidhar, Research Assistant, India, 2014

James Newling, Research Assistant, England, 2013

Laurent Nguyen, Research Assistant, Switzerland, Lausanne, 2011

Jean-Marc Odobez, Senior Researcher, France/Switzerland, Clarens, 2001

Pedro Oliveira Pinheiro, Postdoc, Brazil, 2012

Dimitry Palaz, Research Assistant, Switzerland, Martigny, 2011

Nikolaos Pappas, Research Assistant, Greece, 2012

Novi Patricia, Research Assistant, Indonesia, 2012

Francisco Pinto, Postdoc, Portugal, 2013

Andrei Popescu-Belis, Senior Researcher, France/Romania, 2007

Blaise Potard, Postdoc, France, 2013

Xiao Pu, Research Assistant, China, 2014

André Rabello Dos Anjos, Postdoc, Brazil, 2010

Dhananjay Ram, Research Assistant, India, 2014

Ramya Rasipuram, Research Assistant, India, 2010

Marzieh Razawi, Research Assistant, Iran, 2013

Lakshmi Saheer, Research Assistant, India, 2008

Darshan Santini, Research Assistant, India, 2012

Ashtosh Sapru, Research Assistant, India, 2011

Samira Sheikhi, Research Assistant, Iran, 2010

Gyorgy Szaszak, Postdoc, Hungary, 2012

Mohammad Javad Taghizadeh, Research Assistant, Iran, 2011

Pedro Tome, Postdoc, Spain, 2014

Raphaël Ullman, Research Assistant, Switzerland, Lausanne, 2012

Matthias Vanoni, Research Assistant, France, 2013

Sree Harsha Yella, Research Assistant, India, 2010

Development Engineers

Philip Abbet, Senior Development Engineer, Switzerland, Conthey, 2006

Olivier Bornet, Senior Development Engineer, Switzerland, Pont de la Morge, 2004

Milos Cernak, Senior Development Engineer, Slovakia, 2011

Frederic Dubouchet, Development Engineer, Switzerland, Lens, 2014

Samuel Gaist, Development Engineer, Switzerland, Ardon, 2013

Salim Kayal, Development Engineer, Switzerland, Vevey, 2011

Vasil Khalidov, Senior Development Engineer, Russia, 2010

Christine Marcel, Development Engineer, France, 2007

Florent Monay, Development Engineer, Switzerland, Choëx, 2008

François Moulin, Development Engineer, Switzerland, Vollèges, 2009

Alexandre Nanchen, Senior Development Engineer, Switzerland, Martigny, 2008

Mario Rosario, Engineer, Portugal, 2014

Hugues Salamin, Development Engineer, Switzerland, Dorénaz, 2014

Flavio Tarsetti, Senior Development Engineer, Switzerland, Martigny, 2008

Administrative Staff

First name and last name, position, country of origin, (place of residence), joined Idiap

Sergio Calabretta, Industrial Relations, Italy, Martigny, 2014
Antoine Dorsaz, Financial Assistant, Switzerland, Fully, 2012
Christophe Ecoeur, Program Manager, Switzerland, Collombey, 2010
Fellay Martina, Program Manager, Austria, 2012
François Foglia, Deputy Director, Switzerland, Saxon, 2006
Edward-Lee Gregg, Financial Manager, United States of America, 2004
Léonore Miauton, Program Manager Head, Switzerland, Chexbres, 2012
Sylvie Millius, Administrative Assistant, Switzerland, Vétroz, 1996
Yann Rodriguez, Industrial Relations, Suisse, Vollèges, 2006
Nadine Rousseau, Administrative Assistant, Switzerland, Saxon, 1998

System Engineers

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
Louis-Marie Plumel, System Administrator, France, 2011
Vincent Spano, Webmaster, Switzerland, Martigny-Combe, 2004

Interns

First name and last name, country of origin, home institution

Idiap interns generally spend between six and ten months at the research institute. Some are students at the Ecole polytechnique fédérale de Lausanne (EPFL) and spend time at Idiap as part of their degree studies. Others come on student exchange programs set up with European projects in which Idiap participates.

Krysten Ansermoz, Switzerland, EPFL, Lausanne
Umut Avci, Turkey, University of Trento (ITA)
Julia Biolley-Franzetti, Switzerland, University of Lausanne
Lisa Bornet, Switzerland, HES-SO, Sierre
Maxime Bourlard, Switzerland, University of Neuchâtel
Fabian Brix, Germany, EPFL, Lausanne
Aleksandra Cerekovic, Croatia, University of Zagreb (HRV)
Gilberto Chavez Martinez, Mexico, CIMAT (MEX)
Joane Cheseaux, Switzerland, University of Geneva
Lukas De Oliveira, Switzerland, EPFL, Lausanne
Ailbhe Finnerty, Ireland, University of Trento (ITA)
Manon Fournier, Switzerland, University of Geneva
Catherine Gasnier, France, EPFL, Lausanne
Branislav, Gerazov, Macedonia, University of Skopje, (MKD)
Serife Kucur, Turkey, EPFL, Lausanne
Cesar Laurent, Switzerland, EPFL, Lausanne

Matthieu Mabillard, Switzerland, Migros Valais
Navid Mahmoudian, Iran, University Amvikabir (IRN)
Alvaro Marcos, Spain, University of Alcala (ESP)
Lesly Miculicich, Peru, University of Fribourg
Loris Millius, Switzerland, Collège La Planta, Sion
Thanh Trung Phan, Vietnam, Telecom, Paris (FRA)
Guillem Quer, Switzerland, Computer Programmer Independent, Barcelona (ESP)
Sandrine Revaz, Switzerland, University of Fribourg
Dairazalia Sanchez, Mexico, University Ensenada of Baja California (MEX)
Alexandre Sierro, Switzerland, HES-SO
Kelly Tiraboschi, Italy, University of Lausanne
José Velasco, Spain, University of Alcala (ESP)
David Yambay, USA, University of Clarkson, NY (USA)

Visitors

First name and last name, country of origin, home institution

Visitors are researchers or industry representatives who only spend a few days or weeks at the institute, some to strengthen inter-institutional relationships and others to get an insight into the work carried out at Idiap.

Srinivas Narayanan, Invited Professor, University of California, Berkeley, CA (USA)
Pierre Lison, Belgium, University of Oslo (NOR)
Shogo Okada, Japan, Institute of Technology, Tokyo (JPN)
Salvador Ruiz Correa, Mexico, CIMAT (MEX)

FOUNDATION COUNCIL 2014

The Foundation council is responsible for the economic and financial management of Idiap. It defines the institute's structure, appoints its director, and—more generally—ensures Idiap's development and defends its interests.
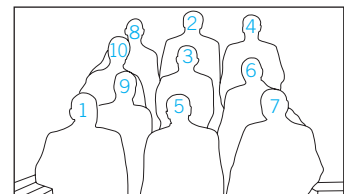
1 **Mr. Olivier Dumas**, President
Independent manager and business consultant

Mr. Jean-Daniel Antille, Vice-President *(not pictured)*
Manager of the Regional Office for the Economic Development of French-Speaking Valais

2 **Prof. Karl Aberer**, Vice-President
Vice-President for Information Systems,
Ecole polytechnique fédérale de Lausanne (EPFL)

3 **Mr. Marc-André Berclaz**
Operational Director of the EPFL Valais Wallis Outpost

4 **Mr. Stefan Bumann**
Head of the Service des Hautes Ecoles (SHE), State of Valais

Mr. Marc-Henri Favre *(not pictured)*
President of the city of Martigny

5 **Prof. Christian Pellegrini**
Senior Member of the Foundation Council of Idiap
Honorary Professor in the Faculty of Sciences
at the University of Geneva

6 **Mr. Jean-René Germanier**
Member of the National Council

7 **Mr. Patrick Furrer**
Vice-Chancellor Research and Innovation, HES-SO

8 **Mr. Jordi Montserrat**
Regional Manager Venturelab

Mr. Dominique Perruchoud *(not pictured)*
President of the Board of Cimark SA

Mr. Walter Steinlin *(not pictured)*
Swisscom University Relations
President of the Commission for Technology
and Innovation (CTI)

Not members of the Foundation Council

9 **Prof. Hervé Bourlard**
Director of Idiap

10 **Mr. François Foglia**
Deputy Director of Idiap

ADVISORY BOARD

The Advisory Board is comprised of members of the scientific community selected by Idiap's management for their exceptional skills and avant-garde vision. Although their role is strictly advisory, their counsel is frequently sought and proves to be invaluable when making decisions regarding research, training, and technology transfer.

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

Prof. Anil K. Jain
Distinguished Professor, Department of Computer Science
& Engineering
Michigan State University, USA

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

Prof. Kenji Mase
Professor, Graduate School of Information Science
Nagoya University, Japan

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

Prof. Klaus-Robert Müller
Professor of Computer Science, TU Berlin
Director, Bernstein Focus on Neurotechnology,
Berlin, Germany

Dr. David Nahamoo
Senior Manager, Human Language Technologies, IBM Research
Yorktown Heights, NY, USA

Prof. Gerhard Sagerer
Rector, University of Bielefeld,
Bielefeld. Germany

Prof. Bernt Schiele
Max-Planck-Director, MPI Informatics
Professor at Saarland University,
Saarbrücken, Germany

Prof. Bayya Yegnanarayana
Professor and Microsoft Chair, International Institute
of Information Technology
Hyderababd, India

CITY OF MARTIGNY

CANTON OF VALAIS

SWISS CONFEDERATION
The State Secretariat for Education, Research and Innovation (SERI)

www.loterie.ch

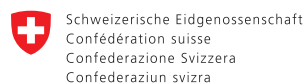**www.swisscom.com**

www.groupemutuel.ch

www.epfl.ch

www.theark.ch

www.ideark.ch

www.snf.ch

Innovation Promotion Agency CTI

www.kti.admin.ch

www.haslerstiftung.ch

SCIENTIFIC INSERTS

IDIAP RESEARCH AREAS

To face its continuous growth and diversification in different, but complementary, research directions, while building upon its general institutional theme of "Human and Media Computing" (as defining ourselves on our web site www.idiap.ch), Idiap's expertise mainly spans five research areas, which are summarized in Figure 1 below. In that table, each research area is also described in a few keywords. As described later, in 2014, those generic research areas were covered by 10 research groups.

Idiap Research Areas (horizontal)	Keywords
Perceptual and cognitive systems	Speech and audio processing, computer vision, handwriting recognition, document processing, computational cognitive sciences, robotics, natural language processing, machine translation
Human and social behavior	Web and mobile social media, social interaction sensing, social signal processing, verbal and nonverbal communication analysis, mobile phone sensing, computational social science
Information and presentation interfaces	Multimedia information systems, user interfaces, contextualization, personalization, system evaluation, mobile HCI using Big Data, data driven services
Biometrics	Speaker recognition, face recognition, multimodal biometric fusion, mobile biometry, spoofing and anti-Spoofing
Machine learning	Statistical and neural network based ML, computational efficiency, online learning, multi-sensor processing, very large datasets

Figure 1: Overview of Idiap research areas.

• **Perceptual and cognitive systems**: Speech processing; Natural language understanding and translation; Document and text processing; Vision and scene analysis; Multimodal processing; Cognitive sciences and Robotics.
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 (with a particular emphasis on human sensing), multimodal interaction, computational cognitive systems, robotics, and methods for automatically training such systems.

• **Human and social behaviour**: Social media; Mobile media; Social interaction analysis; Social signal processing.
This area spans methods for analyzing human and social behaviour from a variety of information sources with the goals of understanding social phenomena and developing human-centered applications. Idiap has investigated new approaches for behavioral analysis in face-to-face communication, online interaction in social media sites like YouTube, and smartphone-mediated interaction.

• **Information and presentation interfaces**: Multimedia information systems, User interfaces; System evaluation.
Information processing by computers must be accompanied by human-computer interfaces that present information and receive input in an efficient and usable way, possibly acquiring information from users in a non-disruptive way. Current research directions at Idiap focus on multimedia information systems, search and recommendation, and interactive information retrieval, and several new projects are currently being initiated in those directions.

• **Biometric person recognition**: Face recognition (detection-localization–identification-verification); Speaker identification and verification; Multimodal biometric person recognition; Counter-measures to spoofing attacks.
Biometric person recognition (Biometrics) 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, EEG, ECG, ear shape, body odour, body salinity, vascular). Idiap is currently a recognized leader in that field, fully exploiting our multi-disciplinary expertise in image processing, computer vision,

pattern recognition and machine learning. This expertise has recently been confirmed by the official launch of a 'Swiss Center for Biometric Security Research and Testing" (http://www.biometrics-center.ch/) hosted at Idiap.

- **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. Today, Idiap is also recognized as a key leader in that field with new trends towards "collaborative machine learning", deep neural network architectures, and large-scale distributed learning algorithms.

- **New research themes initiated in 2014:** In addition to these "horizontal" research themes, new activities recently started to be developed towards bio-medical applications (through the new Bioimaging Group) and Environmental Modeling (through the new group in "Uncertainty Quantification and Optimal DesignDecision Support").

1. Speech and Audio Processing

Key scientific outputs: We have always used both neural network and Markov model based approaches, leading to the KL (Kullback Leibler) HMM approach to fusion of neural network and HMM based speech recognition, especially for multilingual speech modelling. Use of techniques from HMM based speech recognition in HMM based speech synthesis resulted in a unified approach to speech recognition and synthesis. The group was well placed take advantage of recent advances in "deep" neural networks (DNNs) technology, and the associated Kaldi automatic speech recognition (ASR) toolkit. Several key contributions in microphone arrays, including ad-hoc microphone arrays, are well recognized. Sparse model based approaches have led to novel algorithms for source separation. The agglomerative clustering approach to speaker diarization is well regarded. Finally, the work in this group also resulted in a several spin-offs, including dev-audio, Koemei and, more recently (end of 2014), recapp.

Much of the group's recognition has come from contributions to the speech and audio community; these include the AMI corpus[1] and the Juicer ASR decoder[2]. More recently, the MediaParl[3] and Walliserdeutsch[4] databases have been released; other contributions such as ISS and SSP are on the GitHub site[5]. This highlights that, although the group has traditionally worked with English speech, many recent projects are multi-lingual in nature. This has in turn become a focus of the group; one on which we are well placed to capitalise given our geographical location. The group is also very active in contributing to the Kaldi open-source ASR toolkit[6].

Automatic Speech Recognition

In recent years, at Idiap the ASR research activities have been expanded from mono-lingual to cross-/multi-lingual processing. More specifically, in addition to focusing on "majority" languages other than English such as, French, German, Idiap is actively carrying research in several ASR directions, including:

- Robust parametrization and acoustic modeling.
  We are still investigating new features (e.g., posterior-based features) and new acoustic models (new forms of hidden Markov models, such as KL-HMM, or artificial neural networks) that are more robust to noise and acoustic environments, as well as to speaker variability (e.g., accented speech). Over the last 12 months, our research has largely focussed on the use of DNNs for far-field speech recognition; novel approaches to adapt DNNs across different domains and languages; development of articulatory feature (phonological feature) based ASR system; and development of a unified framework for recognition and detection (funded through the EC Eurostars project D-Box[7] , SNSF projects like FlexASR[8] and IM2 NCCR sub-project Walliserdeutsch, HASLER Foundation project AddG2SU[9] , Idiap internship project, CTI project ScoreL2[10] , Samsung Electronics Co., Ltd. (South Korea) sponsored project AMASSE[11])

- Cross-lingual and multi-lingual speech recognition.
  Within the scope of the D-Box project, we focus on investigating and exploiting fast acoustic model adaptation techniques in cross-lingual and multi-lingual scenarios. The resulting speech recogniser relies on a hybrid approach, where an artificial neural network acoustic model is bootstrapped using well-resourced data and adapted to the target language (or environment). In the context of an industry sponsored project (Samsung, Inc. South Korea sponsored project DAUM[12]), we are also investigating Markov model based acoustic model adaptation techniques (currently referred to as "subspace Gaussian modeling") for different domains (e.g., accent or dialect).

The recogniser is often required to perform in real-time and to support closed- and open-dictionary (i.e., detection of keywords) decoding graphs. Multi-lingual or domain adaptation is often combined with speaker-adaptation techniques to further improve recognition accuracies.

---

1 http://www.idiap.ch/mmm/corpora/ami

2 http://juicer.amiproject.org/juicer/

3 http://www.idiap.ch/dataset/mediaparl

4 http://www.idiap.ch/dataset/walliserdeutsch

5 https://github.com/idiap

6 http://kaldi.sourceforge.net

7 http://www.idiap.ch/scientific-research/projects/d-box-a-generic-dialog-box-for-multilingual-conversational-applications

8 http://www.idiap.ch/scientific-research/projects/flexible-grapheme-based-automatic-speech-recognition

9 http://www.idiap.ch/scientific-research/projects/flexible-acoustic-data-driven-grapheme-to-subword-unit-conversion

10 http://www.idiap.ch/scientific-research/projects/automatic-scoring-and-adaptive-pedagogy-for-oral-language-learning

11 http://www.idiap.ch/scientific-research/projects/amasse

12 http://www.idiap.ch/scientific-research/projects/domain-adaptation-using-sub-space-models

- Swiss-languages.

  We continuously improve our speech recognisers for Swiss German and Swiss French and also apply the most recent advances in speech technology employing deep neural networks. The recognisers are now also used in commercial products of the Idiap spinoff recapp IT AG. Work on speech recognition for Walliserdeutsch (illustrated in Figure 4), one of the most difficult to understand of the Swiss dialects, was published in 2014; the associated database is also available for download. The subject of Swiss dialects remains pertinent as an important local focus, with plans underway to expand to other regions.
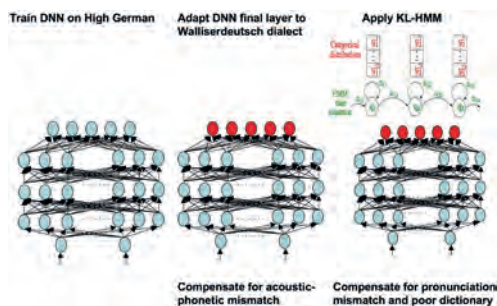

Figure 4: Illustration of development of Walliserdeutsch ASR system, merging Deep Neural Net- work (DNN) technology and a new form of Hidden Markov Model (KL-HMM).

- Template-based ASR.

  Idiap continues to work on template-based or exemplar-based ASR using posterior-based features in the context of projects like SNSF project A-MUSE[13] , SNSF project PHASER[14] and CTI project MultiVEO[15] . In particular, our current research is focussing on linking template-based speech recognition and statistical sequence model based speech recognition. Towards this end there are two on-going efforts. The first focuses on development of a "sparse" HMM framework by establishing a link between statistical speech recognition formalism and the recent theories of compressive sensing and sparse modeling. The framework integrates the advantages of template-based modelling within the HMM sequence matching scheme. The second effort focusses on development of an approach that seamlessly combines both statistical sequence model based ASR system and template-based ASR system into a single ASR system that can yield performance better than the best individual system.

- Grapheme-based ASR.

  Over the past five years, Idiap has been conducting research on the use of alternative sub- word units, specifically graphemes, and development of an ASR approach which is capable of handling both acoustic and lexical resource constraints by learning grapheme-to-phoneme relationship through the acoustic speech signal (funded through the SNSF-project FlexASR and HASLER Foundation project AddG2SU). In the past 12 months, our research activities in that direction have primarily focussed on exploiting the approach for lexical resource development through (a) acoustic data-driven grapheme-to-phoneme conversion and (b) derivation of "phonelike" subword units automatically from the acoustic data.

These research directions are particularly addressing practical challenges such as non-native or accented speech recognition, rapid development of ASR systems for new or under-resourced languages and/or domains. In addition, building on top of the core ASR research, we are also actively carrying out research on development of novel approaches for (a) utterance verification, (b) keyword spotting (spoken term detection), (c) objective assessment of coded speech and synthetic speech, and (d) automatic assessment of second language learners (funded through SNSF project A-MUSE, Idiap internship project, EC Eurostars project D-Box, armasuisse[16] , CTI project ScoreL2[17]). Keyword spotting is also addressed in the Applied Machine Learning group, as discussed in Section 2.7.

Speech Synthesis

- Text-to-speech synthesis (TTS).

  Although newer than ASR, TTS is now an established venture for the speech group at Idiap. TTS is central to three ongoing projects: SIWIS[18] (Spoken Interaction With Interpretation in Switzerland), RECOD[19] , on low bit rate speech coding (see below), D-BOX, and SP2[20] (SCOPES project on speech prosody).

  The TTS work at Idiap is largely aligned with the current trend in the field towards statistical synthesis, which uses the same technology as ASR. SIWIS aims to do speech to speech translation in Swiss languages, including the

---

13 http://www.idiap.ch/scientific-research/projects/adaptive-multilingual-ASR-TTS

14 http://www.idiap.ch/scientific-research/projects/parsimonious-hierarchical-automatic-speech-recognition

15 https://www.idiap.ch/scientific-research/projects/high-accuracy-speaker-independent-multilingua ternship project, EC Eurostars project D-Box, armasuisse16

16 http://www.ar.admin.ch/internet/armasuisse/en/home.html

17 http://www.idiap.ch/scientific-research/projects/automatic-scoring-and-adaptive-pedagogy-for-oral-language-learning

18 http://www.idiap.ch/project/siwis/

19 http://www.idiap.ch/scientific-research/projects/low-bit-rate-speech-coding

20 http://www.idiap.ch/scientific-research/projects/scopes-project-on-speech-prosody

concept of accents. A key theme of SIWIS is adaptation of speech prosody, i.e., spoken emotion. This will ensure that, e.g., questions and stressed phrases remain as such after translation. To this end, we are developing novel models of speech prosody. The goal of SIWIS are mirrored in the SP2 project with a focus on eastern-European languages. Figure 5 illustrates one of our recent approaches to decompose the pitch contour (superimposed on top of the speech signal) into "atomic" patterns (displayed below the speech signal and pitch contour), which can be related to the muscular twitches at the glottal source.
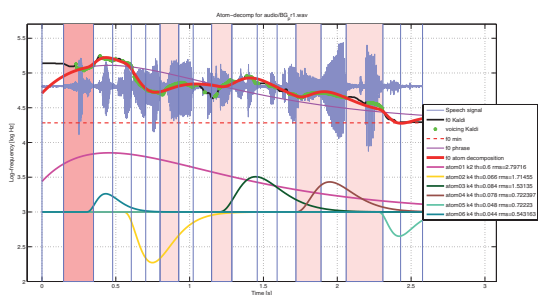

Figure 5: Modelling prosody using distinct atomic muscular twitches.

SIWIS and SP2 highlight the requirement for a collaborative approach: Speech to speech translation requires a large infrastructure that is difficult to maintain at a single site. To achieve this, SIWIS brings together partners at ETHZ and the Universities of Geneva and Edinburgh; SP2 adds partners from Budapest, Skopje and Novi Sad.

R&D work in TTS is also one of active areas for D-BOX project. Besides integrating the engine into an open-gaming platform through the implemented API supporting real-time multi-player response, research in fast adaptation of TTS engine towards target speaker or different language is carried out.

- Speech coding
  Another research area requiring TTS is speech coding, where very low bit rates can be achieved by using a speech recogniser followed by a synthesiser. Under the RECOD project funded by armasuisse, the technology lends itself to operate at 200–300 bits per second. We have developed a prototype of a syllable-context phonetic vocoder deployed for English and German languages. The low-latency speech coder runs incrementally, and both segmental (phonemes with durations) and supra-segmental (syllable stress with boundaries) are analysed, transmitted and re-synthesised.

Speaker Recognition and Speech Analytics

Starting from 2014, within the scope of the SIIP (EC FP7 project)[21] , there has been a rejuvenation of core speaker recognition research in the Speech and Audio Processing group. In particular, as illustrated in Figure 6, Idiap is involved in developing a cutting edge suspect identification solution for identifying speakers captured in lawfully intercepted calls, including multiple types of social-media information streams. In addition to core speaker recognition, the technology will capitalise on fusion of other speech analytics algorithms such as gender, age, language, or accent identification, keyword and taxonomy spotting and voice cloning detection. Speaker recognition is also addressed in the Biometric Person Recognition group, as discussed in Section 2.9.

Figure 6: Illustration of suspect identification system exploiting lawfully intercepted calls, including multiple types of social-media information streams.

Audio Processing

Besides ASR, TTS and speaker recognition, Idiap has been active in other areas of generic audio signals (pre-)processing and metadata extraction, including:

- Sparse structured representations.
  Idiap continues theoretical studies of sparse structured representations of audio signals, more specifically on model-based sparse component analysis, with different practical applications, such as source separation and localisation based on a small number of microphones, room acoustic modelling, and new beamforming techniques to improve recognition of speech recorded from far-field microphones. In 2014, our research focused on exploiting generic low-dimensional structures such as low-rank representations for improving room acoustic modeling (funded through SNSF project PHASER).

- Microphone arrays.
  Idiap also works on audio recorded from microphone arrays, especially those organized in "ad-hoc" geometries (funded originally through the NCCR IM2, and more recently PHASER). Recently, we mainly focused on the automatic calibration of arbitrary microphone array geometries, fusing Euclidean constraints into generic mathematical concepts. This in turn enables the localization required for high quality data acquisition. In this context, we also consider environmental ambiguities such as reverberation and overlapping speech, thus requiring multi-source localization using ad-hoc microphones in reverberant rooms. In addition to that, we also investigated (a) robust array configurations for ad hoc arrays for optimal source localization and (b) alternative approaches to sensor calibration based on non-negative matrix factorization and cross correlation properties in a diffuse sound field.

- Speaker diarization.
  In speaker diarization — the task of annotating temporal regions of audio recordings with labels indicating "who spoke when" — Idiap focuses on processing spontaneous overlapping speech through the use of multiple sources of information investigated in the context of various projects, including the SNSF project RODI[22] , SNSF project DIMHA[23] and EU project inEvent[24] . Over the last 12 months, our research has focussed on (a) linking and uniquely identifying speakers across large speech databases by making use of multiple modalities, (b) extraction of novel features in the frameworks of i-vectors and ANNs, and (c) scaling the KL-HMM approach, originally developed for ASR, to speaker diarization.

- Audio indexing
  In the context of the Hasler Foundation project SESAME[25] , Idiap aims at advancing the state-of-the-art in speech processing and spoken language understanding for automatically structuring data from spoken cultural heritage archives. Technical challenges include speaker segmentation and clustering, role recognition as well as topic/story segmentation and tracking over very heterogeneous. In 2014, our research focused on unsupervised topic segmentation using social role information.

---

22 http://www.idiap.ch/scientific-research/projects/role-based-speaker-diarization

23 https://www.idiap.ch/scientific-research/projects/diarizing-massive-amounts-of-heterogeneous-au set_language=fr

24 http://www.inevent-project.eu/

25 http://www.idiap.ch/scientific-research/projects/searching-swiss-audio-memories

- Very low bit-rate speech coding.
  Over the last three years, Idiap has become active in the domain of very low bit-rate speech coding through a collaboration with armasuisse, in the context of the RECOD project. We have proposed a novel architecture for low bit rate coding based on neural networks. This coder differs from previous attempts by replacing HMMs by deep neural networks, and by using the phonological speech representation instead of the phonetic one. In addition, the coder further lowers the operating latency.

- Perceptual background noise analysis for telecommunication systems.
  Idiap has been investigating how different types of background noises in speech are perceived by human listeners. In the context of a CTI collaboration with the industrial partner SwissQual AG, Idiap has developed a novel model of human noise perception. The model targets an upcoming ITU standard for the quality assessment of new wideband (so-called "HD Voice") telephony systems.

2. Computer Vision and Learning

Key scientific outputs: Our work has resulted in several novel algorithms for learning in high dimension, based on joint modeling of feature response, and a new "reservoir" strategy to leverage memory in an on-line setting. We have also developed a novel learning-based approach to pose estimation from very low quality image, and a hard-sample harvesting procedure with sub-linear computing time. In a long-lasting collaboration with EPFL, we have developed a multi-camera tracking algorithm able to combine the detection of pedestrians and vehicles in a principled manner.

Machine Learning

- Joint feature selection.
  To classify signals using very large number of measurements, or features, it is often desirable to apply a pre-processing step during learning which selects a "good" sub-set of them. Many criteria have been developed for that task over the years, but while it is easy to measure how good a single feature is, these techniques often fail at taking into account the joint quality of a set of features.
  We have developed new methods which relies on a joint models of feature responses, under which we explicitly measure the information content of the subset of features we are selecting. Our main contribution is a "Gaussian compromise", to approximate the entropy of a mixture of Gaussian densities, for which no analytical or tractable expression exist, and that we use to approximate the mutual information.

- Learning in large dimension feature space.
  In the context of the MASH project[26] we have designed new Boosting procedures to learn with hundreds of families each of thousands or millions of features, by either sampling a few features from each family before the learning starts, and to use this features to estimate at every Boosting step the most promising feature family, or by optimizing on-the-fly the number of samples and features to look at at every boosting step.
  Our current work addresses the design of artificial neural networks in that context. We investigate novel algorithmic organizations of the computation coupled with sparse learning to reduce the cost for their evaluation.

- Sub-linear hard sample extraction.
  In the SNSF project DASH, we are interested in speeding-up "hard sample" extraction. Most of the state-of-the-art detection algorithms are trained in a sequential manner, improving iteratively the detector by training it with samples miss-classified by its previous version.
  Since the performance of detectors can be extremely good, finding a sufficient number of hard samples is computationally-intensive, and in all available methods, has a cost proportional to the number of images to visit.
  Our approach consists in using the prior structure provided for any large collection of images, that is its recursive decomposition into homogeneous sub-families, to concentrate computation in a principled manner. We organize the set of images into a recursive tree consistent with the similarity between images, and adapt the Monte-Carlo Tree Search to recursively apply a bandit approach to the exploration-exploitation dilemma.
  In practice, our approach alternate between looking into set of images and images it has poor information about, and images from which it has observed a high success rate in finding hard samples.

- High-dimension density similarity measures.
  We studied the clustering and classification of distributions of points in high-dimension spaces. To leverage the remarkable ability of random forests to cope with such situations, we have proposed a novel metric, which looks at the Entropy of the empirical distribution of a group of samples in the leaves of a forest build for another group of points.
  If the two groups follow the same distribution, this entropy should be similar to that of the initial group, hence maximum. We applied this metric to the study of the growth dynamic of neurons, and its modulation by their genotypes in the framework of a collaboration with the University of Geneva, the University of Basel, and EPFL.

Figure 7: We have developed a novel strategy to find informative training examples in an image that concentrates the computation on difficult images and parts of images (left). The Digit-Arena project aims at replacing ads on perimeter boards automatically in broadcast video (right).

Scene understanding

- Playground learning with pose-indexed features.
  Related to our work on the stationary features, we are working in collaboration with the Universitat Politècnica de Catalunya, on the development of a new technique to leverage high-quality HD training images, to build models able to cope with extremely degraded test images of very low resolutions. The key idea is to use the high-quality signal to build strong priors, namely 3-D models, and to design pose-indexed features automatically from these models. Ultimately, we analytically define measurements in the images, parametrized by the pose of the object whose presence we want to check, and train a predictor from these measurements.

- Multi-camera tracking.
  We have a long-lasting collaboration with the CVLab at EPFL around the design of a robust multi- camera tracking system, which now takes place in the context of the SNSF WILDTRACK project, in collaboration with ETHZ. Our most recent results consist of a new model and optimization scheme to handle properly the appearance / disappearance of "containees" (e.g. pedestrians) given the presence / absence of "containers" (e.g. cars).
  Our system is composed of the Probabilistic Occupancy Map, and a flow-based convex multi- target tracker. The former estimates in individual time frames marginal probabilities of presence of people by minimizing the Kullback-Leibler divergence between the "true" posterior under a joint generative model, and a product law. The latter connects these detections properly by optimizing a flow in a spatio-temporal graph. Both components are extremely efficient computationally and process in real-time multiple camera streams to produce state-of-the-art results. Open-source implementations of both components are available under the GPL3 license.

- Advertisement substitution in sport events.
  In a collaboration wit E-S. Concept. SA, we are developing a novel pipeline of algorithms to perform advertisement substitution on-the fly in broadcast video streams.

___

26 http://mash-project.eu

The objective of this collaboration is to deliver a solution able to automatically and accurately localize perimeter advertisement during a sport event, and replace it with another one. We have already implemented a demonstration pipeline which demonstrates the feasibility of such a system, in particular in term of stability of the board positions in the image plan.

3. Social Computing

Key scientific outputs: Published studies on multimodal inference of hirability in employment interviews; published studies on discovery of multi-dimensional socio-affective attributes in conversational social video (YouTube) integrating automatic audio-visual analysis and video crowdsourcing; mobile crowdsourcing experiences to understand urban phenomena in cities; 10 EPFL PhD students graduated since 2002.

Face-to-face interaction

We are developing computational models to analyze dyadic and small-group interactions in workplace settings using a variety of sensors (cameras, Kinect devices, microphone arrays, smartphones, and wearables), and automatically infer key social variables for psychologists and recruiters (Figure 8). This is collaborative work with the University of Lausanne, Cornell University, and Vatel Switzerland. In the dyadic case, in the context of organizational behavior situations, we have investigated connections between hirability and personality. In the small-group case, we have studied links between behavioral cues and variables like emergent leadership and personality traits. In all cases, nonverbal cues are automatically measured from audio and video streams.

Figure 8: Sensors available in the social sensing lab. (1) Kinect V2, (2) Microcone, (3) Q-Sensor, (4) Android smartphone, (5) Google Glass.

This research was conducted in the framework of two SNSF projects, namely UBImpressed[27] : Ubiquitous Frist Impressions and Ubiquitous Awareness, Sinergia Interdisciplinary

27 http://www.idiap.ch/project/ubimpressed

Program, and SOBE[28], an Ambizione Research Fellowship. The group's research in this domain was covered by the European Commission as a H2020 Project Story[29] .

Social video analysis

Our work in this line develops methods to characterize users of conversational social video in social media sites like YouTube through the use of (1) video crowdsourcing techniques through platforms like Amazon Mechanical Turk to collect interpersonal judgments, which scales up to large amounts of data while relying on demographically diverse viewers; and (2) automatic extraction of behavioral features including prosody, body motion, and facial expressions, which enables high-level social analysis. Our research has shown how certain impressions of personality and mood can be automatically inferred from nonverbal and verbal cues.

This research was supported by the NISHA[30] project (NTT-Idiap Social Behavior Analysis Initiative) funded by NTT Communication Science Laboratories, Japan, and by an EPFL Innogrant Award. This work received press coverage at the national level[31] .

Figure 9: Left: SenseCityVity: Young people in Mexico document and reflect upon urban concerns. Right: Mapping social media activity in Nairobi from geo-localized tweets.

Urban crowdsourcing and mobile social media mining

Our work in this domain spans several research lines. First, we have designed and implemented a mobile crowdsourcing methodology to engage youth in cities to collect multimedia items (geo-localized images and videos) that document urban concerns (Figure 9). Furthermore, we have studied how urban awareness can be crowdsourced and then modeled to understand how sensitive image observers are with respect

28 http://www.idiap.ch/scientific-research/projects/multimodal-computational-modeling-of-nonverbal-social-behavior-in-face-to-face-interaction-1

29 http://ec.europa.eu/programmes/horizon2020/en/news/cracking-code-unspoken-language

30 http://www.idiap.ch/project/nisha/

31 http://www.computerworld.ch/news/artikel/epfl-persoenlichkeitstests-mittels-youtube-65015/

to specific urban issues. This work was done in the context of the SenseCityVity project[32] (Mobile Sensing, Urban Awareness, and Collective Action) funded by EPFL Center for Cooperation and Development (CODEV), and carried out with a multidisciplinary team from IPICYT, Mexico. In other work, through the SNSF Youth@Night[33] project (A Multidisciplinary Study of Young People's Going Out and Drinking Behaviors, in collaboration with Addiction Switzerland and the University of Zurich) and the Mirror Africa project [34] (a joint initiative with EPFL AfroTech Lab), we are investigating the use of mobile crowdsourcing and mobile social media mining (like Twitter and FourSquare) to characterize phenomena at urban scale in Swiss and African cities, including characterization of regions according to patterns of human activity and place ambiance (Figure 9). The work in Africa was covered in the media by The Economist[35].

Visual Analysis of Maya Hieroglyphs

We are developing methods for automatic visual analysis of hieroglyphs for the three most celebrated ancient Maya codices, which are located in museums in Europe – Dresden, Paris, Madrid (Figure 10). Our methods are designed to support specific expert needs and include visual retrieval and classification of glyphs. This is a collaboration with Maya archeologists from the University of Bonn and computer scientists from the University of Geneva, and is supported by the SNSF-DFG project MAAYA[36] project (Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy).

Figure 10: Preprocessing steps to generate Maya glyph data for visual analysis: (a) color images, (b) binarized images, (c) reconstructed glyphs by expert epigraphists.

---

32 http://www.idiap.ch/project/sensecityvity

33 https://www.youth-night.ch/

34 http://actu.epfl.ch/news/surveying-african-cities-using-twitter/

35 http://www.economist.com/blogs/baobab/2014/05/twitter-kenya

36 http://www.idiap.ch/project/maaya/

4. Perception and Activity Understanding

Key scientific outputs: The group is well known for its work on video sequence analysis, probabilistic tracking and non-verbal behavior extraction. In 2006, it developed a unique video stabilization algorithm that is currently applied automatically on all super-8 and videos digitized by the Cinetis company. In 2007, the Klewel company was created using its slide OCR technology. During the period 2009-2014, it published 17 journal papers and more than 50 conference papers, and received 1 best paper and 2 best student paper awards. The group has pioneered work on unsupervised surveillance scene activity, visual attention modeling, and its recent work on gaze estimation from cheap RGB-Depth sensors is under patenting.

Multiple person tracking

This research line has been finalized with the thesis end of A. Heili. It resulted in an enhanced tracking-by-detection system for surveillance applications that leverages several benefits important in this context, and which are illustrated in Fig. 11. Given the individual human detections at each instant, the problem is set as a graph partitioning problem (within a Conditional Random Field CRF framework) such that detections corresponding to the same person are associated together based on multicue measurements (position, image-based motion, multiple patch colors, etc). Difficulties lie in the graph specification (links between nodes and association costs) and its optimization.

Important novel features of our model are the following. Links are created between detections that can be a few seconds apart, rather than only a few frames apart. All links are used in defining and evauating the cost function, not only those between consecutive nodes from the same track. This allows assessing a solution using both short-term and long-term links and defining costs not only in terms of detection node similarities, but also dissimilarities. Importantly, pairwise cue measurement costs (defined as similarity/dissimilarity likelihood ratio) are sensitive to the time interval between two detections, integrate multi-cue association reliability factors, and all cue likelihood distribution parameters are learned in an original and unsupervised fashion, thus accounting for the scene specificities while avoiding tedious annotations. Finally, global label costs favoring the creation or deletion of tracks in specific scene regions are also exploited. Altogether, this results in a well conditioned optimization criterion that can be easily be maximized. Results equal or better than the state-of-the-art were obtained on several benchmark data of the field. The work received the best paper award at the annual IEEE Performance Evaluation of Tracking and Surveillance (PETS) workshop. The software has been licensed for research purposes to a large leading IT company active in camera manufacturing.
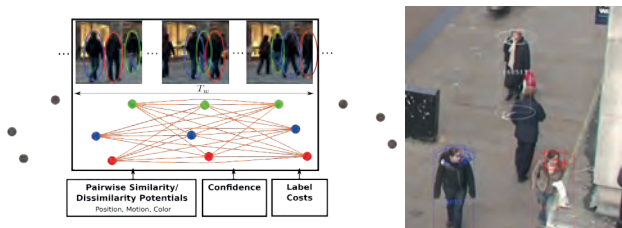
Figure 11: Left: the multi-person tracking framework. Human detections at each frame are associated based on long-term multi-cue and time-interval sensitive discriminative costs. Right: estimated head and body orientation of people.
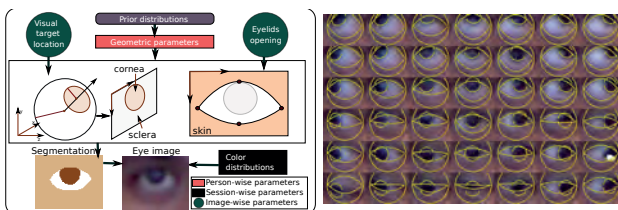

Figure 12: Gaze estimation. Left: The generative process combining geometric eye modeling with a segmentation based eye image processing approach avoiding geometric feature extraction (e.g. ellipse fitting). Right: exemple of eye segmentation resulting from the modeling.

Behavior cue extraction in surveillance scenes.

Along with person tracking, we have continued our work on body and head pose estimation (see Fig. 11, Right) which are important cues for activity analysis. Our previous work was addressing the problem through the joint adaptation to new scene of head and body pose classifiers. While the method accounted for the different coupling (between head and body pose due to anatomical constraints, between motion and body pose when people are walking), it did not handle the discrepancy problems resulting from the mismatch between training and test data. That is, it assumed that similar features from the training and test set where corresponding to the same pose, an assumption that does not hold well in practice. To address this problem, we have proposed an additional synchronized manifold alignment approach which learn two feature projection mappings enforcing that train and test features in the projected space form a single smooth pose manifold. In this way, the method leverages on the availability of weak pose labels (from people motion) in the test data to learn better pose classifiers.

Interaction modeling and gaze extraction.

Our long-standing work on non-verbal behavior cue extraction and human interaction modelling has been continued in the context of the SNSF funded project G3E, which more specifically addresses gaze extraction from cheap RGB-D (depth) cameras like Kinect. Building on our the head-pose free gaze estimation framework, we have proposed several extensions to enhance the performance at different levels (gaze estimation,

software system, usability). In particular, we have proposed a model that relies on a geometric understanding of the 3D gaze action and generation of eye images. By introducing a semantic segmentation of the eye region within a generative process, the model (i) avoids the critical feature tracking of geometrical approaches requiring high resolution images; (ii) decouples the person dependent geometry from the ambient conditions, allowing adaptation to different recording conditions without retraining the gaze model; (iii) allows the exploitation of priors to train from few samples; (iv) is capable of gaze extrapolation allowing for less restrictive training schemes. The figure 12 illustrates the process and some results.
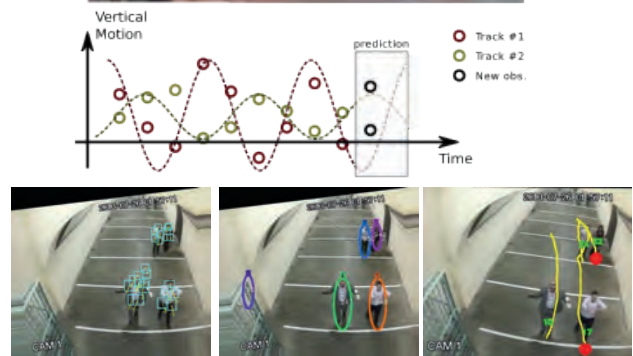

Figure 13: Illustration of the bobbing. During motion, people shoulder move up and down (as well as left and right) according to the gait pattern. Analysis of their visual motion (extracted from their body parts) provides information about the gait frequency as well as when the heel strikes the floor (red dots on the right image).

We have also started a collaboration with the 'Speech and Hearing group' of KTH on the analysis of interaction. More precisely, we have recorded a dataset featuring meetings of 4 people where 3 of them defend their PhD topic and grant request. Automatic turn-taking, gaze tracking and head nodding are currently automatically extracted to learn better and more fine grained interaction models, esp. regarding gaze coordination near speaker changes and backchannels.

5. Robot Learning & Interaction

Task-adaptive motion encoding

The group investigates the problem of encoding movements based on a limited number of demonstrations collected in various situations, such as different positions of objects for a manipulation task, with the aim of further synthesizing new movements that are robust to perturbations. Fig. 14 illustrates the challenge.

Key scientific outputs: We developed an approach simultaneously modeling the observed motion in multiple coordinate systems or frames of reference. The different projections of the original data reveal regularities that are exploited to generalize and adapt the task to new situations in a probabilistic manner. Notably, the retrieved variability and correlation information is used to regulate the stiffness of the robot in order to generate safe and natural movements. This also enables new haptic communication capability that has great potential in human-robot collaboration. Aspiring to reproducible research, the group released open source codes in two versions (either C++ or Matlab, by ensuring full compatibility with GNU Octave)[37].

Learning by imitation, emulation and exploration

Learning from demonstration encompasses a wide range of imitation strategies, from actions mimicking to the extraction of high level intents underlying the actions (goal emulation). The re-use of previously observed actions combined with an objective-based representation is not only pertinent from a computational point of view: it also has roots in the way humans acquire skills by efficiently combining imitation and emulation. Children often employ as default strategy accurate copying, based on which progressive elaborations can later be built to face new situations (copy-all correct-later strategy). Social learning, play, practice, exploration and feedback provide varied opportunities to correct partially or wrongly assimilated aspects of previously observed actions, which can lead to more versatile adaptation mechanisms. We take the perspective that action-level and goal-level copying can be combined to transfer skills from humans to robots, but also from robots to humans and in-between robots.

Key scientific outputs: We developed an approach to exploit context-dependent objective functions, that contrasts with standard inverse optimal control techniques in which the same objective is defined throughout the task. The observed actions are encoded together with associated objective functions, which is exploited to search for new solutions by stochastic optimization.
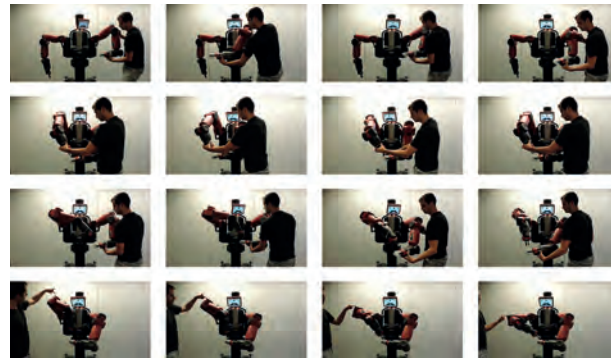
*Figure 14: Incremental learning with the Baxter robot at Idiap. The taught task consists of holding a cup horizontally with one hand, and holding a sugar cube above the cup with the other hand. The demonstrations are provided in two steps by kinesthetic teaching, namely, by holding the arms of the robot and moving them through the task while the robot compensates for the effect of gravity on its limbs. This procedure allows the user to move the robot arms without feeling their weight and without feeling the motors in the articulations, while the sensors are used to record position and force information. Here, the data are recorded in several frames of reference. During reproduction, the robot is controlled by following a minimal intervention principle, where the impedance parameters of the robot (stiffness and damping of a virtual spring pulling the robot arms) are set in accordance to the extracted variation and coordination patterns. **First row:** Short demonstration to indicate how to hold a cup horizontally. **Second row:** Short demonstration to show how to hold a sugar cube above the cup. **Third row:** Manual displacement of the left arm to test the learned behavior (the coordination of the two hands was successfully learnt). **Last row:** Combination of the two tasks within a minimal intervention controller. Here, the user pushes the robot to show that the robot remains soft for perturbations that do not conflict with the learned task constraints.*

---

6. Artificial Cognitive Systems

Key scientific outputs: The group is internationally recognized for its work on life long learning, adaptive control of dexterous prosthetic hands and benchmarking activities. In 2010, we developed the first max-margin based transfer learning algorithm for visual detection. The algorithm is the current state of the art and is considered a reference in the field. Since 2011 we have been creating the first publicly available database of surface ElectroMyoGraphy (sEMG) recordings of hand postures from intact and amputated subjects. The database is the largest existing for number of postures and number of subjects recorded, and it has become the reference benchmark in the field. From 2007 to 2013 we organized the Robot Vision Task challenge under the ImageCLEF umbrella, which has seen the participation of more than 100 teams over the years. As a further recognition of the value of our benchmarking activities, during 2013 and 2014 Prof. Caputo was the main organizer of ImageCLEF.

Semantic Spatial Understanding of Indoor Places

The work on semantic spatial understanding of indoor places, started in 2005, has been continued in the context of several projects, including the VISION@HOME project, funded by the Swiss National Science Foundation. It proposes the innovative Situated Vision paradigm, developing 3D visual perception capabilities from the view of the robot, its task and the environment it operates in.

The Situated Vision approach is inspired by recent work in cognitive science, neuroscience and interdisciplinary work in EU projects: it fuses qualitative and quantitative cues to extract and group 3D shape elements and relate them to affordance categories. Task-oriented and 3D attention mechanisms let the robot execute primitive actions to exploit the affordances perceived. Perception integrates quantitative and qualitative shape information from multiple 2D and 3D measurements. The analysis of the shapes is used to find instances of semantic 3D concepts, such as providing support to objects, enclosing space, etc. that can be used to those spatial concepts to find semantic entities, such as table surfaces, cupboards, closets, drawers and to learn which perceived affordances belong to which object category.

Over the last year, we casted the semantic spatial modeling problem into that of learning from multiple cues. Our contributions have been a principled online Multi Kernel Learning algorithm able to combine optimally multiple features while providing theoretical guarantees on the expected performance, and a general learning to learn algorithm, based on the cue integration paradigm, able to leverage over prior knowledge in an extremely flexible and adaptive manner, over a wide range of application scenarios. This last algorithm, the first of its kind in the literature, achieves state of the art results on benchmark databases widely used in the literature and is currently implemented in a running robot system.

Advanced Hand Prostheses

The work on advanced hand prostheses has been initiated in 2011 with the SNSF funded NINAPRO[38] project. Its goal is to develop a family of algorithms able to significantly augment the dexterity, and reduce the training time, for sEMG controlled prosthesis. Indeed, the state of the art in hand prosthetics does not offer more than 2-3 degrees of freedom and a very coarse control of the force, as there is no haptic feedback. Patients interface with the prothesis via surface electromyo-graphy (sEMG), recorded using surface electrodes. Learning how to control the device through many input sEMG channels is a long and difficult process for most patients, that therefore settles for limited and very simplified movements (open/close).

This contrasts with recent advances in mechatronics, thanks to which mechanical hands gifted with many degrees-of-freedom and force control are being built. There is a need for prosthetic hands able to naturally reproduce a wide amount of movements and forces, while at the same time requiring a lower effort in learning how to control hand postures. This goes beyond mechatronic dexterity: the real challenge is how to provide patients with a cheap, easy and natural way of controlling the prosthesis. By testing its findings on a very large collection of data, NINAPRO has the ambition to pave the way for a new generation of prosthetic hands. The work is organized along four themes: data acquisition and analysis (Theme 1), augmented dexterity: posture classification (theme 2), augmented dexterity: natural control (theme 3) and adaptive learning (Theme4). Prof. B. Caputo is the coordinator of the project, and the principal investigator in the Themes 2 and 4 (Figure 15).

---

38 http://www.idiap.ch/project/ninapro/

The objective of theme 2 (augmented dexterity) is to push the current state of the art in prosthetic hand posture classification from handling a maximum of 12 postures up to 40-50. Over the last year, we have demonstrated the importance of using multi modal features to increase the robustness of posture classification and therefore of actual control of the prosthesis, by merging information acquired from sensors applied on the stump of the user with those recorded by an accelerometer mounted on the prosthesis. This work has resulted in a publication on the international journal IEEE Transaction on on Neural Systems and Rehabilitation Engineering.
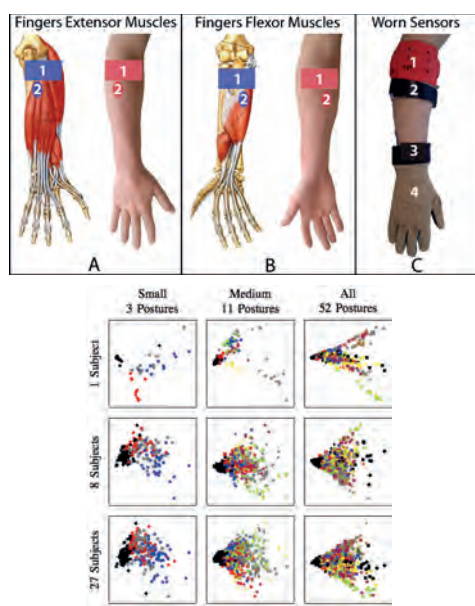

Figure 15: One of the most important results of the NinaPro project has been the creation of the very first open access database of sEMG recordings of hand and arm movements from intact subjects and amputees. As of today, the database contains recording of more than 50 postures from over 100 intact subjects and 15 amputees. On the left, we show the placement of sensors with respect to the muscles activations; on the right, a probabilistic analysis of a first batch of data, illustrating as the signals acquired from the different subjects become more and more difficult to interpret as the number of subjects and hand postures grows.

The goal of theme 4 (adaptive learning) is to develop learning algorithms to better interpret the sEMG signals acquired from the patients, with the ultimate goal of boosting the learning process necessary for the patient to effectively use the prosthesis. Our plan is to build pre-trained models of various data postures, on the data acquired in theme 1, and adapt these general models to the needs of individual users as new data will became available using adaptive online learning methods. Over the last year, we did build over our results obtained in the large-margin classification framework, and we explored how to maintain the stability and effectiveness of our method when in presence of a very large number of prior models. This would be indeed the case when leveraging over priors derived

from the actual experience of intact subjects, from which it is relatively easy to acquire large amounts of data and hence build strong prior models. We developed a new version of our previous algorithm able to scale well over a very large number of priors in terms of computational efficiency, while preserving a very high accuracy in terms of hand posture recognition.

7. Applied Machine Learning

Overview

Key scientific outputs: The Applied Machine Learning group has been a pioneer in applying deep learning approaches to Natural Language Processing (NLP). In that matter, the group released a state-of-the-art syntactic parsing tool, which completes the previously released "SENNA" software, a renowned tool for a collection of NLP tasks. We have also investigated advanced ways for representing words and phrases as vectors, while keeping the syntax and semantic information. Concerning computer vision, we have been working on the object segmentation task with a specific recurrent deeplearning architecture. We have also developed a new state-of-the-art segmentation system relying only on weakly labeled data, as well as a state-of-the-art system for captioning images. In the speech processing field, we developed new algorithms for training deep learning systems from raw speech (waveform) without the need of phonemes (or grapheme) segmentation.

Deep Learning

Many state-of-the-art Natural Language Processing, Computer Vision or Audio Processing systems address a task by applying simple statistical models to carefully hand-crafted features. In other words, the researchers themselves discover intermediate data representations relevant for the task of interest, by engineering task-specific features. In contrast, deep learning algorithms aim at learning several layers of features representing raw data, with an increasing level of abstraction. Our group investigates both new generic deep learning algorithms, as well as algorithms specific to the following applications:

- Semantic Word, Phrase and Document Representations. The SODS project, funded by the Swiss Hasler Stiftung[39] is interested by a new type of semantic search (in contrast to existing search engine, which are mostly keyword-based). In that respect, we have a huge interest in investigating new algorithms for representing words, phrases, as well as documents in a vector space, while retaining both the syntactic and semantic information. These representations (see Table 1) are trained on large unlabeled text corpora, without the need of any supervision. They are essential for a lot of

Natural Language Processing tasks, where they are used as powerful features carrying semantic information.

- **Natural Language Processing (NLP).**
In an ongoing collaboration with NEC Laboratories America[40] , we investigate new general machine learning algorithms for various NLP tasks. We leverage semantic representations from the SODS project in a single deep network architecture which outputs tags for several NLP tasks ranging from syntax analysis (Part of Speech tagging, chunking, syntactic parsing) to semantic analysis (Name Entity Recognition, Semantic Role Labeling). Our architecture is state-of-the- art both in accuracy and speed performance, while having the ability to compose phrases while retaining both syntactic and semantic information.

AMERICAN AIRLINES	CHICAGO BULLS	HOME PLATE	PRESIDENT OF THE UNITED STATES
BRANIFF AIRLINES	DENVER NUGGETS	RIGHT FIELDER	PRESIDENT COOLIDGE
ALOHA AIRLINES	SEATTLE SUPERSONICS	CENTER FIELDERS	PRESIDENT EISENHOWER
BRANIFF AIRWAYS	CLEVELAND CAVALIERS	THE OUTFIELD FENCE	U.S. PRESIDENT DWIGHT EISENHOWER
JETBLUE AIRWAYS	BOSTON CELTICS	LEADOFF BATTER	PRESIDENT TRUMAN
BRANIFF INTERNATIONAL AIRWAYS	DALLAS MAVERICKS	THE INFIELD	PRESIDENT REAGAN

Table 1: Examples of phrases (first row) and their five closest neighbors (below) in our phrase vector representation space. Note that phrase neighbors might have a different length, and are relevant both in terms of semantic or syntax measure

- **Object Segmentation.**
Object segmentation can be viewed as a classification task, which assigns a label to all pixels in an image according to the class they belong to. Most systems rely on costly graphical models to take a global decision from local classifiers. We investigated new efficient end-to-end systems based on recurrent convolutional neural networks (see Figure 16), alleviating the need of any hand-crafted feature and allowing the modeling of complex spatial dependencies with a low inference cost. Compared to existing approaches, our system can be trained using only weakly labeled data, having only the class information of the objects present in one image, rather than the full labeled segmentation

(a)

(b)

(c)

Figure 16: Our recurrent scene parser over original image (a) segments the image as shown in (b). Due to its recurrent nature, it can be fed again with its own prediction (b) and the original image (a), which leads to (c): most mistakes are corrected.

- **Image Captioning.**
This sub-project is at the frontier between Natural Language Processing and Computer Vision, involving the two students from the SODS and Object Segmentation projects. We developped a system which generates sentences according to the representation of an image provided by a deep learning model. It leverages semantic vector phrase representations. The system reaches state-of-the-art performance, both in speed and quality of the generated sentences. See examples of caption generation in Figure 17.

- **Spoken Term Detection.**
The DeepSTD project, funded by the Swiss Hasler Stiftung[41], is concerned about applying new deep learning techniques to audio processing, with a particular interest in fast spoken term detection. We developed a new system which can label phonemes (or graphemes) directly from the raw speech signal (waveform), using convolutional neural networks. We proposed new state-of-the-art techniques for training these type of models in a end-to-end manner, without the need of any segmentation.

---

40 40 http://ml.nec-labs.com

41 http://www.haslerstiftung.ch

Technology Transfer

Our group puts a strong emphasis on the reproducibility of all our research findings, such that our tools can be integrated easily in more complex systems or products. Tools we developed include:

- Face Technologies.
  This year we started a collaboration with the KeyLemon[42] startup, through the CLEAR CTI project. We are in charge of developing state-of-the-art algorithms for multi-pose face detection, head pose detection, facial feature detection, and gender detection. In 2013, we delivered a state-of-the-art version of our multi-pose face and gender detection system (see Figure 18), which runs in real time with good performance even in adverse conditions.
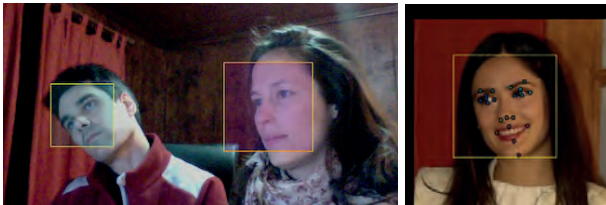

Figure 18: Our real time face detector detects multi-pose faces and labels gender (pink for female, bluish for male) in adverse conditions (left). It also outputs 22 precise facial features, in multi-pose conditions (right).

- SENNA, a Natural Language Processing (NLP) tagger.
  SENNA[43] is a standalone software distributed under a non-commercial license, that we develop through our ongoing collaboration with NEC Laboratories America[44]. It outputs a host of Natural Language Processing predictions, at blazing speed while holding state-of-the-art accuracy performance. Predictions include part-of-speech tags, chunking, name entity recognition, semantic role labeling and syntactic parsing.

- Torch, an Open-Source Machine Learning Platform.
  We are maintaining Torch[45] an open-source (BSD license) fast and scalable Matlab-like system, which serve us as an efficient platform to implement all our machine learning algorithms. Torch leverages Luajit, an extremely fast scripting language, which allows us to quickly develop complex demos and prototypes. Torch is widely spread, present in many international academic and private institutions (including Facebook, Google DeepMind, or Twitter). Thanks to its open-source nature, external contributors wrote various packages for a large number of application domains.

---

42 https://www.keylemon.com

43 http://ml.nec-labs.com/senna

44 http://ml.nec-labs.com

45 http://www.torch.ch

8. Computational Bioimaging

Key scientific outputs: Recent milestones include the reconstruction of 3D volumes of the beating embryonic heart at frame rates equivalent to over 1000 volumes per second, temporal super-resolution for sensitive fluorescence cameras, and observation and quantitation of heart development in animal models.

High-Speed Fluorescence Microscopy

Fluorescence microscopy has become a central tool for the study of biological structures, from single molecules to whole organisms. Fluorescence light emitted by dim samples is scarce and limits the frame rate at which dynamic processes can be observed with sufficient signal-to-noise ratio or without severe motion blur artifacts. Our group (together with UCSB graduate student Kevin Chan) has proposed a temporal super-resolution algorithm to reconstruct a high temporal resolution image sequence from multiple low temporal resolution acquisitions. Each acquisition is shifted in time by a subframe delay and a cost minimization is used to reconstruct the high temporal resolution sequence. Via experiments on a dynamic phantom objet, we demonstrated that our temporal super-resolution algorithm increases the bandwidth by a factor of 1.5.

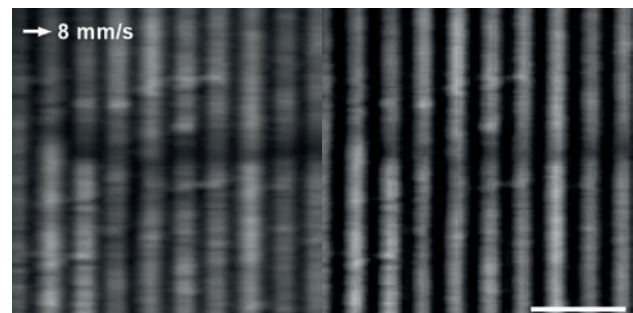*Figure 19: Temporal super-resolution reduces motion blur. Left: Frame from a movie of horizontally moving stripes acquired at 30 fps exhibits motion blur. Arrow indicates direction of motion, and length corresponds to distance traveled during shutter aperture. Right: The corresponding frame after multi-cycle temporal super-resolution reconstruction from eight sequences acquired at 30 fps. Scale bar is 1 mm.*

9. Biometric Person Recognition

Key scientific outputs: The Biometrics group has been pioneering the work on mobile biometrics (face and speaker recognition) and on anti-spoofing in face recognition by sharing the first open databases, organizing the first International competitions and producing the first reproducible research studies in the domain. In mobile biometrics, the group demonstrated that the same probabilistic session variability modeling techniques can be applied both to speaker and to face recognition. The group proposed also an exact and scalable formulation of Probabilistic Linear Discriminant Analysis (PLDA), a state-of-the- art technique in speaker recognition, hence solving a major limitation. In face anti-spoofing, the group demonstrated that the current trend in discriminant-based anti-spoofing is prone to over-fitting hence resulting in a lack of generalization on unseen spoofing attacks.

Face and Speaker Recognition

We leveraged prior work on distribution modeling for part-based face recognition using session variability modeling techniques. Session variability modeling aims to explicitly model and suppress detrimental within-class (inter-session) variation. We examined two techniques to do this, inter-session variability modeling (ISV) and joint factor analysis (JFA), which were initially developed for speaker recognition. Next, we explored Total Variability modeling (TV), so called i-Vectors originally proposed for speaker recognition, for the task of face recognition, and we reached boosted performance for the task of bimodal face and speaker recognition (Figure 20).
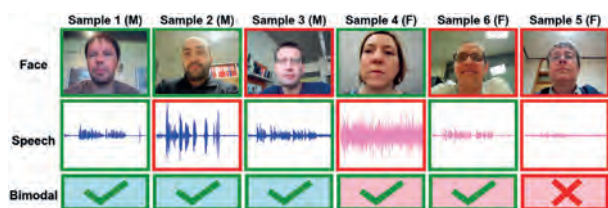

Figure 20: An illustration of face and speaker recognition.

We also developed recently a scalable formulation of Probabilistic Linear Discriminant Analysis (PLDA). PLDA is a probabilistic model that has been shown to provide state-of-the-art performance for both face and speaker recognition. However, it has one major drawback, at training time estimating the latent variables requires the inversion and storage of a matrix whose size grows quadratically with the number of samples for the identity (class). We developed a scalable derivation which is theoretically equivalent to the non-scalable solution and so obviates the need for a variational approximation. The main advantage of the proposed approach is a gain in performance when using more training samples per identity (class), which is made possible by the proposed scalable formulation of PLDA.

As a consequence, on July 31 2014 the **biometrics group was ranked 1st** among 130+ participants to the NIST i-Vector Machine Learning challenge on Speaker Recognition (ivectorchallenge. nist.gov) using this novel PLDA combined with an efficient hierarchical clustering.

Anti-Spoofing

One important aspect of biometric systems is their reliability not only when assaulted by impostors, but also under different types of attacks. One possible security treat is spoofing attacks: an action of outwitting a biometric sensor by presenting a counterfeit biometric evidence of a valid user. It is a direct attack to the sensory input of the biometric system and the attacker does not need previous knowledge about the recognition algorithm. Most of the biometric modalities are not resistant to spoofing attacks: a biometric system is usually designed to only recognize identities without concern whether the sample comes from a live person or not. Despite the existence of very sophisticated biometric systems nowadays, the task of implementing anti-spoofing schemes for them has attracted much less attention.

Recently, we advanced the state-of-the-art in anti-spoofing to printed photo attacks in face recognition by proposing public benchmarks and effective solutions based on texture modeling, motion detection or spatio-temporal pattern recognition. We also organized the first International Competition on the topic.

Reproducible Research

Over the last year the group consolidated a major evolution of the Bob (www.idiap.ch/software/ bob/) software library to support its research activities. Bob is a free signal processing and machine learning toolbox. The toolbox is designed to meet the needs of researchers by reducing development time and efficiently processing data. Firstly, Bob provides a researcher-friendly Python environment for rapid development. Secondly, efficient processing of large amounts of multimedia data is provided by fast Cpp implementations of identified bottlenecks. The Python environment is integrated seamlessly with the Cpp library, which ensures the library is easy to use and extensible. Thirdly, Bob supports reproducible research through its integrated experimental protocols for several databases. Finally, a strong emphasis is placed on code clarity, documentation, and thorough unit testing. Bob is thus an attractive resource for researchers due to this unique combination of ease of use, efficiency, extensibility and transparency. Bob is an open-source library and an ongoing community effort.

Biometrics Evaluation and Testing

Within the EU BEAT project[46] , coordinated by Idiap, we built a biometry-independent platform for Biometrics research, development and certification. By making use of such a system, academic or industrial partners enable users to easily compare results from distinct algorithms or/and parameterizations with minimal interaction using one or potentially many biometric traits. Here are three practical examples:

1. A teacher asks students to implement a specific illumination normalization algorithm in a pre-configured face recognition tool-chain. The teacher would like to receive the examination results from his group of students;

2. A researcher (PhD student or Professor) wishes to test new ideas for pre-processing (i.e. to solve translation, rotation, and scaling problems due to different positions and downward pressure of the finger) in a fingerprint recognition system. She/He wishes to upload the new algorithm to an existing fingerprint-recognition tool-chain and measure its impact on the overall system performance;

3. A scientific challenge is organized in which users contribute with features that will be used in a pre-configured key-point localisation algorithm, for face recognition. The user wishes to upload his/her own feature-extraction algorithm to the contest and observe how that performs.

One can use the platform to generate attestation (a.k.a. certificates) to any produced experimental results. Attestation can be used to support formal certification processes or improve evidence on published academic articles (Figure 21).
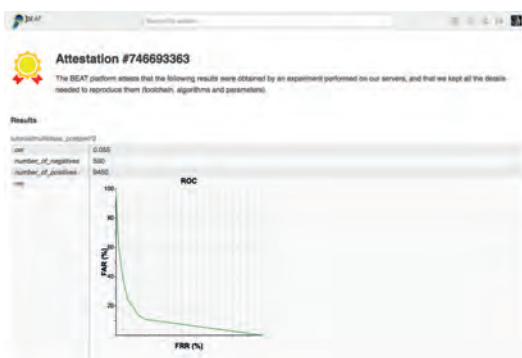

Figure 21: A snapshot of a BEAT attestation.

Swiss Center for Biometrics Research and Testing

In 2014, the Idiap Research Institute launched the "Swiss Center for Biometrics Research and Testing" (www.biometrics-center.ch), a competence center within the Institute following recent successes in coordinating International research projects in Biometrics (MOBIO, TABULA RASA and BEAT). The aim of this center is to serve as a legacy for these projects and to push for industry-driven research and testing in biometrics. The center will mainly promote two instruments: (1) the coordination of a cooperative research consortium, and (2) the deployment of the BEAT evaluation platform. These instruments were presented on November 2014 during a kick-off to researchers and stake-holders in biometrics (40+ participants). Consequently to this kick-off, we decided to elevate the visibility of the cooperative research consortium to a more European level. We concluded an agreement with the European Association for Biometrics (composed of 160 members) to transform the cooperative research consortium initiated in Switzerland into an European cooperative research consortium but still coordinated in Switzerland by the center. We expect to attract a critical mass of partners to fund the European cooperative research consortium.

Regarding the BEAT evaluation platform, we were able to speed-up its development for a deployment late 2015. As a consequence, we have been contacted already by institutions and companies to use the platform for evaluation and hence we plan to monetize the platform by providing commercial licenses.

Finally, all these activities related to the center attracted the attention of large companies and led to a license agreement with Safran Morpho and a research agreement with Google.

10. Natural Language Processing

Key scientific outputs: Firstly, a PhD thesis completed in 2014 has demonstrated that semantic analyses of discourse connectives and verb tenses in multi-sentence texts are beneficial to statistical MT, an innovation that has been received with a lot of interest in the community. Secondly, we have integrated the diversity-preserving algorithms for keywords extraction and document ranking into an improved "content linking system" with a new interface, which recommends Wikipedia pages in real-time to conversation participants. Thirdly, in ground-breaking papers, we have shown that sentiment analysis of user comments improves recommendation, and that it is possible to predict individual sentiment aspects and relate them to text excerpts that support the predictions. Fourthly, snippet recommendation techniques have been successfully used for audio-visual lecture summarization.

---

46 www.beat-eu.org/

Text-Level Machine Translation

We focus on a problem that is not often targeted in the current statistical machine translation (SMT) paradigm: the translation of words that depend on long-range relationships between sentences. The main examples are discourse connectives (which express argumentative relations between clauses), pronouns and noun phrases (which depend on the entities they refer to), and verb tenses (which convey the narrative structure of a text). Connectives and verb tenses have been mainly addressed in the COMTIS SNSF Sinergia project[47] , and the study of their use for SMT has been completed in the first year of the MODERN SNSF Sinergia project[48].

We implemented new specialized features for the disambiguation of discourse connectives, showing that a large part of the performance is gained by the use of syntactic and dependency structures, but that additional improvements are brought by specific features such as temporal information and synonym/antonym detection. We successfully integrated discourse connective labels into SMT to improve the coherence and readability of translations from English into four target languages: French, German, Italian, and Arabic. The labels were annotated automatically over large data sets, and then used to train and test phrase-based factored SMT models. This is, to our knowledge, the first study to improve connective translation, hence text coherence and readability, based on source-side contextual features.

We dealt in a somewhat similar way with verb tenses, which also depend on relations across sentences, and raise numerous challenges when translating from English to French. Using a precise alignment of EN/FR verb phrases, we trained a classifier predicting directly the translated tense, and showed that when combined with SMT, the translation of verb phrases was improved in terms of tense but also lexical choice.

We have shifted our focus towards the translation of noun phrases and pronouns across entire texts. In a first experiment, in collaboration with the U. of Zurich, we enforced the consistent translation of nouns that appear in compounds. An Chinese-to-English example is shown in Figure 22; for German-to- French, an example is the word Nordwand, which translates into face nord, and indicates that another mention of Wand should be translated by face and not by mur. The situation is more frequent in Chinese due to multicharacter words. We proposed a method to identify such cases and post-edit the output of a baseline SMT system, leading to above 15% improvement of translations.

Figure 22: Identifying and post-editing Chinese compounds to improve their translation.

Moreover, we have designed a system to post-edit pronouns in English-to-French translations (it ambiguous between il and elle) using a number of surface features. We also implemented an evaluation metric dedicated to this problem (APT, for Accuracy of Pronoun Translation). The metric and the system will be used for the shared task on pronoun translation proposed at DiscoMT 2015, the second workshop on Discourse and MT, which the NLP group co-organizes at EMNLP 2015.

Diverse Keyword Extraction and Document Ranking for Just-in-Time Document Retrieval

We have continued work on the application framework known as the Automatic Content Linking Device (ACLD), which is a real-time multimedia retrieval system that enriches a conversation with suggestions of potentially relevant documents. We have proposed a method, using a submodular optimization criterion, to extract from a conversation fragment (e.g., most recent 20 seconds) a precise yet diverse set of keywords that reflect the diversity of topics hypothesized in the fragment. These keywords appeared as more relevant than competing methods, and allowed the retrieval of better recommendations. We also proposed a method for re-ranking results using a diversity criterion, with application to implicit but also to explicit queries. For the latter case, we released the AREX dataset[49] which allows objective comparisons between various contextual query-answering methods. These theoretical contributions have been integrated into an end-to-end system, with a new graphical interface shown in Figure 23. The interface allows meeting participants to view the ASR output of each conversation fragment, and to relate the recommendations to words in the transcripts. The interface displays the most recent fragment but can also scroll back in time. – This research is grounded in work on multimodal meeting analysis and meeting browsers in the IM2 Swiss NCCR[50] , and was supported by the Hasler Foundation.

Figure 23: Keyword extraction from a conversation transcript (real-time ASR fragment, top frame) and Wikipedia pages retrieved based on these keywords (bottom frame). When hovering over one result, the corresponding keywords change from green to blue.

Text Analysis for Multimedia Recommendation

The NLP Group develops methods for multimedia information retrieval based on the analysis of content, including sentiment analysis, and taking advantage of the networked structure of social media repositories. This research is supported by the InEvent[51] EU project, by SNSF funding for technology transfer in collaboration with Klewel (a company doing conference recording and broadcasting), and by a CTI/KTI project with Faveeo SA (an information watch company).

First, to study the problem of recommending multimedia networked data, we have built a benchmark dataset from about 1100 TED lectures, including lists of user favorites and ground-truth generic recommendations[52] . We have shown that performing sentiment analysis on user comments improves the accuracy of recommendations, in particular when using machine learning to adjust the sentiment scores to the rating levels.

Figure 24: Examples of weights for keywords extracted in relation to four emotion aspects.

We then defined a model that uses multiple-instance regression (MIR) to predict ratings from user- contributed texts. We proposed a new MIR algorithm which generalizes over the previous ones by assigning individual relevance values to each instance of a bag. The new method supports high-dimensional feature spaces, as required for text regression, and helps to predict both the class label and the content

structure of unlabeled bags, which no previous algorithm can do. The advantages of our method are: learning an optimal method to aggregate instances rather than a fixed one; allowing more degrees of freedom in the regression model; interpreting the weight of an instance as relevance; analyzing the content structure of a bag both at the training and at the prediction stage. The model is used to predict aspect ratings in previously unseen texts or transcripts, demonstrating explanatory power for its predictions. For instance, Figure 24 displays sets of words that have been found to correlate with 4 of the 12 emotions annotated over the TED dataset by its viewers. We evaluated the model on seven publicly available multi-aspect data sets, improving over several existing models: three MIR baselines and two linear models, namely Support Vector Regression and Lasso.

The model using multi-aspect sentiment analysis for recommendation has been implemented as the EMOrec software module for the inEvent project. The module recommends hyper-events based on the similarity of their emotional profile, as inferred automatically from the transcripts. A snapshot of the online demonstrator[53] is shown in Figure 25.
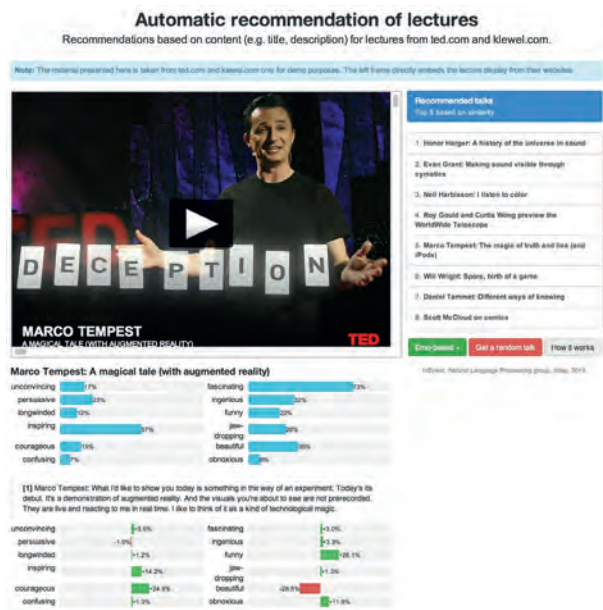

Figure 25: Emotion-based analysis (blue bars under video), paragraph-specific relevance scores

---

Projects funded by European agencies

1. Name   **SODA** (Person Recognition in debate and broadcast news)
    Funding   ANR - Agence Nationale de la Recherche
    Coordinator   Universite du Maine
    Duration   01.11.2011 - 30.09.2014
    Partner(s)   Idiap Research Institute

2. Name   **DBOX** (D-Box: A generic dialog box for multilingual conversational applications)
    Funding   Eurostars Programme
    Coordinator   Mi'pu'mi Games GmbH
    Duration   01.11.2012 - 31.10.2015
    Partner(s)   Universität des Saarlandes, SIKOM Software GmbH, Koemei SA, Idiap Research Institute, Acapela

3. Name   **BEAT** (Biometrics Evaluation and Testing)
    Funding   FP7 - STREP - SEC
    Coordinator   Idiap Research Institute
    Duration   01.03.2012 - 28.02.2016
    Partner(s)   University of Surrey, Universidad Autónoma de Madrid, Tüvit, Tubutak Informatics and Information Security,
    Advanced Technologies Research Center, Sagem Securité S.A., Katholieke Universiteit Leuven, École Polytechnique Fédérale
    de Lausanne, Commissariat à l'Énergie Atomique

4. Name   **EUMSSI** (EUMSSI - Event Understanding through Multimodal Social Stream Interpretation)
    Funding   FP7 - IP - ICT
    Coordinator   Fundacio Barcelona Media
    Duration   01.12.2013 - 30.11.2016
    Partner(s)   Video Stream Networks S.L., Université du Maine, Gottfried Wilhelm Leibniz Universität Hannover, Idiap Research Institute,
    Gesellschaft zur Förderung der angewandten Informationsforschung EV, Deutsche Welle

5. Name   **INEVENT** (Accessing Dynamic Networked Multimedia Events)
    Funding   FP7 - STREP - ICT
    Coordinator   Idiap Research Institute
    Duration   01.11.2011 - 31.10.2014
    Partner(s)   University of Edinburgh, Radvision LTD, Klewel, IBM Israel LTD, Fraunhofer-Gesellschaft zur Förderung der angewandten
    Forschung e.V.

6. Name   **ROCKIT** (Roadmap for Conversational Interaction Technologies)
    Funding   FP7 - CSA - ICT
    Coordinator   University of Edinburgh
    Duration   01.12.2013 - 30.11.2015
    Partner(s)   Vodera Limited, Universität des Saarlandes, LT-Innovate, Idiap Research Institute, Europe Limited SA

7. Name   **SIIP** (Speaker Identification Integrated Project)
    Funding   FP7 - SRTEP - SEC
    Coordinator   Verint System Ltd
    Duration   01.05.2014 - 30.04.2018
    Partner(s)   Sail Labs Technology AG, Rijksuniversiteit Groningen, Police Service of Northern Ireland, Ministério da Justiça, Lisboa, ok2go,
    Laboratorio Di Scienze della Cittadinanza, Lowuendo SpA, The international Criminal Police Ogranization, Inov Inesc Inovacao
    - Instituto de novas Tecnologias, Idiap Research Institute, International Biometrics Group (UK) Limited, Green Fusion Limited
    (Data Fusion International) Cassidian SAS, University of Warwick, Synthema S.R.L, Singularlogic Anonymos Etairia Pliriofiorakon
    Systimaton & Efarmogon Pliroforikis

8.   Name   **SSPNET** (Social Signal Processing Network)
      Funding   FP7 - NoE - ICT
      Coordinator   Idiap Research Institute
      Duration   01.02.2009 - 31.01.2014
      Partner(s)   Universiteit Twente, Universita di Roma Tre, University of Geneva, University of Edinburgh, Delft University of Technology, Queen's University Belfast, Institut de Recherche en Informatique et en Automatique, Imperial College London, German Research Centre for Artificial Intelligence Stefan Institute, Idiap Research Institute, Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., Stichting Centrum voor Wiskunde en Informatica, Centre national de la recherche scientifique, Xerox SAS, Universite Pierre et Marie Curie-Paris 6, University of Bristol, Università degli Studi di Milano, University of Manchester, Helsingin yliopisto, University of Edinburgh

9.   Name   **TABULARASA** (Trusted Biometrics under Spoofing Attacks)
      Funding   FP7 - STREP - ICT
      Duration   01.11.2010 - 30.06.2014
      Coordinator   Idiap Research Institute
      Partner(s)   University of Southampton, University of Oulu, Universita Degli Studi Di Cagliari, Universidad Autónoma de Madrid, Startlab Barcelona SL, Sagem Sécurté S.A., KeyLemon SA, Eurocom, Biometry.com AG, Centre for Science, Society and Citizenship, Institute of Automation Chinese Academy of Sciences

Projects funded by Swiss agencies

1. Name   **MCM-FF** (Multimodal Computational Modeling of Nonverbal Social Behavior in Face to Face Interaction)
   Funding   SNSF - Ambizione
   Coordinator   Idiap Research Institute
   Duration   01.11.2011 - 28.02.2015
   Partner(s)

2. Name   **MAAYA** (Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy)
   Funding   SNSF - DACH
   Coordinator   Idiap Research Institute
   Duration   01.06.2013 - 31.05.2016
   Partner(s)   University of Geneva, University of Bonn

3. Name   **SIVI** (Situated Vision to Perceive Object Shape and Affordances)
   Funding   SNSF - DACH
   Coordinator   University of Bonn
   Duration   01.11.2011 - 30.11.2015
   Partner(s)   Technische Universitaet Wien, RWTH Aachen, Idiap Research Institute

4. Name   **A-MUSE** (Adaptive Multilingual Speech Processing)
   Funding   SNSF - Division II
   Coordinator   Idiap Research Institute
   Duration   01.10.2012 - 30.09.2015
   Partner(s)

5. Name   **DASH** (Object Detection with Active Sample Harvesting)
   Funding   SNSF - Division II
   Coordinator   Idiap Research Institute
   Duration   01.09.2012 - 31.08.2015
   Partner(s)

6. Name   **FLEXASR** (Flexible Grapheme-Based Automatic Speech Recognition)
   Funding   SNSF - Division II
   Coordinator   Idiap Research Institute
   Duration   01.05.2013 - 30.04.2014
   Partner(s)

7. Name   **G3E** (G3E: Geometric Generative Gaze Estimation model)
   Funding   SNSF - Division II
   Coordinator   Idiap Research Institute
   Duration   01.04.2014 - 31.01.2015
   Partner(s)

8. Name   **HFACE** (Heterogeneous Face Recognition)
   Funding   SNSF - Division II
   Coordinator   Idiap Research Institute
   Duration   01.07.2014 - 30.06.2017
   Partner(s)

9. Name   **PHASER** (PHASER: Parsimonious Hierarchical Automatic Speech Recognition)
   Funding   SNSF - Division II
   Coordinator   Idiap Research Institute
   Duration   01.06.2014 - 31.05.2016
   Partner(s)

10. Name   **RODI** (Role based speaker diarization)
    Funding   SNSF - Division II
    Coordinator   Idiap Research Institute
    Duration   01.11.2011 - 31.10.2014
    Partner(s)

11. Name **TRACOME** (Robust face tracking, feature extraction and multimodal fusion for audio-visual speech recognition)
    Funding   SNSF - Division II
    Coordinator École Polytechnique Fédérale de Lausanne
    Duration   01.01.2011 - 31.03.2014
    Partner(s)   Idiap Research Institute
12. Name **YOUTH@NIGHT** (Youth@Night – A multi-disciplinary multi-method study of young people's outgoing and drinking behaviors)
    Funding   SNSF - Interdisciplinary Project
    Coordinator Sucht Schweiz - Research Institute
    Duration   01.01.2014 - 31.12.2016
    Partner(s)   Universität Zürich, Idiap Research Institute

13. Name **DIMHA** (Diarizing Massive Amounts of Heterogeneous Audio)
    Funding   SNSF - Plan de Relance
    Coordinator Idiap Research Institute
    Duration   01.06.2012 - 30.09.2014
    Partner(s)   Koemei SA

14. Name **LOBI** (Low Complexity Binary Features for Robust-to-Noise Speaker Recognition)
    Funding   SNSF - Plan de Relance
    Coordinator Idiap Research Institute
    Duration   01.06.2012 - 30.09.2014
    Partner(s)   KeyLemon SA

15. Name **SP2** (SCOPES Project on Speech Prosody)
    Funding   SNSF - SCOPES
    Coordinator Idiap Research Institute
    Duration   01.04.2014 - 31.03.2016
    Partner(s)   Universizy of Novi Sad, Ss. Cyril and Methodius University, Budapest University of Technology and Economy

16. Name **MODERN** (Modeling discourse entities and relations for coherent machine translation)
    Funding   SNSF - Sinergia
    Coordinator Idiap Research Institute
    Duration   01.08.2013 - 31.07.2016
    Partner(s)   Universiteit Utrecht, Universität Zürich, University of Geneva

17. Name **NINAPRO** (Non-Invasive Adaptive Hand Prosthetics)
    Funding   SNSF - Sinergia
    Coordinator Idiap Research Institute
    Duration   01.01.2011 - 30.06.2014
    Partner(s)   HES-SO Valais, German Aerospace Research Center

18. Name **SIWIS** (Spoken Interaction with Interpretation in Switzerland)
    Funding   SNSF - Sinergia
    Coordinator Idiap Research Institute
    Duration   01.12.2012 - 30.11.2015
    Partner(s)   University of Geneva, University of Edinburgh, Eidgenössische Technische Hochschule Zürich

19. Name **SONVB** (Sensing and Analysing Organizational Nonverbal Behavior)
    Funding   SNSF - Sinergia
    Coordinator Idiap Research Institute
    Duration   01.06.2010 - 30.06.2014
    Partner(s)   Université de Neuchâtel, The Trustees of Dartmouth College Corp.

20. Name **UBIMPRESSED** (UBImpressed: Ubiquitous First Impressions and Ubiquitous Awareness)
    Funding   SNSF - Sinergia
    Coordinator Idiap Research Institute
    Duration   01.01.2014 - 31.12.2016
    Partner(s)   Université de Neuchâtel, Cornell University

21. Name **WILDTRACK** (Tracking in the Wild)
    Funding   SNSF - Sinergia
    Coordinator Idiap Research Institute
    Duration   01.01.2014 - 31.12.2016
    Partner(s)   Eidgenössische Technische Hochschule Zürich, École Polytechnique Fédérale de Lausanne

22. Name   **AROLES** (Automatic Recommendation of Lectures and Snipets)
    Funding   SNSF - Swiss National Science Foundation
    Coordinator Idiap Research Institute
    Duration   01.06.2012 - 30.09.2014
    Partner(s) KLEWEL

23. Name   **ADDG2SU** (Flexible Acoustic Data-Driven Grapheme to Subword Unit Conversion)
    Funding   Hasler Foundation
    Coordinator Idiap Research Institute
    Duration   01.03.2013 - 01.03.2016
    Partner(s)

24. Name   **COHFACE** (COntactless Heartbeat detection for trustworthy FACE Biometrics)
    Funding   Hasler Foundation
    Coordinator Idiap Research Institute
    Duration   01.09.2014 - 31.08.2015
    Partner(s)

25. Name   **DEEPSTD** (Universal Spoken Term Detection with Deep Learning)
    Funding   Hasler Foundation
    Coordinator Idiap Research Institute
    Duration   01.12.2011 - 30.11.2014
    Partner(s)

26. Name   **DEEPSTD-EXT** (Universal Spoken Term Detection with Deep Learning (extension))
    Funding   Hasler Foundation
    Coordinator Idiap Research Institute
    Duration   01.12.2014 - 30.09.2015
    Partner(s)

27. Name   **GENEEMO** (Geneemo: An Expressive Audio Content Generation Tool)
    Funding   Hasler Foundation
    Coordinator Idiap Research Institute
    Duration   01.04.2014 - 30.09.2015
    Partner(s)

28. Name   **IMAGECLEF** (The Robot Vision Task @ ImageCLEF: Towards Web- Robotics)
    Funding   Hasler Foundation
    Coordinator Idiap Research Institute
    Duration   01.01.2013 - 28.02.2014
    Partner(s)

29. Name   **MASH-2** (Massive Sets of Heuristics for Machine Learning II)
    Funding   Hasler Foundation
    Coordinator Idiap Research Institute
    Duration   01.07.2013 - 30.06.2016
    Partner(s)

30. Name   **REMUS** (REMUS: Re-ranking Multiple Search Results for Just-in-Time Document Recommendation)
    Funding   Hasler Foundation
    Coordinator Idiap Research Institute
    Duration   01.01.2014 - 31.10.2014
    Partner(s)

31. Name   **SESAME** (SEarching Swiss Audio MEmories)
    Funding   Hasler Foundation
    Coordinator Idiap Research Institute
    Duration   01.01.2011 - 28.02.2015
    Partner(s)

32. Name   **SODS** (Semantically Self-Organized Distributed Web Search)
    Funding   Hasler Foundation
    Coordinator Idiap Research Institute
    Duration   01.01.2012 - 01.03.2015
    Partner(s)

33. Name   **OMSI ARMASUISSE** (Objective Measurement of Speech Intelligibility)
    Funding   Armasuisse
    Coordinator Idiap Research Institute
    Duration   01.07.2014 - 30.06.2015
    Partner(s)

34. Name   **SENSECITYVITY** (Mobile Sensing, Urban Awareness, and Collective Action)
    Funding   EPFL - Coopération
    Coordinator Idiap Research Institute
    Duration   01.03.2014 - 28.02.2015
    Partner(s)   Instituto Potosino de Investigacion Cientifica y Tecnologica

35. Name   **VALAIS 2015** (Valais*Wallis Digital)
    Funding   Etat du Valais
    Coordinator Idiap Research Institute
    Duration   01.04.2013 - 31.12.2015
    Partner(s)   Valais/Wallis Promotion, Migros Engagement, Société Coopérative Migros Valais,
    Médiathèque Valais et Archives de l'Etat, Ecole professionnelle des arts contemporains

36. Name   **DEMO-NAO** (Demonstrateur NAO)
    Funding   Loterie Romande
    Coordinator Idiap Research Institute
    Duration   01.09.2014 - 31.08.2015
    Partner(s)

Projects funded by industrial partners

1. Name **ARIS** (Automatic French and German Speech Recognition System for Efficient Indexing and Search)
   Funding Fondation The Ark
   Coordinator Idiap Research Institute
   Duration 01.04.2013-31.01.2014
   Partner(s) Koemei, Klewel

2. Name **DIGI-BOARD** (DIGI-Board (Digit-Arena))
   Funding Fondation The Ark
   Coordinator Idiap Research Institute
   Duration 01.04.2014 - 30.09.2014
   Partner(s) E.S. Concept

3. Name **FAR** (Face Alignment based on RGB-D Cameras) Fondation The Ark
   Funding KeyLemon SA
   Coordinator Idiap Research Institute
   Duration 01.07.2014 - 28.02.2015
   Partner(s)

4. Name **LAMI** (Large scale multi-energy information platform)
   Funding Fondation The Ark
   Coordinator Centre de recherches energetiques et municipales
   Duration 01.08.2013 - 28.02.2014
   Partner(s) ESR - Energie Sion Region, Sinergy, Idiap Research Institute, HES-SO Valais

5. Name **SMARTCAP** (Smart Caps Orientation)
   Funding Fondation The Ark
   Coordinator AISA Automation Industrielle SA
   Duration 01.05.2013 - 21.02.2014
   Partner(s) Idiap Research Institute

6. Name **SUVA** (Recomed: Intégration de la transcription vocale dans le dossier
   Funding patient informatisé CRR) Fondation The Ark
   Coordinator Idiap Research Institute
   Duration 01.03.2014 - 31.03.2015
   Partner(s) Clinique romande de réadaptation, Cimark

7. Name **DIGIT-ARENA** (REAL-TIME PERIMETER BOARD CONTENT DIGITAL REPLACEMENT)
   Funding CTI
   Coordinator Idiap Research Institute
   Duration 01.12.2014 - 30.07.2016
   Partner(s) E.S. Concept

8. Name **DOMOCARE** (DomoCare - A new Home Care Preventive Protocol)
   Funding CTI
   Coordinator Idiap Research Institute
   Duration 01.10.2013 - 30.09.2014
   Partner(s) DomoSafety S.A.

9. Name **FAVEO** (Accelerating online information discovery through context-driven and behaviour-based personalization of search)
   Funding CTI
   Coordinator Idiap Research Institute
   Duration 01.09.2014 - 31.03.2015
   Partner(s) Faveeo, École Polytechnique Fédérale de Lausanne

10. Name **FEDARS** (Feature Extraction from Deep learning Architectures for face Recognition Systems)
    Funding CTI
    Coordinator Idiap Research Institute
    Duration 01.04.2013 - 30.09.2014
    Partner(s) KeyLemon SA

11. Name **MULTIVEO** (High Accuracy Speaker-Independent Multilingual Automatic Speech Recognition System)
    Funding CTI
    Coordinator Idiap Research Institute
    Duration 01.11.2013 - 31.05.2015
    Partner(s) Veovox

12. Name   **SCOREL2** (Automatic scoring and adaptive pedagogy for oral language learning)
    Funding   CTI
    Coordinator Idiap Research Institute
    Duration   01.06.2014 - 31.01.2016
    Partner(s) SpeedLingua

13. Name   **VIDEOPROTECTOR** (Morphean VideoProtector)
    Funding   CTI
    Coordinator Idiap Research Institute
    Duration   01.06.2014 - 31.12.2015
    Partner(s) HES-SO Fribourg

14. Name   **AAMASSE** (Acoustic Model Adaptation toward Spontaneous Speech and Environment)
    Funding   IND - Industrial project
    Coordinator Idiap Research Institute
    Duration   01.05.2014 - 31.12.2014
    Partner(s) Samsung Electronics Co

15. Name   **EMMA1** (Expression Mimics Marker Analysis)
    Funding   IND - Industrial project
    Coordinator Idiap Research Institute
    Duration   01.03.2014 - 28.02.2015
    Partner(s) DoctorWeb

16. Name   **NEC** (NEC collaboration)
    Funding   IND - Industrial project
    Coordinator Idiap Research Institute
    Duration   01.05.2012 - 30.04.2014
    Partner(s) NEC

17. Name   **NTT** (NISHA - NTT Idiap Social Behaviour Analysis Initiative)
    Funding   IND - Industrial project
    Coordinator Idiap Research Institute
    Duration   15.11.2012 - 30.09.2014
    Partner(s) NTT

Projects awarded in 2014 and starting in 2015

1. Name   **BIOWATCH** (Biowatch)
   Funding   Fondation The Ark
   Coordinator Idiap Research Institute
   Duration   01.01.2015 - 30.04.2015
   Partner(s) Biowatch SA

2. Name   **DEXROV** (Effective Dexterous ROV Operations in Presence of Communications Latencies)
   Funding   H2020 - RIA - BG
   Coordinator Space Applications Services
   Duration   01.03.2015 - 31.08.2018
   Partner(s) Interuniversity Center of Integrated Systems for the Marine Environment, Jacobs University Bremen,
               Idiap Research Institute, Graal Tech s.r.l.,COMEX SA.

3. Name   **RECAPP** (Making speech technology accessible to Swiss people)
   Funding   Hasler Foundation
   Coordinator Idiap Research Institute
   Duration   01.01.2015 - 30.09.2016
   Partner(s)

4. Name   **ATAP** (ATAP - Continuous Authentication Program) Funding   IND - Industrial Project
   Coordinator Idiap Research Institute
   Duration   15.01.2015 - 14.04.2015
   Partner(s) Google Laboratorio Di Scienze Della Cittadinanza

Book / Book Chapters

1. Ivana Chingovska, André Anjos and Sébastien Marcel, Anti-spoofing: Evaluation Methodologies, in: Encyclopedia of Biometrics, Springer US, 2014

2. André Anjos, Ivana Chingovska and Sébastien Marcel, Anti-Spoofing: Face Databases, in: Encyclopedia of Biometrics, Springer US, 2014

3. Sandrine Zufferey and Andrei Popescu-Belis, Discourse connectives: theoretical models and empirical validations in humans and computers, in: Papers dedicated to Jacques Moeschler, University of Geneva, 2014

4. Stan Z.Li, Javier Galbally, André Anjos and Sébastien Marcel, Evaluation Databases, in: Handbook of Biometric Anti-Spoofing, pages 247-278, Springer-Verlag, 2014

5. Ivana Chingovska, André Anjos and Sébastien Marcel, Evaluation Methodologies, in: Handbook of Biometric Antispoofing, Springer, 2014

6. André Anjos, Jukka Komulainen, Sébastien Marcel, Abdenour Hadid and Matti Pietikainen, Face Anti-spoofing: Visual Approach, in: Handbook of Biometric Anti-Spoofing, pages 65-82, SpringerVerlag, 2014

7. Francois Fleuret, Horesh Ben Shitrit and Pascal Fua, Re-Identification for Improved People Tracking, in: Person Re-Identification, pages 311-336, Springer, 2014

8. Daniel Gatica-Perez, Oya Aran and Dinesh Babu Jayagopi, Understanding small groups, in: Social Signal Processing, Cambridge University Press. Editors J. Burgoon A. Vinciarelli, M. Pantic and N. Magnenat-Thalmann., 2014

Articles in Scientific Journals

1. Trinh-Minh-Tri Do, O. Dousse, Markus Miettinen and Daniel Gatica-Perez, A Probabilistic Kernel Method for Human Mobility Prediction with Smartphones, in: Pervasive and Mobile Computing, 2014

2. Alessandro Vinciarelli and Gelareh Mohammadi, A Survey of Personality Computing, in: IEEE Transaction on Affective Computing, 2014 Mohammad J. Taghizadeh, Reza Parhizkar, Philip N. Garner, Hervé Bourlard and Afsaneh Asaei, Ad Hoc Microphone Array Calibration: Euclidean Distance Matrix Completion Algorithm and Theoretical Guarantees, in: Signal Processing, 2014

3. Charles Dubout and Francois Fleuret, Adaptive Sampling for Large Scale Boosting, in: Journal of Machine Learning Research, 15:1431-1453, 2014

4. Ivana Chingovska, André Anjos and Sébastien Marcel, Biometrics Evaluation Under Spoofing Attacks, in: IEEE Transactions on Information Forensics and Security, 9(12):2264-2276, 2014

5. Oya Aran, Joan-Isaac Biel and Daniel Gatica-Perez, Broadcasting oneself: Visual Discovery of Vlogging Styles, in: IEEE Transactions on Multimedia, 16(1):201-215, 2014

6. Manfredo Atzori, Arjan Gijsberts, Ilja Kuzborskij, Simone Heynen, Anne-Gabrielle Mittaz Hager, Olivier Deriaz, Claudio Castellini, Henning Müller and Barbara Caputo, Characterization of a Benchmark Database for Myoelectric Movement Classification, in: Transactions on Neural Systems and Rehabilitation Engineering, 2014

7. Nikolaos Pappas and Andrei Popescu-Belis, Combining Content with User Preferences for NonFiction Multimedia Recommendation: A Study on TED Lectures, in: Multimedia Tools and Applications, Special Issue on Content Based Multimedia Indexing:1–23, 2014

8. Lakshmi Saheer, Junichi Yamagishi, Philip N. Garner and John Dines, Combining Vocal Tract Length Normalization with Hierarchical Linear Transformations, in: IEEE Journal of Selected Topics in Signal Processing Special Issue on Statistical Parametric Speech Synthesis, 8(2):262 - 272, 2014

9. Denise Frauendorfer, Marianne Schmid Mast, Dairazalia Sanchez-Cortes and Daniel Gatica- Perez, Emergent Power Hierarchies and Group Performance, in: International Journal of Psychology, 2014

10. Mohammad J. Taghizadeh, Philip N. Garner and Hervé Bourlard, Enhanced Diffuse Field Model for Ad Hoc Microphone Array Calibration, in: Signal Processing, 101:242-255, 2014

11. Alexandre Heili, Adolfo Lopez-Mendez and Jean-Marc Odobez, Exploiting Long-Term Connectivity and Visual Motion in CRF-based Multi-Person Tracking, in: Transactions on Image Processing, 2014

12. Tiago de Freitas Pereira, Jukka Komulainen, André Anjos, José Mario De Martino, Abdenour Hadid, Matti Pietikainen and Sébastien Marcel, Face liveness detection using dynamic texture, in: EURASIP Journal on Image and Video Processing, 2014:2, 2014

13. Laurent Son Nguyen, Denise Frauendorfer, Marianne Schmid Mast and Daniel Gatica-Perez, Hire Me: Computational inference of hirability in employment interviews based on nonverbal behavior, in: IEEE Transactions on Multimedia, 2014

14. Maryam Habibi and Andrei Popescu-Belis, Keyword Extraction and Clustering for Document Recommendation in Conversations, in: IEEE/ACM Transactions on Audio Speech and Language Processing, 2014

15. M. S. Malekzadeh, S. Calinon, D. Bruno and D. G. Caldwell, Learning by Imitation with the STIFF- FLOP Surgical Robot: A Biomimetic Approach Inspired by Octopus Movements, in: Robotics and Biomimetics, 1(13):1-15, 2014

16. Stefan Duffner and Jean-Marc Odobez, Leveraging Colour Segmentation for Upper-Body Detection, in: Pattern Recognition, 47(6):2222-2230, 2014

17. Joan-Isaac Biel and Daniel Gatica-Perez, Mining Crowdsourced First Impressions in Online Social Video, in: IEEE Transactions on Multimedia, 16(7), 2014

18. Srikanth Madikeri, Asha T and Hema A Murthy, MODIFIED GROUP DELAY FEATURE BASED TOTAL VARIABILITY SPACE MODELLING FOR SPEAKER RECOGNITION, in: International Journal of Speech Techonology, 2014

19. Denise Frauendorfer, Marianne Schmid Mast, Laurent Son Nguyen and Daniel Gatica-Perez, Nonverbal Social Sensing in Action: Unobtrusive Recording and Extracting of Nonverbal Behavior in Social Interactions Illustrated with a Research Example, in: Journal of Nonverbal Behavior, 38(2):231-245, 2014

20. Sree Harsha Yella and Hervé Bourlard, Overlapping speech detection using long-term conversational features for speaker diarization in meeting room conversations., in: Audio, Speech and Language processing, IEEE/ACM Transaction on, 22(12):1688-1700, 2014

21. Miranti I. Mandasari, Manuel Günther, Roy Wallace, Rahim Saedi, Sébastien Marcel and David Van Leeuwen, Score Calibration in Face Recognition, in: IET Biometrics:1-11, 2014

22. Nesli Erdogmus and Sébastien Marcel, Spoofing Face Recognition with 3D Masks, in: IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, 2014

23. Arjan Gijsberts, Rashida Bohra, David Sierra González, Alexander Werner, Markus Nowak, Barbara Caputo, Maximo A. Roa and Claudio Castellini, Stable Myoelectric Control of a Hand Prosthesis using Non-Linear Incremental Learning, in: Frontiers in Neurorobotics, 2014

24. Afsaneh Asaei, Mohammad Golbabaee, Hervé Bourlard and Volkan Cevher, Structured Sparsity Models for Reverberant Speech Separation, in: IEEE/ACM Transaction on Audio, Speech and Language Processing, 2014

25. Remi Emonet, Jagannadan Varadarajan and Jean-Marc Odobez, Temporal Analysis of Motif Mixtures using Dirichlet Processes, in: IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), 36(1), 2014

26. Arjan Gijsberts, Manfredo Atzori, Claudio Castellini, Henning Müller and Barbara Caputo, The Movement Error Rate for Evaluation of Machine Learning Methods for sEMG-based Hand Movement Classification, in: Transactions on Neural Systems and Rehabilitation Engineering, 2014

27. Lucia Teijeiro-Mosquera, Joan-Isaac Biel, Jose Luis Alba-Castro and Daniel Gatica-Perez, What Your Face Vlogs About: Expressions of Emotion and Big-Five Traits Impressions in YouTube, in: IEEE Transactions Affective Computing, 2014

Articles in Conference Proceedings

1. Kenneth Alberto Funes Mora and Jean-Marc Odobez, 3D Gaze Tracking and Automatic Gaze Coding from RGB-D Cameras, in: IEEE Conference in Computer Vision and Pattern Recognition, Vision Meets Cognition Workshop, Columbus, Ohio, USA, 2014

2. Paul Gay, Elie Khoury, Sylvain Meignier, Jean-Marc Odobez and Paul Deleglise, A Conditional Random field approach for audio-visual people diarization, in: Proceedings of the 39th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) , 2014

3. M. S. Malekzadeh, S. Calinon, D. Bruno and D. G. Caldwell, A Skill Transfer Approach for Continuum Robots - Imitation of Octopus Reaching Motion with the STIFF-FLOP Robot, in: In Proc. of the AAAI Symp. on Knowledge, Skill, and Behavior Transfer in Autonomous Robots, Arlington, VA, USA, pages 49-52, 2014

4. S. Calinon, D. Bruno and D. G. Caldwell, A task-parameterized probabilistic model with minimal intervention control, in: Proc. IEEE Intl Conf. on Robotics and Automation (ICRA) , 2014

5. Mohammad J. Taghizadeh, Afsaneh Asaei, Philip N. Garner and Hervé Bourlard, Ad-Hoc Microphone Array Calibration from Partial Distance Measurements, in: Proceeding of 4th Joint Workshop on Hands-free Speech Communication and Microphone Arrays, 2014

6. Mohammad J. Taghizadeh, Afsaneh Asaei, Philip N. Garner and Hervé Bourlard, Ad-Hoc Microphone Array Calibration from Partial Distance Measurements, in: Proceedings of the 4th Joint Workshop on Hands-free speech communication and Microphone Arrays, 2014

7. Sree Harsha Yella, Andreas Stolcke and Malcolm Slaney, Artificial neural network features for speaker diarization, in: IEEE Spoken Language Technology workshop, South Lake Tahoe, USA, 2014

8. Laurent El Shafey, Elie Khoury and Sébastien Marcel, Audio-Visual Gender Recognition in Uncontrolled Environment Using Variability Modeling Techniques, in: International Joint Conference on Biometrics, Clearwater, Florida, USA, 2014

9. Adolfo Lopez-Mendez, C. E. I Westling, Remi Emonet, M. Easteal, L. Lavia, H. J. Witchel and Jean-Marc Odobez, Automated Bobbing and Phase Analysis to Measure Walking Entrainment, in: IEEE International Conference on Image Processing (ICIP) , Paris, 2014

10. Alvaro Marcos-Ramiro, Daniel Pizarro-Perez, Marta Marron-Romera and Daniel Gatica-Perez, Automatic Blinking Detection towards Stress Discovery, in: Proc. ACM Int. Conf. on Multimodal Interaction, Istanbul, 2014

11. Rui Hu, Carlos Pallan, guido krempel, Jean-Marc Odobez and Daniel Gatica-Perez, Automatic Maya Hieroglyph Retrieval Using Shape and Context Information, in: ACM MM, pages 4, 2014

12. Philip N. Garner, David Imseng and Thomas Meyer, Automatic Speech Recognition and Translation of a Swiss German Dialect: Walliserdeutsch, in: Proceedings of Interspeech, 2014

13. Alvaro Marcos-Ramiro, Daniel Pizarro-Perez, Marta Marron-Romera and Daniel Gatica-Perez, Capturing Upper Body Motion in Conversation: an Appearance Quasi-Invariant Approach, in: Proc. ACM Int. Conf. on Multimodal Interaction, Istanbul, 2014

14. Paul Gay, Gregor Dupuy, Jean-Marc Odobez, Sylvain Meignier and Paul Deleglise, Comparison of Two Methods for Unsupervised Person Identification in TV Shows, in: 12th International Workshop on Content-Based Multimedia Indexing, 2014

15. Matthias Vanoni, Pedro Tome, Laurent El Shafey and Sébastien Marcel, Cross-Database Evaluation With an Open Finger Vein Sensor, in: IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BioMS) , Rome, Italy, 2014

16. Cristina Grisot and Thomas Meyer, Cross-linguistic annotation of narrativity for English/French verb tense disambiguation, in: 9th Edition of the Language Resources and Evaluation Conference, Reykjavik, Iceland, 2014

17. Pranay Dighe, Marc Ferras and Hervé Bourlard, Detecting and Labeling Speakers on Overlapping Speech using Vector Taylor Series, in: INTERSPEECH, 2014

18. A. Sapru and Hervé Bourlard, Detecting speaker roles and topic changes in multiparty conversations using latent topic models, in: Proceedings of Interspeech, 2014

19. Petr Motlicek, David Imseng, Milos Cernak and Namhoon Kim, Development of Bilingual ASR System for MediaParl Corpus, in: Proceedings of the 15th Annual Conference of the International Speech Communication Association (Interspeech 2014) , Singapore, ISCA, 2014

20. Hamid Behravan, Ville Hautamäki, Sabato Marco Siniscalchi, Elie Khoury, Tommi Kurki, Tomi Kinnunen and Chin-Hui Lee, Dialect Levelling in Finnish: A Universal Speech Attribute Approach, in: The 15th Annual Conference of the International Speech Communication Association, 2014

21. Marc Ferras, Stefano Masneri, Oliver Schreer and Hervé Bourlard, Diarizing Large Corpora using Multi-modal Speaker Linking, in: INTERSPEECH 2014, 2014

22. Leonidas Lefakis and Francois Fleuret, Dynamic Programming Boosting for Discriminative MacroAction Discovery, in: International Conference on Machine Learning, 2014

23. Umut Avci and Oya Aran, Effect of nonverbal behavioral patterns on the performance of small groups, in: ICMI Workshop on Understanding and Modeling Multiparty Multimodal Interactions, Istanbul, Turkey, 2014

24. Olivier Canévet and Francois Fleuret, Efficient Sample Mining for Object Detection, in: Proceedings of the 6th Asian Conference on Machine Learning (ACML) , Nha Trang, Vietnam, 2014

25. Maryam Habibi and Andrei Popescu-Belis, Enforcing Topic Diversity in a Document Recommender for Conversations, in: Proceedings of the Coling 2014 (25th International Conference on Computational Linguistics) , Dublin, Ireland, 2014

26. Sharid Loaiciga, Thomas Meyer and Andrei Popescu-Belis, English-French Verb Phrase Alignment in Europarl for Tense Translation Modeling, in: The Ninth Language Resources and Evaluation Conference, Reykjavik, Iceland, 2014

27. Nikolaos Pappas and Andrei Popescu-Belis, Explaining the Stars: Weighted Multiple-Instance Learning for Aspect-Based Sentiment Analysis, in: Conference on Empirical Methods in Natural Language Processing, Doha, Qatar, 2014

28. Adolfo Lopez-Mendez, Florent Monay and Jean-Marc Odobez, Exploiting Scene Cues for Dropped Object Detection, in: 9th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. , 2014

29. David Imseng, Blaise Potard, Petr Motlicek, Alexandre Nanchen and Hervé Bourlard, Exploiting un-transcribed foreign data for speech recognition in well-resourced languages, in: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, Florence, IT, 2014

30. Kenneth Alberto Funes Mora, Florent Monay and Jean-Marc Odobez, EYEDIAP: A Database for the Development and Evaluation of Gaze Estimation Algorithms from RGB and RGB-D Cameras, in: Proceedings of the ACM Symposium on Eye Tracking Research and Applications, Safety Harbor, Florida, United States of America, ACM, 2014

31. Paul Gay, Elie Khoury, Sylvain Meignier, Jean-Marc Odobez and Paul Deleglise, Face identification from overlaid texts using Local Face Recurrent Patterns and CRF models, in: IEEE International Conference on Image Processing 2014, Paris, IEEE, 2014

32. Asha T, Saranya M S, Karthik Pandia D S, Srikanth Madikeri and Hema A Murthy, Feature Switching in the i-vector Framework for Speaker Verification, in: Proc. of Interspeech 2014, pages 5, 2014

33. Kenneth Alberto Funes Mora and Jean-Marc Odobez, Geometric Generative Gaze Estimation (G3E) for Remote RGB-D Cameras, in: IEEE Computer Vision and Pattern Recognition Conference, Columbus, Ohio,USA, pages 1773-1780, IEEE, 2014

34. Elie Khoury, Laurent El Shafey, Marc Ferras and Sébastien Marcel, Hierarchical speaker clustering methods for the NIST i-vector Challenge, in: Odyssey: The Speaker and Language Recognition Workshop, 2014

35. Aleksandra Cerekovic, Oya Aran and Daniel Gatica-Perez, How Do You Like Your Virtual Agent?: Human-Agent Interaction Experience through Nonverbal Features and Personality Traits, in: Human Behavior Understanding, pages 1-15, Springer, 2014

36. Pierre-Edouard Honnet and Philip N. Garner, Importance of Prosody in Swiss French Accent for Speech Synthesis, in: Nouveaux cahiers de linguistique francaise, 2014

37. Alexandre Heili, Jagannadan Varadarajan, Bernard Ghanem, Narendra Ahuja and Jean-Marc Odobez, Improving Head and Body Pose Estimation through Semi-supervised Manifold Alignment, in: International Conference on Image Processing, 2014

38. A. Sapru, Sree Harsha Yella and Hervé Bourlard, Improving Speaker Diarization using social role information, in: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, 2014

39. Sree Harsha Yella, Xavier Anguera and Jordi Luque, Inferring social relationships in a phone call from a single party's speech, in: ICASSP, 2014

40. Sree Harsha Yella and Hervé Bourlard, Information Bottleneck based Speaker Diarization of Meetings using Non-speech as Side Information, in: ICASSP, 2014

41. Sree Harsha Yella and Hervé Bourlard, Information Bottleneck based Speaker Diarization of Meetings using Non-speech as Side Information, in: ICASSP, 2014

42. Elie Khoury, Tomi Kinnunen, Aleksandr Sizov, Zhizheng Wu and Sébastien Marcel, Introducing I-Vectors for Joint Anti-spoofing and Speaker Verification, in: The 15th Annual Conference of the International Speech Communication Association, 2014

43. Gulcan Can, Jean-Marc Odobez and Daniel Gatica-Perez, Is That a Jaguar? Segmenting Ancient Maya Glyphs via Crowdsourcing, in: Proc. ACM Int. Workshop on Crowdsourcing for Multimedia, Orlando, 2014

44. Dimitri Palaz, Mathew Magimai.-Doss and Ronan Collobert, Joint Phoneme Segmentation Inference and Classification using CRFs, in: Global Conference on Signal and Information Processing, 2014

45. Leonidas Lefakis and Francois Fleuret, Jointly Informative Feature Selection, in: International Conference on Artificial Intelligence and Statistics, 2014

46. D. Bruno, S. Calinon and D. G. Caldwell, Learning adaptive movements from demonstration and self-guided exploration, in: Proc. IEEE Intl Conf. on Development and Learning and on Epigenetic Robotics (ICDL-EpiRob) , Genoa, Italy, pages 160-165, 2014

47. L. Rozo, S. Calinon and D. G. Caldwell, Learning Force and Position Constraints in Human-robot Cooperative Transportation, in: Proc. IEEE Intl Symposium on Robot and Human Interactive Communication (Ro-Man) , Edinburgh, Scotland, UK, pages 619-624, 2014

48. T. Alizadeh, S. Calinon and D. G. Caldwell, Learning from demonstrations with partially observable task parameters, in: Proc. IEEE Intl Conf. on Robotics and Automation (ICRA) , 2014

49. Novi Patricia and Barbara Caputo, Learning to Learn, from Transfer Learning to Domain Adaptation: A Unifying Perspective, in: Proceedings of the Computer Vision and Pattern Recognition, 2014

50. Adrian Penate-Sanchez, Francesc Moreno-Noguer, Juan Andrade-Cetto and Francois Fleuret, LETHA: Learning from High Quality Inputs for 3D Pose Estimation in Low Quality Images, in: Proceedings of the International Conference on 3D vision, 2014

51. Yogesh Singh Rawat, Chidansh A. Bhatt and Mohan S. Kankanhalli, Mode of Teaching Based Segmentation and Annotation of Video Lectures, in: International Workshop on Content-Based Multimedia Indexing, 2014

52. Afsaneh Asaei, Hervé Bourlard, Mohammad J. Taghizadeh and Volkan Cevher, Model-based Sparse Component Analysis for Reverberant Speech Localization, in: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing, 2014

53. Pranay Dighe, Marc Ferras and Hervé Bourlard, Modeling Overlapping Speech using Vector Taylor Series, in: Odyssey: The Speaker and Language Recognition Workshop, 2014

54. Novi Patricia, Tatiana Tommasi and Barbara Caputo, Multi-Source Adaptive Learning for Fast Control of Prosthetics Hand, in: Proceedings of the International Conference on Pattern Recognition, 2014

55. Marc Ferras and Hervé Bourlard, Multi-source Posteriors for Speech Activity Detection on Public Talks, in: INTERSPEECH, 2014

56. Ngoc Thang Vu, David Imseng, Daniel Povey, Petr Motlicek, Tanja Schultz and Hervé Bourlard, Multilingual Deep Neural Network based Acoustic Modeling For Rapid Language Adaptation, in: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, Florence, 2014

57. Chidansh A. Bhatt, Nikolaos Pappas, Maryam Habibi and Andrei Popescu-Belis, Multimodal Reranking of Content-based Recommendations for Hyperlinking Video Snippets, in: ACM International Conference on Multimedia Retrieval, 2014

58. D. Bruno, S. Calinon and D. G. Caldwell, Null space redundancy learning for a flexible surgical robot, in: Proc. IEEE Intl Conf. on Robotics and Automation (ICRA) , 2014

59. Marzieh Razavi, Ramya Rasipuram and Mathew Magimai.-Doss, On Modeling Context-Dependent Clustered States: Comparing HMM/GMM, Hybrid HMM/ANN and KL-HMM Approaches, in: International Conference on Acoustics, Speech, and Signal Processing, 2014

60. Marzieh Razavi and Mathew Magimai.-Doss, On Recognition of Non-Native Speech Using Probabilistic Lexical Model, in: Proceedings of the 15th Annual Conference of the International Speech Communication Association (Interspeech 2014) , 2014

61. Pedro Tome, Matthias Vanoni and Sébastien Marcel, On the Vulnerability of Finger Vein Recognition to Spoofing, in: IEEE International Conference of the Biometrics Special Interest Group (BIOSIG) , Darmstadt, Germay, 2014

62. Barbara Caputo and Novi Patricia, Overview of the ImageCLEF 2014 Domain Adaptation Task, in: ImageCLEF 2014: Overview and analysis of the results, 2014

63. Sree Harsha Yella, Petr Motlicek and Hervé Bourlard, Phoneme Background Model for Information Bottleneck based Speaker Diarization, in: Interspeech 2014, 2014

64. Sree Harsha Yella, Petr Motlicek and Hervé Bourlard, Phoneme Background Model for Information Bottleneck based Speaker Diarization, in: Interspeech, Singapore, 2014

65. Sara Bahaadini, Afsaneh Asaei, David Imseng and Hervé Bourlard, Posterior-based Sparse Representation for Automatic Speech Recognition, in: Proceeding of Interspeech, 2014

66. Pierre-Edouard Honnet, Alexandros Lazaridis, Jean-Philippe Goldman and Philip N. Garner, Prosody in Swiss French Accents: Investigation using Analysis by Synthesis, in: Speech Prosody, 2014

67. Joël Legrand and Ronan Collobert, Reccurent Greedy Parsing with Neural Networks, in: Proceedings of ECML 2014, 2014

68. Pedro H. O. Pinheiro and Ronan Collobert, Recurrent Convolutional Neural Networks for Scene Labeling, in: 31st International Conference on Machine Learning (ICML), 2014

69. Steve Renals, Jean Carletta, K Edwards, Hervé Bourlard, Philip N. Garner, Andrei Popescu-Belis, Dietrich Klakow, A Girenko, Volha Petukhova, P Wacker, A Joscelyne, C Kompis, S Aliwell, W Stevens and Y Sabbah, ROCKIT: Roadmap for Conversational Interaction Technologies, in: Proceedings of the 2014 Workshop on Roadmapping the Future of Multimodal Interaction Research including Business Opportunities and Challenges (RFMIR '14), pages 39-42, ACM, 2014

70. Olivier Canévet, Leonidas Lefakis and Francois Fleuret, Sample Distillation for Object Detection and Image Classification, in: Proceedings of the 6th Asian Conference on Machine Learning (ACML), Nha Trang, Vietnam, 2014

71. Laurent El Shafey and Sébastien Marcel, Scalable Probabilistic Models: Applied to Face Identification in the Wild, in: 8th European Biometrics Research and Industry Awards, European Association for Biometrics, Darmstadt, Germany, 2014

72. Marco Fornoni and Barbara Caputo, Scene Recognition with Naive Bayes Non-linear Learning, in: Proceedings of the 22nd International Conference on Pattern Recognition, IEEE, 2014

73. S. Calinon, Skills Learning in Robots by Interaction with Users and Environment, in: In Proc. of the Intl Conf. on Ubiquitous Robots and Ambient Intelligence (URAI), Kuala Lumpur, Malaysia, pages 161-162, 2014

74. Elie Khoury, Laurent El Shafey and Sébastien Marcel, SPEAR: An open source toolbox for speaker recognition based on Bob, in: Proceedings of the 39th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014

75. Milos Cernak, Alexandros Lazaridis, Philip N. Garner and Petr Motlicek, Stress and Accent Transmission In HMM-Based Syllable-Context Very Low Bit Rate Speech Coding, in: Interspeech, 2014

76. Alexandros Lazaridis, Pierre-Edouard Honnet and Philip N. Garner, SVR vs MLP for Phone Duration Modelling in HMM-based Speech Synthesis, in: Speech Prosody, 2014

77. Alexandros Lazaridis, Elie Khoury, Jean-Philippe Goldman, Mathieu Avanzi, Sébastien Marcel and Philip N. Garner, SWISS FRENCH REGIONAL ACCENT IDENTIFICATION, in: Odyssey: The Speaker and Language Recognition Workshop, 2014

78. Alexandros Lazaridis, Jean-Philippe Goldman, Mathieu Avanzi and Philip N. Garner, Syllable based Regional Swiss French Accent Identification using Prosodic Features, in: Nouveaux cahiers de linguistique francaise, 2014

79. Volha Petukhova, Martin Gropp, Dietrich Klakow, Anna Schmidt, Gregor Eigner, Mario Topf, Stefan Srb, Petr Motlicek, Blaise Potard, John Dines, O. Deroo, Ronny Egeler, Uwe Meinz and Steffen Liersch, The DBOX Corpus Collection of Spoken Human-Human and Human-Machine Dialogues, in: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14), Reykjavik, Iceland, European Language Resources Association (ELRA), 2014

80. Daniel Gatica-Perez, Carlos Pallan Gayol, Stephane Marchand-Maillet, Jean-Marc Odobez, Edgar Roman-Rangel, Guido Krempel and Nikolai Grube, The MAAYA Project: Multimedia Analysis and Access for Documentation and Decipherment of Maya Epigraphy, in: Proc. Digital Humanities Conference, Lausanne, 2014

81. Gyorgy Szaszak, Tamas Gabor Csapo, Philip N. Garner, Branislav Gerazov, Zoran Ivanovski, Geza Nemeth, Balint Toth, Milan Secujski and Vlado Delic, The SP2 SCOPES Project on Speech Prosody, in: DOGS2014 - Digital speech and image processing, 2014

82. Fabio Celli, Bruno Lepri, Joan-Isaac Biel, Giuseppe Riccardi, Daniel Gatica-Perez and Fabio Pianesi, The Workshop on Computational Personality Recognition 2014, in: Proceedings of the ACM International Conference on Multimedia, 2014

83. Salvador Ruiz-Correa, Darshan Santani and Daniel Gatica-Perez, The Young and the City: Crowd- sourcing Urban Awareness in a Developing Country, in: The 1st International Conference on IoT in Urban Space (Urb-IoT 2014), 2014

84. Xinchao Wang, Engin Turetken, Francois Fleuret and Pascal Fua, Tracking Interacting Objects Optimally Using Integer Programming, in: Proceedings of the European Conference on Computer Vision, pages 17-32, 2014

85. Philip N. Garner, Rob Clark, Jean-Philippe Goldman, Pierre-Edouard Honnet, Maria Ivanova, Alexandros Lazaridis, Hui Liang, Beat Pfister, Manuel Sam Ribeiro, Eric Wehrli and Junichi Yamagishi, Translation and Prosody in Swiss Languages, in: Nouveaux cahiers de linguistique francaise, 2014

86. Remi Emonet, E. Oberzaucher and Jean-Marc Odobez, What to Show? Automatic Stream Selection Among Multiple Sensors, in: International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2014

87. Catharine Oertel, Kenneth Alberto Funes Mora, Samira Sheikhi, Jean-Marc Odobez and Joakim Gustafson, Who Will Get the Grant ? A Multimodal Corpus for the Analysis of Conversational Behaviours in Group, in: International Conference on Multimodal Interaction, Understanding and Modeling Multiparty, Multimodal Interactions Workshop, Istanbul, Turkey, ACM, 2014

88. Rémi Lebret and Ronan Collobert, Word Embeddings through Hellinger PCA, in: 14th Conference of the European Chapter of the Association for Computational Linguistics, 2014